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

Florian Knip
University of Mannheim
Mannheim, Germany
fknip@mail.uni-mannheim.de

Christian Bikar
University of Mannheim
Mannheim, Germany
cbikar@mail.uni-mannheim.de

Bernd Pfister
University of Mannheim
Mannheim, Germany
bpfister@mail.uni-mannheim.de

Ansgar Scherp
University of Mannheim
Mannheim, Germany
ansgar@informatik.uni-mannheim.de

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 people use and interact with such applications is fundamental for improving the functionality and the user interface (UI) design. In our two-step field study, we developed mobEx, a mobile app for the exploration of social media data on Android phones. mobEx unifies the data sources of related commercial apps in the market by retrieving information from various providers. We focused on providing suitable UI features for exploring and selecting faceted data. We have collected qualitative user feedback via a questionnaire. We also conducted a quantitative user study in which we evaluated user-generated logging data over a period of three weeks with a group of 18 participants. The goal was to better understand the user needs and how users interact with the application. Our findings include the following: A grid-navigation is the preverable choice when exploring a large spaces of faceted data. Further, the evaluation shows that the timewheel, a novel search feature, can serve as an intuitive way to explore time-dependent resources such as events.

Keywords

field study, mobile exploration, faceted navigation, faceted data

1. INTRODUCTION

Smart phone applications (short: apps) for nearby search of points of interests (POI) have gained huge popularity among smart phone 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 POIs nearby. 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. We developed mobEx, a mobile application for Android phones that retrieves data from different data providers such as Qype, Twitter, LastFM, Eventful, or Google Places and other open web sources such as DBpedia or OpenPOI. This data is being aggregated via an entity-resolution-concept, and enables the user to explore the retrieved information in an intuitive manner. For this reason, mobEx offers unique search features that enable the user to find people, locations, organizations, and events nearby. 'mobEx' is a composition of the two terms 'mobile' and 'exploration' indicating the focus on the characteristics of exploratory search.

Providing a user interface (UI) that enables the users to easily explore this large space of hierarchical data is one of the major challenges when developing such apps. Especially in terms of usability, latest touchscreen technology and various screen sizes on mobile devices offer high potential different realizations options, but at the same time confront developers with many challenges. The main research question is to understand how people use and interact with such applications in order to improve the functionality and the UI design. In case of nearby search scenarios, the application should provide intuitive access to places and events in the area. For the exploration of time-dependent entities (e.g. events or shops with opening hours), mobEx offers the time-wheel, an intuitive search tool that allows the user to filter events according to his or her prevered time. Further, mobEx offers two different approaches for faceted navigation. A classical

[1] `http://qype.com`
[2] `http://www.aroundmeapp.com`
[3] `http://foursquare.com`
[4] `http://wikitude.com`

list-based approach and an alternative grid-based approach. Both navigation approaches have been compared [6] with the result that the grid-based approach requires more clicks and more time, yet, has a higher user satisfaction.

Overall, the goal of our study was to find out, first (1), if the subjects understood the metaphor of the novel time-wheel and how they use it. Second, (2) what type of navigation approach do the users prefer when they can choose between a grid-based or a list-based navigation menu.

We structured the study into two parts, a qualitative and a quantitative one. The qualitative user study was conducted with a beta version of mobEx and served as feedback to improve the UI design. Hence, the evaluation of the qualitative feedback 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 adujusting time.)
After conducting the changes, we conducted a quantitative study with a final version of mobEx, in which we collected and analyzed user-generated logging data of 18 subjects over a period of three weeks.

The remainder of this paper is organized as follows: In the next section, we will introduce mobEx along with its UI features. In Section 2, we will give an overview of the two studies, especially regarding the participants and the overall procedure. In Section 3 and Section 4, the qualitative and quantitative user study will be presented in detail. In Section 5, we examine work that is either related to our evaluation strategy, as well as other applications that focus on faceted search. Finally, the overall conclusion of the experiment is presented in section 6.

