

A Tool for Conducting User Studies on Mobile Devices

Luca Costa
Università della Svizzera italiana (USI)
luca.costa@alumni.usi.ch

Mohammad Aliannejadi*
University of Amsterdam
m.aliannejadi@uva.nl

Fabio Crestani
Università della Svizzera italiana (USI)
fabio.crestani@usi.ch

ABSTRACT

With the ever-growing interest in the area of mobile information retrieval and the ongoing fast development of mobile devices and, as a consequence, mobile apps, an active research area lies in studying users' behavior and search queries users submit on mobile devices. However, many researchers require to develop an app that collects useful information from users while they search on their phones or participate in a user study. In this paper, we aim to address this need by providing a comprehensive Android app, called Omicron, which can be used to collect mobile query logs and perform user studies on mobile devices. Omicron, at its current version, can collect users' mobile queries, relevant documents, sensor data as well as user activity and interaction data in various study settings. Furthermore, we designed Omicron in such a way that it is conveniently extendable to conduct more specific studies and collect other types of sensor data. Finally, we provide a tool to monitor the participants and their data both during and after the collection process.

KEYWORDS

mobile search; context; user study; field study; mobile app; sensors

ACM Reference Format:

Luca Costa, Mohammad Aliannejadi, and Fabio Crestani. 2020. A Tool for Conducting User Studies on Mobile Devices. In *2020 Conference on Human Information Interaction and Retrieval (CHIIR '20)*, March 14–18, 2020, Vancouver, BC, Canada. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3343413.3377985>

1 INTRODUCTION

With the technological advances in developing mobile devices, we witnessed an evolution of mobile devices as well as the ways people access information through these devices. Research on mobile information retrieval (mobile IR) is aimed at “enabling users to carry out, using a mobile device, all the classical IR operations that they were used to carry out on a desktop” [8]. The evolution of mobile devices as well as their applications (a.k.a. apps) not only enables users to carry out the “classical IR” tasks on their mobile devices, but also identifies new methods of interaction and information access.

*Work done while Mohammad Aliannejadi was affiliated with Università della Svizzera italiana (USI).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHIIR '20, March 14–18, 2020, Vancouver, BC, Canada

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-6892-6/20/03...\$15.00

<https://doi.org/10.1145/3343413.3377985>

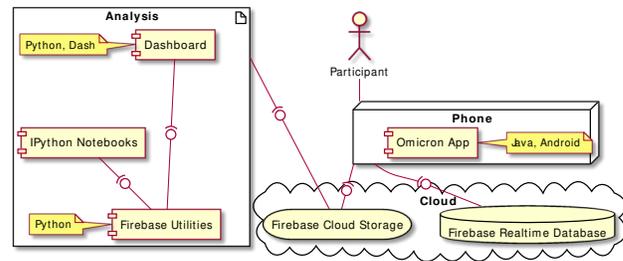


Figure 1: System Overview

Research on mobile IR started as early as 2006 when Kamvar and Baluja [11] studied query logs of Google mobile search. Since then, there has been growing interest in studying this area both in industry and academia. Early studies mainly focused on understanding users information needs on mobile devices [12] and exploring conventional Web-based IR approaches on these devices [6]. More recently, along with advances in technology, researchers have explored various aspects of mobile IR. For instance, Song et al. [15] studied and found significant differences in search patterns done using iPhone, iPad, and desktop. Crestani and Du [7] conducted a comparative study on mobile spoken and written queries showing that spoken queries are longer and closer to natural language. Sohn et al. [14] conducted a diary study in which they found that contextual features such as activity and time influence 72% of mobile information needs. Carrascal and Church [5] studied user interactions concerning mobile apps and mobile search, finding that users' interactions with apps have an impact on search. Harvey and Pointon [9] found that fragmented attention of users while searching on-the-go, affects their search objective and performance perception. Also, Aliannejadi et al. [1] confirmed the findings of this study by studying people's behavior while searching in different contexts through a field study. More recently, researchers indicated the need for a universal mobile search framework and found that commercial mobile search engines such as Google and Bing are not the preferred means of information access for the majority of users' information needs [3]. As it is obvious, above-mentioned studies either had access to commercial search logs or developed a specific app for their study.

