

Evaluation of a User-centric QoE-based Recommendation Tool for Wireless Access

Michalis Katsarakis, Vasileios Theodosiadis, and Maria Papadopouli*
Department of Computer Science, University of Crete, and
Institute of Computer Science, Foundation for Research and Technology-Hellas

ABSTRACT

This paper presents the u-map, a user-centric crowd-sourcing recommendation system that enables mobile users to select the best service provider in a certain region. Via u-map, users upload information about their profile, quality of experience (QoE) feedback for a service, network measurements, and position in the database of u-map. In our earlier work, we analyzed the performance of u-map, in terms of power consumption, responsiveness, and scalability. To understand how users assess such systems in general, and how they perceive the performance of u-map and its services specifically, we performed a field study. The study evaluates the GUI of the u-map client and the impact of u-map on the discovery of new networks for improving the QoE of VoIP calls. The paper presents a methodology for performing such studies, analyzes the collected measurements, models the user satisfaction about u-map, and highlights the main findings.

1. INTRODUCTION

Wireless access, use, and traffic demand are on a fast rise. Wireless networks manifest periods of severe impairment, causing severe degradation of the quality of experience (QoE) of services. Due to the diverse set of services, network operators, and users, the estimation of the QoE is challenging and largely under-explored. Furthermore, there are no automated real-time QoE diagnostic tools. The traditional mobile network access paradigm is both static and fragmented:

* This work is supported by the General Secretariat for Research and Technology in Greece with a Research Excellence, Investigator-driven grant, 2012, and by a Google Faculty Research Award, 2013, PI Maria Papadopouli (mgp@ics.forth.gr).

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.

C2B(1)D'15, August 17-21, 2015, London, United Kingdom

© 2015 ACM. ISBN 978-1-4503-3539-3/15/08...\$15.00

DOI: <http://dx.doi.org/10.1145/2787394.2787402>

Users subscribe to or prepay for specific cellular operators/providers for long-term network access. However, the technology that allows user devices to select the appropriate operator and/or base stations (BS) is ready. Moreover, the Body of European Regulators for Electronic Communications (BEREC) in recent reports envisages measures for *consumer empowerment*, boosting *consumer choice*, *network transparency* and developing mechanisms for *data collection and analysis*. New technological and business paradigms (e.g., the network virtualization, MVNOs, access sharing) provide more options to users and providers. Motivated by the need to empower customers to *select the appropriate service provider/operator* and make educated decisions, especially when visiting a new region, we developed the u-map.

The u-map supports a *crowd-sourcing* monitoring system that collects *real-time* information about the performance of the network and services. This information includes network measurements as well as customer feedback and opinion scores about services. The u-map can then apply statistical analysis and data mining techniques on these data to recommend the most appropriate service provider, enabling users to make educated selections. Providers/operators can also query u-map to obtain feedback about the quality of their network and services in an inexpensive, almost real-time manner. U-map can act as a “semi-automatic” alarm detection system: it can enable providers to better understand a target user population, model their QoE, and detect problems with their network infrastructure and services.

Databases have been proposed to maintain spectrum or physical-layer based information (e.g., [1–6]), focusing on spectrum availability/usage, “whitespace”, and interference. Micro-Blog [7] users can upload measurements captured by their cameras, GPS, accelerometers, or health monitors. MyExperience [8] collects various measurements about device usage, user context information, and environmental measurements. OpenSignal maps [1] and Sensorly [9] focus on signal strength measurements. Farshad *et al.* [10] utilize crowdsensing for urban WiFi characterization. Portolan [11] also performs traceroute-like measurements for the discovery of the Internet topology. WiScope [12] opportunisti-

cally collects transport-layer measurements on laptops to provide a coarse-grained view of the wireless landscape. WiScape [12] provides a coarse-grained view of the wireless landscape by opportunistically collecting transport-layer measurements. MCNet [13] measures the WiFi performance using the crowdsourcing paradigm on Android devices. WeFi [14] allows the automatic connection to the best available network in terms of throughput. Typically monitors focus on either the mobile device [15–19] or the infrastructure. Unlike these approaches, u-map integrates a richer set of data, encompassing cross-layer measurements (e.g., RSSI values to QoE scores) and user preferences (e.g., about data rate and price).

Our earlier work [20, 21] presented the main concept of u-map and a preliminary performance analysis that focused on the responsiveness and scalability of u-map. This work extends the analysis by performing a field study. The field study was conducted in a two-week period and included 21 users, aiming to a) evaluate the overall performance of u-map, its Graphical User Interface (GUI), and its main functionalities, e.g., provider selection, improvement of QoE, b) highlight the most important parameters that affect the user experience and future adoption of u-map, and c) address the main weaknesses of u-map. The study also provided a methodology for a) modeling the overall user satisfaction about u-map, b) assessing the services that u-map provides and the need for such systems, and c) learning about the user requirements and profiles. The paper is structured as follows: Section 2 describes the architecture of u-map, while Section 3 focuses on its evaluation. Finally, Section 4 summarizes our conclusions and future work plan.