2. INVESTIGATED OBJECT - MOBEX

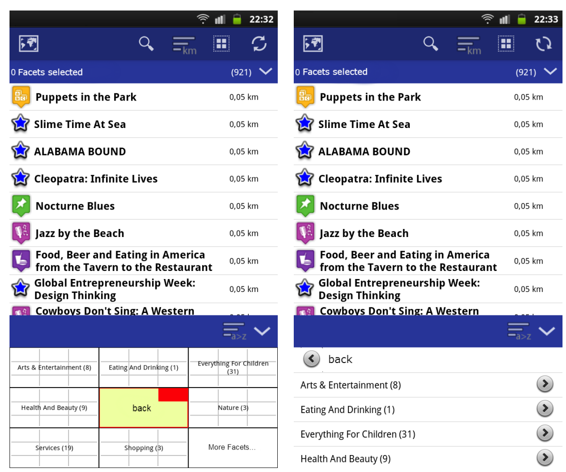
2.1 Scenario

MobEx was designed to address everyday search scenarios while being on the go. The application encourages the user to explore the surrounding area by providing information about people, places, events, and organizations. For instance, the scenario of being in an unknown area or city for a business trip or on vacation looking for upcoming events during the time of the stay. Yet, even in a familiar environment, one might look for a bakery in 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 - certain place, certain time. MobEx addresses these mobile search problems by retrieving data from different web sources according to the users current location and displays the retrieved entities both in a result list and as icons on a map. It offers special UI features such as a (1) time-wheel for the time-dependent exploration and a (2) faceted navigation menu for exploring the retrieved entities along different hierarchically organized categories. The UI can be divided into two major parts, one is the map view where the web resources are visualized on the map, the other one is the facet view where the sources can be filtered and searched according to categories or search terms. The following sections will explain the application and its features in detail.

2.2 Navigation Type

MobEx provides a faceted navigation menu which encourages the user to search in a exploratory manner [5]. Faceted navigation leverages the meta-data fields of the retrieved resources through a hierarchical taxonomy of categories [1].
The mobEx navigation menu is inspired by FaThumb, a search application for large data sets on mobile phones [4]. In contrast to FaThumb, which was designed for physical numerical pads, mobEx makes use of today's touchscreen technology and mobile web access.

MobEx provides two different kinds of navigation menus that both work with the same content, yet each offering a different layout and various navigation functions. Figure 1 shows the two versions of the menu. In both cases, the navigation menu takes the bottom third of space on the touchscreen. Via a simple switch button in the settings menu, the user can choose between the grid menu (Figure 1a) and the list menu (Figure 1b). 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 via non-intruding speech bubbles.

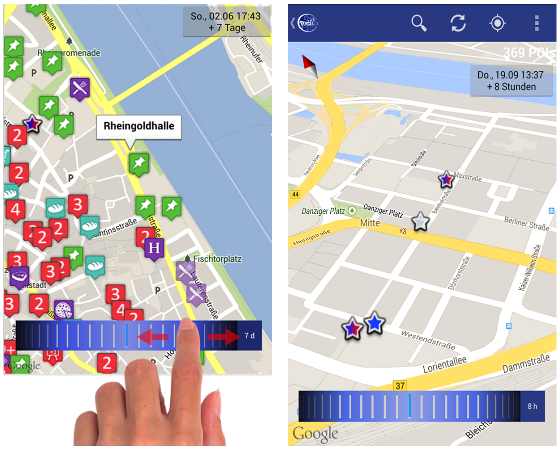


(a) Grid Navigation Menu (b) List Navigation Menu

Figure 1: Comparison of Grid vs. List Navigation

2.2.1 *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 resources in brackets. On the bottom right corner, in the rectangle "other facets", further facets are displayed. The result list can be filtered via long-press on a facet. With short-press on a facet, the user can navigate deeper into the facet structure and explore its sub-facets. To select facets



(a) The time-wheel can be spun with a left or right swipe

(b) The opacity of the icons changes when spinning the time-wheel

Figure 2: The time-wheel can be used to look for events within a specific time frame

without sub-facets a short press is sufficient. Facets can be selected and unselected in an arbitrary combination. By default, all facets are unselected and only the resources from the selected facets show up in the result list. The result list always shows the resources of facets that are selected. When selecting the facet "restaurants", only resources that are under this facet are displayed. Some facets have a light gray grid as background which indicates that these facets have at least one more layer of sub-facets. A characteristic of the grid navigation is the colored rectangles, which are integrated into the back button. Thes rectangles work like a breadcrumb trail and allow user can keep track of the current position within the facet structure. In case of Figure 1a, the red border indicates that the current position is one layer beneath the top layer and the red rectangle on the top right states that on the first level the top right facet was selected.

2.2.2 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 on the right of each entry indicates that the facet has at least one more sub-facet. The back-button is on 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.

The time-wheel allows the user to filter events by time. As the user spins the time-wheel, the time frame changes and the events on the map appear or disappear accordingly, see Figure 2b. Events and other POIs such as shops with opening hours within the chosen time frame are displayed and events/POIs outside of the time frame are transparent. The bigger the time distance the lower the opacity of the icon, see Figure 2a. Turning right shifts the time interval towards the future while turning left shifts it backwards to the present. The user can also adjust the time interval (i.e., the period in which events will be shown) via the switch on the right edge of the wheel widget. Here, the user can choose intervals ranging from single hours up to several weeks. With this interval size, we covered typical use cases such as "Looking for today's nightlife events" or "Planning a two week vacation in New York". Furthermore, the selected time also effects the appearance of icons that represent places where opening hours are available (e.g. shopping centers). The current date and time including the current duration of the interval is displayed in the upper right corner.