Due the large number of studies that have been carried out in this area, the need for a comprehensive tool that facilitates data collection in mobile IR research is clear. To this end, we have developed, tested, and open-sourced Omicron, to provide researchers with a tool that is comprehensive and extendable for conducting a large variety of mobile IR tasks. Omicron enables task-based laboratory

and field studies. It can also serve as a light-weight mobile search engine allowing researchers to collect mobile query logs. Moreover, Omicron provides a simple way of configuring its background data collection service. The background service collects raw sensor data as well as high-level user activity information together with user interaction data while they use the app. Therefore, in this paper we aim to:

- Familiarize the reader with the Omicron app and the monitoring dashboard.¹
- Give an overview of the architecture and the implementation details.
- Elaborate on the usefulness of our system in relation to the need of an open-source customizable tool to conduct reproducible user studies.

Successful deployment and use of Omicron in our previous experiments indicates its usability as a tool for both conducting user studies and collecting search logs on mobile devices.

2 OMICRON

Here we present the details of the user interface, collected data, and implementation of Omicron.

System overview. An overview of the system is presented in Figure 1. As we can see, the user interacts with Omicron, while the app reads and stores the data on a cloud service (i.e., Google Cloud). It is worth noting that Omicron stores raw sensor data, as well as user interaction data on Google Cloud Storage. On the other hand, it reads and updates the task-related data in the Firebase Realtime Database. This choice was motivated by the relative cost and effectiveness of the two services to store a large amount of data as both services does not require building and maintaining the infrastructure and offer a usage-based billing plan which make them serviceable for large-scale studies and convenient for small ones. Finally, as we see the data analysis and monitoring component retrieves the data from the cloud server and visualizes it, which could be helpful for monitoring an ongoing study.

User interface. Omicron presents itself depending on the action that the participant is required to take. After a short tutorial (Figure 2a), a demo task plus a demographic and background survey, participants are presented with a screen where they can freely search and test the application. The app guides the participants by showing an informative text and by hiding non-required actions at every step. When the user receives a task via a notification, the main screen changes and awaits that the participant starts the task (Figure 2b). Figure 2c shows a sample mobile SERP where the participants consult search results and bookmark the ones that helped them complete the task. This screen is reached when the participant starts a task. Finally, as we see in Figure 2d, similar to an optional pre-task questionnaire, the participants is required to complete a post-task questionnaire.

Collected data. Apart from the participants' input data, Omicron also collects their interactions within itself (i.e., taps and scrolling). Moreover, a background service constantly collects the phone's

Listing 1: Excerpt of the configuration file

```
// Collectors
Bool shouldRecordAppUsage = true;
Bool shouldRecordLocation = true;
Bool shouldRecordAccelerometer = true;
Bool shouldRecordBattery = true;
...
// Sample rates (ms)
// app usage interval period
long USAGE_INTERVAL = 1000*60*60*24;
// sample rate location recording
long locationRate = 1000*60*3;
// sample rate
long sampleRate = 1000 *60*2;
[...]
```

sensors data. In summary, Omicron collects the following sensor data:

- GPS
- accelerometer
- gyroscope
- ambient light
- WiFi
- cellular

Furthermore, it collects other available phone data that can be used to better understand a user's context. The additional collected data are as follows:

- battery level
- the screen on/off events
- apps usage statistics
- apps usage events

Note that apps usage statistics indicate how often each app has been used in the past 24 hours, whereas apps usage events provide more detailed app events.² Apps usage events record user interactions in terms of:

- launching a specific app
- interacting with a launched app
- closing a launched app
- installing an app
- uninstalling an app

This service collects the data at a predefined time interval and securely transfers it to the cloud service.