3. USER STUDIES

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

3.1 Participants

3.1.1 Descriptive Statistics

We conducted studies with 18 participants, comprising of twelve male and six female. These participants stay the same over the entire evaluation process (both evaluation parts). All participants were required to own an Android phone with the version 2.3.3 Gingerbread or higher. Ages ranged from 18 - 31 (mean: 25.5, SD: 3.03). Seven participants were studying business informatics. Five others were studying in other fields. The other six participants did or still do an apprenticeship. Over 90% of the participants were German natives.

3.1.2 Acquisition

The participants have been recruited by asking the family, friends, class and flat mates of the development team. A personal reference between our participants 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 participant during the study more reliable compared to a group of non personal committed participants.

3.2 Overall Procedure

The evaluation was planed over two steps. First a qualitative user study followed by a quantitative user study.

As shown in Figure 3 we started with the first study at the end of April and finished the phase on the 5th of May. Afterwords, we aggregated the feedback and prioritized according to three categories before we started to implement the feedback. After that we conducted the second step of the evaluation.

The qualitative user study was conducted with a questionnaire and served as feedback at an early stage of development process. The following study has been conducted as a quantitative analysis. Within this part we asked the participants to use the app regularly and with their issues arise during the normal day live.

As it can be seen from Figure 3 we launched the second questionnaire, the last part of the second study, 2 weeks after the quantitative study was finished. Similar to the first questionnaire, this part was also finished within a period of

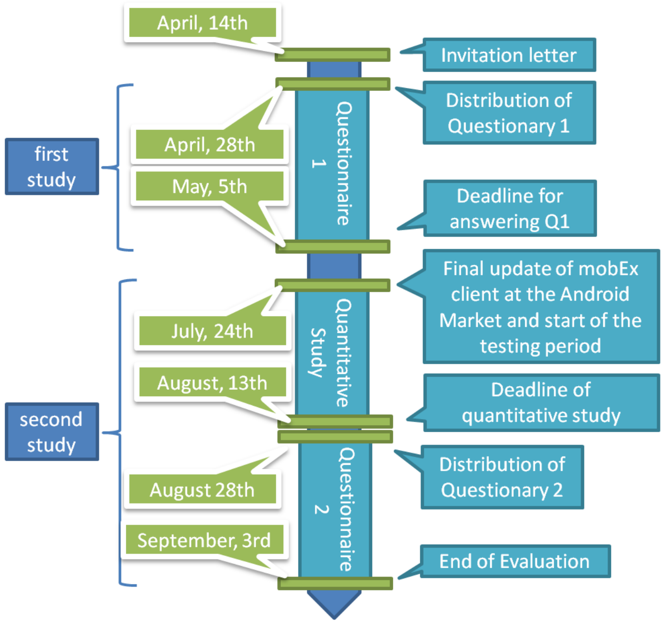


Figure 3: Overview of the time line of the procedure of the study settings

one week and - at the same time - marked the end of the entire evaluation period.

4. STUDY I: QUALITATIVE USER STUDY

4.1 Goals of study I

The goal of study I was to find out if the users understood the application, what users liked about it, what they did not like, and what further improvements or changes they suggest. Particularly the time-wheel was under special investigation since it is a novel search tool which cannot be found elsewhere.

4.2 Briefing of the participants

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

4.3 Programs and Technologies

The mobEx application was published on Google Play[5], 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 two questionnaires were done with Google Docs[6]. Google Docs offers a function to compose questionnaires and to collect the answers in a spreadsheet. To ensure, that each participant answered the questionnaires exactly once, an individual random pseudonym was used. That was sufficient to be assured not to loose the answers to any participant's questionnaire.

4.4 Evaluation Results

The evaluation summary of our two serialized questionnaires can be shown in Figure 4 There it is shown an abbreviation of all question we sequentially asked in both qustionnaires. The evaluation summary is shown below. So was the satisfaction with the loading time with a mean of 2.4 (the likert scale in the entire study means 1 as strongly agree which is in all cases a positive statement and 5 as strongly disagree that is in all cases a negative one) and the standard deviation with 1.0 lower as the satisfaction with the overall reactiveness (mean: 1.7; SD: 0.8) of the App. As it is shown at the third column in Figure 4 the app was with more than 70 % rated as innovative. A third Example of using this graphic is the trend about the "first impression" over time (between the first and the second survey). This question was rated with "same or equal": 8 participants, "better":8 participants and "worst" with 2 participants. There were no abstention from voting.

	satisfaction with loading time	satisfaction with the reactivness	innovativ	love the layout	the time wheel was helpful	correct interpretation of the time wheel	recommendation to friends	love the app	correct description of the manner of this app	user age	love the app	difference "liked the app" between first and second	recomendation	use futher on	keep first impression of the app	prefer grid	prefer list	end without success	stability of the app	love the layout	love the speed (reactivness and loadtime)	completeness	data quality	recognise deviation	feeling the some entries are missing	disturbed by the missing entries	intuitiv grid / list sorting
	Questionnaire 1										Questionnaire 2																
mean	2,4	1,7			3,4		2,7	2,2		25,5	2,6	0,4						2,7	1,9	2,3	2,1	2,4	2,2	2,7	2,6	3,2	2,2
median	2,0	2,0			4,0		3,0	2,0		25,0	2,0	0,0						3,0	2,0	2,0	2,0	2,0	2,0	2,0	2,0	3,0	2,0
Varianz	1,0	0,6			1,3		0,3	0,4		9,2	0,6	0,6						1,2	0,9	0,7	0,6	1,1	1,0	1,2	1,8	1,2	0,4
SD	1,0	0,8			1,2		0,6	0,6		3,0	0,8	0,8						1,1	1,0	0,8	0,8	1,0	1,0	1,1	1,3	1,1	0,6
"yes"			13	11		11			17				7	13		12	9										
"no"			5	7		2			0				11	5													
N/C			0	0	2	5			1																		
same (0)															8												
better (1)															8												
worst (-1)															2												

Figure 4: Overall analysis results of the qualitative study part

After the complete evaluation process we identified that more than 61% liked the overall application, compared to approximately 70% at the beginning of the evaluation. Furthermore, 2 out of 3 participants are going to use this application in the future. 16 of our participants answered that the impression of the application did not went worse over the evaluation process and 8 users attested a better impression at the second questionnaire (after the quantitative evalua-

[5]https://play.google.com/store

[6]https://docs.google.com

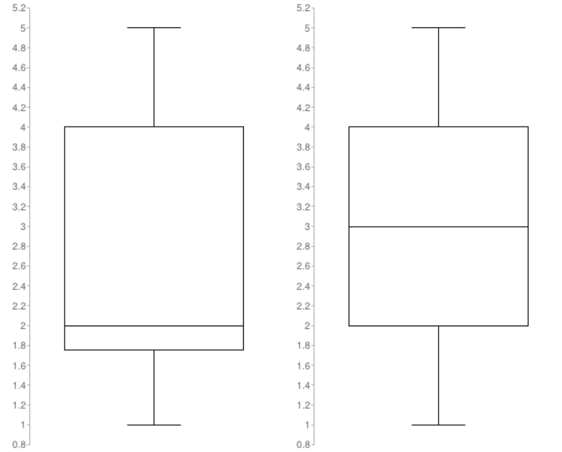
tion process). Two participants, who expressed a negative opinion of our app in both questionnaires, explicitly complained about the data incompleteness of mobEx.

Nine participants preferred the grid navigation mode in the facet screen whereas six participants preferred the list navigation. The remaining participants did not have a preference for list or grid. In the first questionnaire we asked about the usability of the time-wheel on a scale from 1 to 5 (5 point likert scale). Out of 18 times, this question had been answered 10 times with value 4 or worse (mean: 3.4; SD: 1.2). Just one participant was totally satisfied with usability of the time-wheel and three of our participants rated the usability with 2. 50% of the participants who rated the time-wheel negative (with a likert scale value of 4 or higher) did not fully understand the advantages of it. In contrast, all of our participants interpreted the functionality and the use case of mobEx in the correct manner. 13 participants liked (rating better 2 at a likert scale) the layout of the user interface at the very end of our evaluation period. This number increased compared to the beginning (11 participants liked that). Just one participant downrated the layout in the second questionnaire compared to its first rating. In the second questionnaire, approximately 20% complained about the performance of the application and criticized the latency until the events are presented on the client. This result doesn't change to the beginning of the evaluation.

The completeness (statement: In my opinion the data entries reflect the environment in a complete manner.) as well as the data quality (In my opinion the data quality is very good.) of the app was rated on average with 2.4 and 2.2 (each with a SD: 1.0) on a 5 point likert scale (1 strongly agree, 5 strongly disagree). 10 participants (55%) did not even recognize that some entries were totally missing over the entire questionnaire (rating 2 or better on the 5 point likert scale (mean:2.6; SD:1.3) for the statement: I did not have the feeling that some events or places were missing(Figure 5a).) but the others complained about the missing ones as soon as they recognized (5 participants did so with a rating 4 or worse on the 5 point likert scale) (Figure 5b). Close to 50% of the participants complained about the missing entries (rating 4 or worst(mean:3.2 SD: 1.1) at statement: It did not bother me that events or places were missing.) as soon as those entries attracted attention to them.