2.1 Implementation Details

Omicron is created as a tool to conduct task-based user studies. It is crucial to have a way to schedule tasks, have a search interface to execute queries (in our case we used Bing) and gather sensor data at the end of the study.

Tasks. A task is defined by an ID, a start time and end time, a title, a description and a series of metadata (e.g., the time the participant start the task). For instance, we could have a task with the title

¹Code available at: <https://github.com/aliannejadi/Omicron>

²<https://developer.android.com/reference/android/app/usage/package-summary>

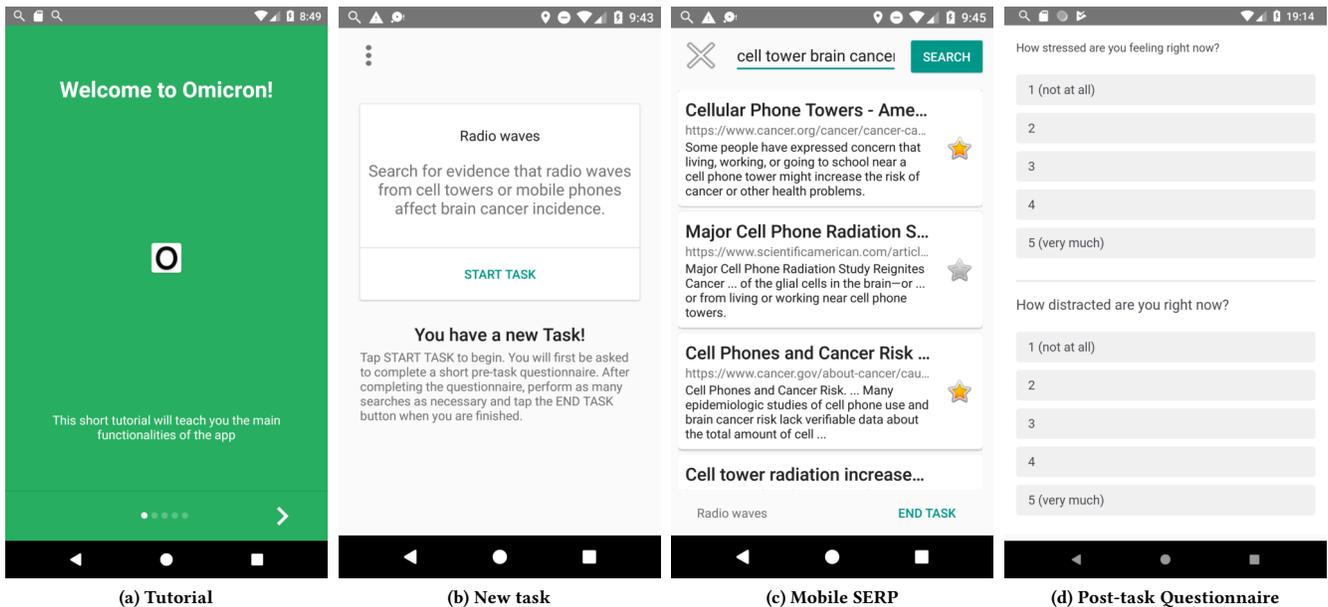


Figure 2: Different user interfaces of Omicron.

"Food" and in the description specify that the participant will need to search for a recipe for dinner. We would then set a date and we would specify an interval of hours in which this task will need to be completed.

A task can be defined in various ways. For instance, each participant can have a series of tasks assigned to them on a particular day. The way tasks can be assigned to a participant or a group of them is via the application ID, an identifier generated from the instance of the installation of Omicron. Tasks can be changed or added during the experiment and they can be rescheduled if the participant does not manage to do them in the assigned slot.

Configuration. Omicron can be customized by editing a single configuration file. Listing 1 is a small excerpt of the configuration file. It can be seen that the gathering of the various data can be conveniently switched on or off. The interval for sampling or transferring the data can also be modified easily. It is also possible to set the link for a survey after the installation and pre- and post-task questionnaires. Some of the sensor data can be chosen to be transmitted only over WiFi to consume less mobile data.