In addition, the top 20% of the most active users in our focus group evaluated the overall quality as well as the design of the application consequently positive. At the end of our evaluation 11 participants are likely going to recommend mobEx to friends. In contrast to the first questionnaire this is an improvement of almost 100 %.

In the first questionnaire we also asked in opened questions for concrete feedback about possibilities for further improvement and received some qualitative, constructive Feedback from our participants and analyzed it via an excel sheet where we categorize the feedback in categorization / client / detail view / facet / map / server / general or tweets issues



(a) I did not have the feeling that some events or places were missing

(b) It did not bother me that events or places were missing.

Figure 5: Boxplot Overview of the Evaluation

first. Afterwords we aggregated the feedback how often this specific issue were mentioned. When there are no brackets with further information behind any single entry attached it means that we received this feedback exactly once. This list is the aggregate constructive feedback received from our participants during the first study. The numbers behind several entities there marks how often it was mentioned in our response.

The (blue) colored items were not solved because of time or strong technical difficulties. The (green) colored items were rejected because of contrarian team decisions and non-colored issues were solved before the second study started.

- Categorization
  - Category names are useless, confusing or too abstract (10 times)
  - Breadth and depth are inconsistent
  - No consistent language (mixed up German/English)
- Client
  - The loading time of data is too long (14 times)
  - Change the description
  - Tutorial that explains the app might be helpful
  - Offer an "offline" mode
- Detail View
  - Do not show the entire URL; make it shorter (2 times)
  - The detail dialog fills up the hole screen; it is not possible to click outside of it to close it
- Facets
  - Switching between list and map mode doesn't work out fluently

  - There was no ”deselect all” button found (better place it at the top, not at the end of the filter view) (4 times)
  - Use ”standard icons”; others make the user confusing
  - Change layout (reduce the thick action bar)
  - The ”back” button in the middle of the facet disturbs a lot (remove it)(2 times)

- Map
  - Map zooms in instead of showing the events after a click on a clustered event; that’s confusing (2 times)
  - Too huge or useless / insensitive ’timewheel’ to much different date setting possibilities (9 times)
  - Better show the weekday when time is set
  - Time representation should not be longer than dd:mm:yy hh:mm + xx minutes | hours | days | ...

- Server
  - No complete data provided (2 times)
  - Include POIs from open street maps; they are often better than Google Maps information
  - For many restaurants opening hours are not available (2 times)
  - More events should be shown, like cinema / cultural events or others
  - Show the menu of restaurants

- Team
  - Logo looks old fashioned
  - There is no more functionality than in googlemaps app (3 times)

- Tweets
  - Tweets without follower is not useful; better use the own twitter account
  - Tweets from over a year do not have any good usecase
  - show the tweet - message and not just the user name
  - Movement patterns by tweet analysis are possible! Privacy?

4.5 Discussion of Results

The user suggestions were very helpful to make improvements and add features to the application. However, one issue was out of our influence and therefore hard to tackle. The issue of ”insufficient and incomplete data” which was criticized by several participants is mainly determined by the quality of the data sources. 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 be 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 together in order to provide more complete information about a POI and also to eliminate duplicate results. Yet, the usage of multiple data providers could not entirely solve the problem of incomplete data. This issue were especially mentioned during the open questions in the first qualitative study. But it holds also true for the second questionnaire in which this questions especially arise.

Some participants 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. Better information about opening hours could tremendously improve the user satisfaction in the future.

Remarkably, more participants preferred the grid over the list, even though list-based navigation menus are more common for today’s smartphone application. (preferred the grid: 12; preferred the list:9) One reasons for that may be the layer level indicator which is not available in the list-based navigation. It keeps track of the current position in the facet tree and may be supportive, especially when browsing large spaces of data.

Another point of criticism - with partly contradictory user feedback - was the granularity of the facet structure. Some participants 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 best level of granularity and the preferred granularity always depends on each individual user.

4.6 Implications and Reactive Actions

Our reactions were focused on the constructive feedback about how we can improve our app. This list was introduced already above and was discussed during a team meeting session. We also prioritized each single entry during this session and decided which issue should be and which can not be fixed within our project. Quite often also several possible solutions for these issues have been discussed.

We did not react to the colored issues because of several reasons. In the most cases it was because of unfeasible solutions in matter of time or technical possibilities (blue) in some cases we, as a team, just thought differently about the prioritization, the functionality of the solutions or, in our opinion, it was just useful to keep the current solution (green). To offer an offline mode in mobEx, we would have needed to change the internal data structure and implement a persistent database at the client side. Furthermore it was rated with a low priority because it was just mentioned once. Thus, this suggestion has not been followed further. The server issues would also have ended up in huge effort by including additional data providers or it would be just unfeasible to show the restaurant menu (no structured data available), opening hours or an entire data stock (bad database quality). The ”tweets” feedback were discussed but also not solved because of its low priority.

The non-colored issues were eliminated until the quantitative study began.

5. STUDY II: QUANTITATIVE USER STUDY

In the second step, we conducted a user study where the participants were asked to use the app regularly, in the best case daily, for their own searches that usually arise. Within this 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 participants to answer a second questionnaire after the three week logging period.

5.1 Goals of Study II

The goal of this study was to get insights about how mobEx is used. We focused especially on how users interacted with the time-wheel and the grid or list navigation. 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.

5.2 Logging methods

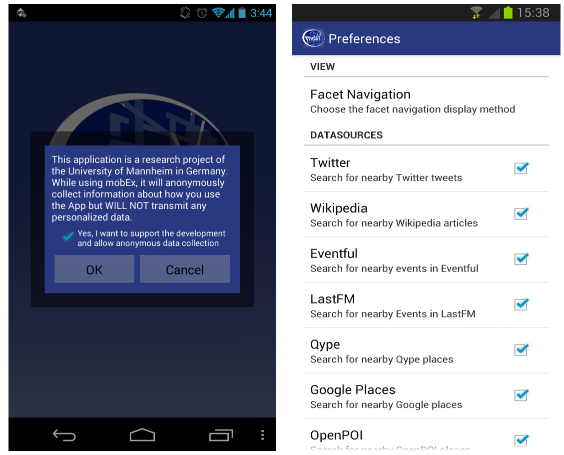
For logging the user's activity, we used Google Analytics[7]. This free service helps to analyze user traffic and can be customized to capture individual events. The data can be examined via a web application. As long as a user has not opted out of logging and transferring the data to Google Analytics, every activity such as a click on a button or the interaction with the time-wheel results in an event that is logged. The user is asked for his consent at the first start of the app. A dialog (see Figure 6a) containing a short statement and a pre-checked check box for acknowledging the statement is shown. In order to accept, users have to click the 'OK' button. If users don't accept they need to uncheck the box and confirm with 'OK'. A click on 'Cancel' closes the app. A study has shown that this approach results in a good share of users who give their consent [3].

5.3 Briefing of the participants

On July 21st, we started the quantitative part of our evaluation by sending a notification to our participants. We asked them to update the app from the Android Market first and to place a shortcut of the app on their home screen. We motivated the participants 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 since the first step.

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

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



(a) The user is prompted to give his consent for logging his activity

(b) The default setting for the selection of data providers

Figure 6: Screenshot of the informed consent and settings menu

5.4 Data Set

The data set used for this study corresponds to the resources 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 (Figure 6b). Before the result is send back to the client, the server uses entity-resolution technique to match similar entities from different data providers together. 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 the same location. By treating time dependent entities differently according to the given time, the data set also differs on the current time and the time that has been set by the user using the time-wheel.

5.5 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 participants 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 8 shows the distribution of the different session durations. In Figure 9 the number of session during the three week period is plotted. 55% of the participants used the application daily, 72% every second day or more. Figure 7 shows how many sessions each user has started during the study. 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, we observed only minor changes. Dbpedia and Twitter have been disabled once and Twitter has been enabled also once.