Dashboard. To be able to monitor the ongoing status of the study, we have built a small Web dashboard. This might be less needed for in lab studies since one can observe participants directly. In the field study setting, however, it is critical to monitor the progress of the study. We found it to be particularly useful to know when a participant was lost (the result of the queries did not answer the task) or if there was a problem in data collection. Jointly with a standard crash and error reporting dashboard, this gave us an overview of how the study was proceeding. Figure 3 shows one of the tables that correlated the queries with some of the other data

that we collected. The dashboard is made in Python and Dash³. Although is not possible to personalize it graphically, it is relatively easy to convert Jupyter Notebooks⁴.

3 USABILITY

We developed Omicron as part of our effort to conduct an in-situ user study on smartphones. We carefully took the required measures in developing an app that runs smoothly on a large variety of Android devices. More importantly, since the participants had to install the app on their phones, we aimed to design an efficient application that does not affect the people's everyday usage experience. Therefore, we aimed to develop Omicron in such a way that it would not slow down the device, would not consume much energy, and would recover from possible crashes without user's involvement.

While conducting our previous studies [1, 2], Omicron was successfully installed and used by over 350 users worldwide. Throughout the studies, we continuously monitored the collection of data, as well as the performance of the app. Several bugs were addressed throughout the study quickly, accompanied by app updates on the Play Store to ensure a positive experience of the participants. We used Omicron to conduct an in-situ user study where we asked the participants to perform pre-defined search tasks [1]. In another effort, we used the Omicron's background service and extended the UI to collect a mobile query search log [2]. We see in Figure 4 the number of reported queries, as well as users who installed and used the app on their phones during this study. It is worth noting that during these studies, Omicron collected and stored over 350 GB of data on cloud servers. In both studies, Omicron was installed and

³<https://plot.ly/dash/>

⁴<https://jupyter.org/>

timestamp	content	jsonResult	AppUsages	UsageEvents	LocationTime	LocationLatitude	LocationLongitude
2018-07-14 10:56:31.940	radio waves cancer	{"_type":"SearchResponse	{'App': 12 WhatsApp 65 S	{'EventType': 0 1 1 2 2 1 :	2018-07-14 10:57:08.872	46. .	9. .
2018-07-16 09:01:28.848	tesla switzerland	{"_type":"SearchResponse	{'App': 21 Instagram 35 F	{'EventType': 0 1 1 2 2 1 :	2018-07-16 09:00:35.479	46. .	9. .
2018-07-16 09:02:14.752	electric cars list	{"_type":"SearchResponse	{'App': 21 Instagram 35 F	{'EventType': 0 1 1 2 2 1 :	2018-07-16 09:02:35.352	46. .	9. .
2018-07-16 13:19:24.195	steak sauce	{"_type":"SearchResponse	{'App': 21 Instagram 35 F	{'EventType': 0 1 1 2 2 1 :	2018-07-16 13:20:13.613	46. .	9. .
2018-07-17 13:50:05.297	restaurant near zurich lal	{"_type":"SearchResponse	{'App': 80 Maps 13 Whats	{'EventType': 0 1 1 2 2 1 :	2018-07-17 13:49:10.079	47. .	8. .
2018-07-17 17:07:55.848	list of 2018 movies	{"_type":"SearchResponse	{'App': 82 Maps 13 Whats	{'EventType': 0 2 1 1 2 2 :	2018-07-17 17:08:48.353	46. .	8. .
2018-07-17 17:09:10.668	best movies to watch	{"_type":"SearchResponse	{'App': 82 Maps 13 Whats	{'EventType': 0 2 1 1 2 2 :	2018-07-17 17:08:48.353	46. .	8. .
2018-07-18 09:04:15.652	biking routes gandria	{"_type":"SearchResponse	{'App': 32 Maps 13 Whats	{'EventType': 0 1 1 2 2 1 :	2018-07-18 09:04:09.689	46. .	9. .
2018-07-18 09:05:22.022	bike routes ticino	{"_type":"SearchResponse	{'App': 32 Maps 13 Whats	{'EventType': 0 1 1 2 2 1 :	2018-07-18 09:06:09.980	46. .	9. .
2018-07-18 09:05:25.915	bike routes ticino map	{"_type":"SearchResponse	{'App': 32 Maps 13 Whats	{'EventType': 0 1 1 2 2 1 :	2018-07-18 09:06:09.980	46. .	9. .
2018-07-18 14:56:55.621	slovenia travel info	{"_type":"SearchResponse	{'App': 32 Maps 82 Insta	{'EventType': 0 2 1 1 2 2 :	2018-07-18 14:56:15.494	46. .	9. .