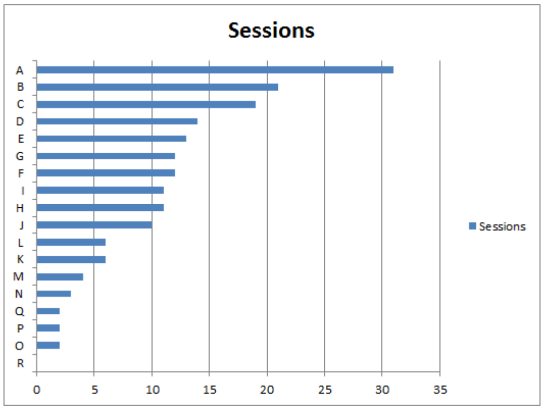


Figure 7: Sessions per User ordered by total number of user sessions

img/sessionDuration.png

Figure 8: Distribution of Session Durations

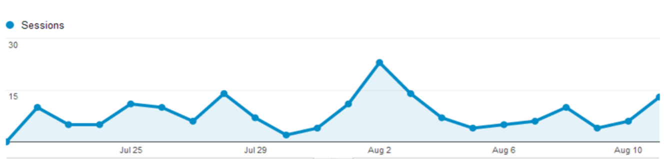


Figure 9: Sessions during the three week evaluation period

List and Grid Navigation

For investigating the acceptance of the grid and list navigation, we took two different measurements. First, the amount of sessions in which grid or list navigation was used and second, how many participants 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 participants switched to grid navigation whereas four times participants switched to list. The initial setting, grid or list, was decided by random. We counted 10 users having list navigation enabled and 8 users having grid navigation enabled at the beginning of the study. Three participants switched from list to grid. We did not observe any switches 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' or 'up' navigation to the upper level were done using the back button and 90% using the back-facet in the middle of the grid. 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 and 81% the top list entry in order to move up to the prior level. Figure 10 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. There is no time interval elapsing within the user has to interact. Participants having the list navigation enabled switched in 37% of the cases directly to the map screen. 36% of the first actions taken were interactions with the event list, (i.e scrolling through the list of events, showing details of an event or searching for an event) and 10% can be assigned to the list navigation (i.e browsing through or selecting facets). The corresponding values for grid navigation users show that in 26% of all first actions the user switched to the map screen, 30% were actions related to the event list and 15% of all actions were interactions with the grid navigation. In both cases actions that result in a switch to the preferences or help screen accounted for the remaining actions.

Another related measurement is to look at how many times users switched to the map as first action in contrast to how many times users first filter the events by selecting a facet with either grid or list navigation. To measure how this value changed within the 3 week period, we measured seven times, each sample containing the sum of three consecutive days due to infrequent usage of the time-wheel. The results are shown in Figure 11.

Time Wheel

The time wheel has been used in 25 of 179 session. In those 25 sessions, we measured 3.4 interactions per session with

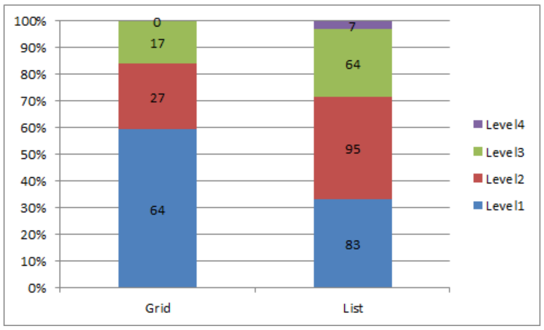


Figure 10: Distribution of how many times users navigated in a level of the facet tree using either grid or list navigation (in percent)

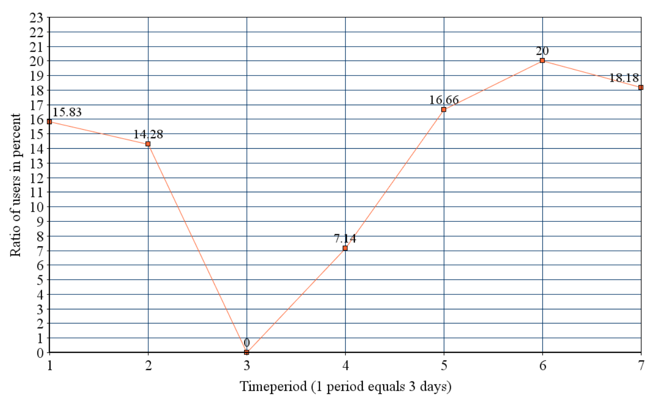


Figure 11: Ratio of users filtering events before switching to the map view

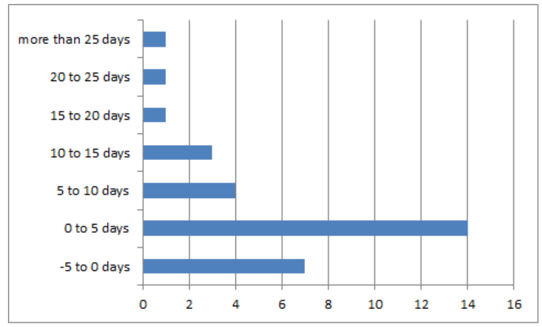


Figure 12: Distribution of the number of days users looked into the future

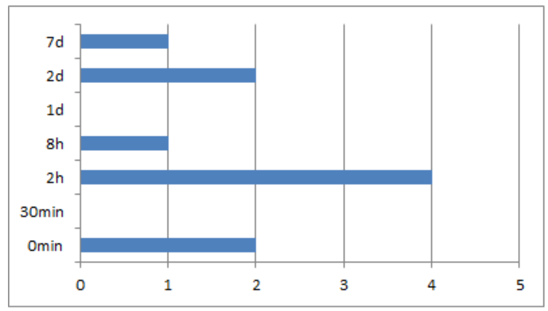


Figure 13: The different values the time interval can be changed to and the number of changes

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 participants 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 were opening the detail screen for an event as subsequent action. We measured the difference between the date that was set by the user, while using the time wheel, and the current date when the user left the map screen. The results summarized in different time spans are shown in Figure 12. Note that no data has been logged if the user did not use the time wheel. Thus, Figure 12 only contains data from users who changed the date with the time wheel. The time interval has been changed only 10 times, so most of the time the user was working with the standard value of 8 hours. The distribution of time intervals that the users have set can be seen in Figure 13. 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 within the three weeks of evaluation has not significantly increased or decreased 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.