Relevant Results

Filter Rows

timestamp	index	link	title
2018-07-22 09:30:09.489000	043	https://www.tripadvisor.com/ShowTopic-g188045-133	Bus Schedule from Lugano to Menaggio - Switzerland
2018-07-22 09:29:36.230000	043	https://www.tripadvisor.com/ShowTopic-g188045-133	Bus Schedule from Lugano to Menaggio - Switzerland

Figure 3: Part of the monitoring dashboard: The tables correlate the queries with the location, starred websites and application usage

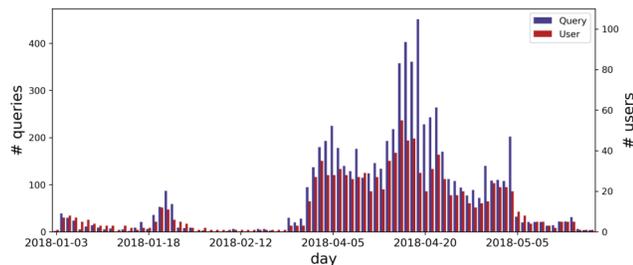


Figure 4: Number of queries and active participants per day, during the course of data collection [2] (best viewed in color).

used on multiple smartphones from various manufacturers featuring different versions of Android OS. These successful experiences enforce the usability of Omicron in the community.

Our software design allows for easy extension to collect sensor or user data that would emerge in the future. Also, the UI can be easily modified to be used for other types of studies. For instance, one could add a voice-based interface to Omicron for user studies based on spoken conversations. Notice that we have tested Omicron only for field studies and on personal phones; however, nothing limits it to be used for controlled lab studies. Therefore, one could use Omicron in a much more controlled experimental settings. The design of the app, as well as the database of tasks, allows for several use cases, not to mention the possibility of extending the current version for a more specific need.

4 CONCLUSIONS AND FUTURE WORK

In this paper we introduced Omicron, an app to collect mobile query logs which includes other system information about the task, and perform user studies on mobile devices. We gave a brief overview of the application and the various components that are part of the architecture. We showed how each piece fit together and how to customize Omicron to conduct a study. Writing a mobile app is

expensive and time-consuming. Omicron offers a starting point for the researchers in the community to develop their required study apparatus on top of Omicron, without worrying about a bespoke application. Furthermore, we have designed Omicron to be easy to customize so that it can adapt to the needs of different studies. It can also be easily configured to perform other types of studies, or be used as a simple tool to collect mobile query log. Finally, it is worth noting that Omicron provides the ground towards more reproducible studies.

In the future, we plan to extend Omicron to be modular and explore gamification for mobile data collection. One of the challenges of building an application that can be reused for different types of user studies is that for instance, one study might need to gather sensor data but it would not make use of the search feature. Unfortunately at the moment, it is hard to do it without adventuring in the code base and knowing Android programming. The architecture could be made more modular to allow both the reuse of low-level components and graphical components, doing this could allow easier reuse of Omicron even for studies that differ considerably from what it was created for. As we know, it is difficult to gather enough participants to conduct mobile IR studies at large scale. By making the studies open via mobile app stores it is possible to sample from a vast audience as shown in studies like [10] and [13]. Every study would need to adapt game elements in a way that incentivize participants to take part in their studies. Therefore the type of study would dictate a different approach. However, there are known game elements that can be reused in many studies, for instance leader boards and points. Also, we plan to extend Omicron to enable studies on conversational search on mobile devices [4].

ACKNOWLEDGMENTS

This research was in part funded by the RelMobIR project of the Swiss National Science Foundation (SNSF). We would like to thank Jacopo Fidacaro, who helped us in developing Omicron as part of his summer internship. Also, we would like to thank Morgan Harvey and Matthew Pointon for their feedback on the app's design.

REFERENCES

- [1] Mohammad Aliannejadi, Morgan Harvey, Luca Costa, Matthew Pointon, and Fabio Crestani. 2019. Understanding Mobile Search Task Relevance and User Behaviour in Context. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, CHIIR '19*. ACM, 143–151.
- [2] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W. Bruce Croft. 2018. In Situ and Context-Aware Target Apps Selection for Unified Mobile Search. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM '18*. ACM, 1383–1392.
- [3] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W. Bruce Croft. 2018. Target Apps Selection: Towards a Unified Search Framework for Mobile Devices. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR '18*. ACM, 215–224.
- [4] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W. Bruce Croft. 2019. Asking Clarifying Questions in Open-Domain Information-Seeking Conversations. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '19*. ACM, 475–484.
- [5] Juan Pablo Carrascal and Karen Church. 2015. An In-Situ Study of Mobile App & Mobile Search Interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*. ACM, 2739–2748.
- [6] Karen Church, Barry Smyth, Keith Bradley, and Paul Cotter. 2008. A large scale study of European mobile search behaviour. In *Proceedings of the 10th Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI '08 (ACM International Conference Proceeding Series)*. ACM, 13–22.
- [7] Fabio Crestani and Heather Du. 2006. Written versus spoken queries: A qualitative and quantitative comparative analysis. *JASIST* 57, 7 (2006), 881–890.
- [8] Fabio Crestani, Stefano Mizzaro, and Ivan Scagnetto. 2017. *Mobile Information Retrieval*. Springer.
- [9] Morgan Harvey and Matthew Pointon. 2017. Searching on the Go: The Effects of Fragmented Attention on Mobile Web Search Tasks. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '17*. ACM, 155–164.
- [10] Niels Henze, Benjamin Poppinga, and Susanne Boll. 2010. Experiments in the wild: public evaluation of off-screen visualizations in the Android market. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction '10*. ACM, 675–678.
- [11] Maryam Kamvar and Shumeet Baluja. 2006. A large scale study of wireless search behavior: Google mobile search. In *Proceedings of the 2006 Conference on Human Factors in Computing Systems, CHI '06*. ACM, 701–709.
- [12] Maryam Kamvar and Shumeet Baluja. 2007. The role of context in query input: using contextual signals to complete queries on mobile devices. In *Proceedings of the 9th Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI '07 (ACM International Conference Proceeding Series)*, Vol. 309. ACM, 405–412.
- [13] Donald McMillan, Alistair Morrison, Owain Brown, Malcolm Hall, and Matthew Chalmers. 2010. Further into the Wild: Running Worldwide Trials of Mobile Systems. In *Pervasive Computing, 8th International Conference, Pervasive '10*. Springer, 210–227.
- [14] Timothy Sohn, Kevin A. Li, William G. Griswold, and James D. Hollan. 2008. A diary study of mobile information needs. In *Proceedings of the 2008 Conference on Human Factors in Computing Systems, CHI '08*. ACM, 433–442.
- [15] Yang Song, Hao Ma, Hongning Wang, and Kuansan Wang. 2013. Exploring and exploiting user search behavior on mobile and tablet devices to improve search relevance. In *Proceedings of the 22nd International World Wide Web Conference, WWW '13*. ACM, 1201–1212.