5.6 Discussion of Results

In this section we discuss the results of the quantitative study together with the ones of the second questionnaire, which we handed out to our participants after the evaluation period. Over the three weeks of logging user data, we could observe a balanced distribution of usage. The participants used the application regularly and produced enough data that was used to create measurements about the behavior of users. Regarding the navigation type, we observed that more participants were using the grid navigation type in the end of the study. Although there were slightly more sessions with list navigation enabled, we think, that the grid navigation was the approach that had a better acceptance among the users. This statement can be supported by the answers of the second questionnaire where 9 participants indicated a preference for the grid navigation, in contrast to seven participants who preferred the list view. Having the list navigation type enabled, more users were using the back button in order to go up to a parent facet. This result is obvious 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 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 10. 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. This shows us that the grid navigation attracted more attention and could persuade more users to use it in order to filter events. 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 and 7% of the participants 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.

5.7 Implications for Future Work

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. Concerning implications of the application itself, it is necessary to add more time dependent events and intensify the focus on the temporal dimension. This would increase the usefulness of the time-wheel. Furthermore, the grid navigation mode can be set as default as it was the preferred view. With regard to the loading time of events it is necessary to implement a caching mechanism to compensate for slow data providers. Furthermore it is useful to acquire at least 40 percent more participants in order to compensate for those ones that do not regularly contribute to the study.

6. RELATED WORK

The procedure of our quantitative field study was based on the experience of Sohn et. al [7]. The authors describe a two-week diary study with 20 participants about mobile information needs. During the experiment period, the participants 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 to design 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 participants use mobEx in their 'natural habitat'.

Another field study that was conducted with users in their 'natural habitat' was done by Niels Henze [3]. 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 publishing process was very similar. We also deployed our application via the Android Market in order to reach the participants and enable them to easily use, rate, or update mobEx from home or while they are on the go. This also enabled us to deploy new versions of the app on a regularly basis during the development phase of mobEx without the need to explicitly notify the users about each update. Nevertheless, during the 3 weeks period of our quantitative study we kept the application stable and just care about the stability and error fixing of the app. We did not need to push such critical updates during our test period.

The idea of a grid-based faceted navigation on mobile phones was initially presented in 2006 by Karlson et al. [4]. 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 participants, 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.

In one of the previous studies [6], the grid-based and list-based approach have been compared via a between group experiment with 24 subjects. In this experiment, participants 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. The results of our field study support the idea that the grid-navigation is the better choice because most subjects preferred it over the traditional list.

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.

7. CONCLUSION

We have presented with mobEx an application for the mobile exploration of social media data, especially places, events, people, and organizations. Unlike existing applications, mobEx provides access to multiple databases and matches the content from different providers to offer a unified view on the data resources. Users are enabled to explore these data resources with unique search features such as the time-wheel and the facet navigation.

One of the main goals of our evaluation was to find out whether the users prefer a list-based or a grid-based approach for facet navigation as well as whether the time-wheel has been used. The findings of the quantitative and the qualitative study taken together indicate that the grid navigation is the better option when exploring large faceted data spaces and thus support findings from the past. However, it has to be kept in mind that the list navigation is faster and requires less clicks. [6]. During the quantitative study more participants switched from list to grid navigation and stayed there.

The qualitative study also uncovered that the purpose of the time-wheel was not well understood at the beginning. We also observed a rare usage of the time-wheel during the quantitative study. We believe that the poor usage was due to a lack of time dependent resources 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.

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 *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1353–1356. ACM, 2008.

[3] Niels Henze. Hit it!: an apparatus for upscaling mobile hci studies. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 1333–1338. ACM, 2012.

[4] 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 *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 711–720, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: 10.1145/1124772.1124878. URL `http://doi.acm.org/10.1145/1124772.1124878`.

[5] Gary Marchionini. Exploratory search: from finding to understanding. *Commun. ACM*, 49(4):41–46, April 2006. ISSN 0001-0782. doi: 10.1145/1121949.1121979. URL `http://doi.acm.org/10.1145/1121949.1121979`.

[6] Mark Schneider and Ansgar Scherp. Comparing a grid-based vs. list-based approach for faceted search of social media data on mobile devices. *Arbeitsberichte aus dem Fachbereich Informatik University Koblenz-Landau*, 01 2012. URL `http://www.uni-koblenz.de/~fb4reports/2012/2012_01_Arbeitsberichte.pdf`.

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