2. SYSTEM ARCHITECTURE

The u-map is a review and recommendation system that follows the client-server architecture. A u-map client, running on mobile devices, stores locally network measurements collected at the background during a session, such as a GSM or VoIP call, web browsing, game, or video streaming. Furthermore, it enables a user to indicate his/her QoE feedback, at the end of the session. It dynamically uploads these network measurements and subjective data on a spatio-temporal geo-database of the u-map server. The u-map server collects, processes, and stores these data. The large size and heterogeneity of the collected measurements, need for data integrity, fault-tolerance, erroneous-data detection (e.g., injected by mis-configured or malicious users), treatment of missing values, protection of user privacy, and access control are critical aspects that need to be considered for the design of u-map. The u-map client (henceforth referred to as *client*) includes a *monitor*, *performance estimator*, *database*, *back-end interface*, and its *GUI*. The monitor runs in the background and records network-related information, such as traffic demand, network data (e.g., network type & provider,

RSSI, interference, packet loss) together with timestamps and position information. Session data is also recorded. The prototype focused on two types of sessions, namely, the GSM and Sipdroid calls. For each session, the start and end time, termination status (i.e., was terminated successfully or abruptly, or was blocked), and user feedback are recorded. The user feedback consists of an opinion score, followed in some cases by an explanation (e.g., as in the Skype). Each session is associated with the network measurements and opinion scores collected during that session. The performance estimator assesses the services of those sessions. The client internal database stores *temporarily* the recorded data using OrmLite, an object-relational mapping package. The back-end interface includes a *secure* HTTP client and the functionality that connects the client to the HTTP server using the JSON data-interchange format for communication. During a connection, a client may upload the recorded information to the u-map server. The client runs on the Android OS. The u-map server includes the set of services and mechanisms to receive data from clients, process, store, and analyze it, as well as allow controlled access to it. The u-map server consists of a *PHP application*, *database*, and *security and privacy* components. The PHP application is a content management system that enables the u-map server and client communication. The registration and login processes are also implemented in the PHP application using a certificate authority to sign and validate client certificates. The server employs the PostgreSQL with the PostGIS extension, which allows the formation of spatial queries. The user privacy is protected by requiring authorization for granting access to the data. The client-to-database connection relies on the end-to-end security that protects the integrity and confidentiality of the submitted data by leveraging standard technologies (e.g., public-private key pairs, TLS). To further protect sensitive information, access is allowed only to aggregate statistics. The u-map server implements a *query rewriting access control* mechanism that allows users to control the information revealed to third parties through a fine-grained discretionary approach. More precisely, access control rules define *who* has access to *what* data; *who* can be a user or a role (e.g., operator, application), and *what* is a query over the data. The u-map server (simply server from now on) runs on a Linux machine.

3. PERFORMANCE ANALYSIS

This section evaluates the usability of u-map, and its impact on improving the QoE of VoIP calls. In our earlier work [20, 22], we analyzed the power consumption, responsiveness, and scalability. For the assessment of the usability, functionality and services of u-map, a *field study* was conducted. We use the terms subjects, participants, or volunteers, interchangeably to refer to these participants. During the field study, *objective measurements* and *subjective scores* were collected via u-map.

Specifically, via u-map clients, the *subjects evaluated the QoE of their VoIP calls*, and these QoE scores, along with network measurements, were uploaded to the server. Furthermore, at the start and end of the study, the participants provided feedback by answering *two questionnaires* (available online at [23]). The data were analyzed using various statistical analysis and machine learning algorithms.

Methodology: To evaluate the user experience when running u-map, a field study was conducted at the premises of FORTH, for a two-week period, with the participation of 21 volunteers. Common practices and guidelines for performing subjective studies [24, 25] were followed. The subject population was composed by graduate students and junior researchers from FORTH, excluding the u-map development team. The study encompassed the following four phases: (i) A 30-minute tutorial was given on the first day of the study by one of the team developers. The tutorial introduced u-map, its functionality, and the client GUI. The users were instructed to perform *at least 3 VoIP calls per day*, while at FORTH. For that, they had to use the Sipdroid VoIP client with u-map and assign their QoE feedback (score) via the u-map GUI. The score is an integer ranging from 1 to 5 (1 corresponds to a poor performance, while 5 the excellent one). (ii) At the end of the first phase, the subjects also answered the *first questionnaire*. This questionnaire focused on the user background and profile (e.g., gender, familiarity with various technologies, Android usage), preferences (e.g., QoE-vs-cost trade-off), traffic demand (e.g., GSM and VoIP call frequency), and expectations about u-map. (iii) For a *2-week period*, the participants used u-map. (iv) At the end of this two-week period, the subjects answered a second questionnaire to assess u-map. They evaluated various functionalities (e.g., GUI, ability to discover new networks and improve the overall QoE) and also provided an overall score. Moreover, participants indicated the significant factors that affect the performance and future adoption of u-map. Feedback for improving u-map was also collected. During the phases (iii) & (iv), there was no interaction between the subjects and u-map developers. The study was anonymous. Each participant was identified by a unique id generated by the u-map client application. The submitted questionnaires included this user id. The server runs on a Linux machine (CPU: 2.66 GHz Core 2 Duo, RAM: 2048 MB) located at FORTH. Four WiFi networks have been deployed at FORTH, namely the *forth public access*, *forth authorized access*, *eduroam*, and *netlab-QoS-test* (experimental single-AP network). One objective is the evaluation of the functionality of u-map to discover a new network that potentially can improve the QoE of VoIP services. For this, a *new* WiFi network was deployed with SSID *PublicWiFi!*. This network was configured to offer an improved VoIP quality compared to the other networks. It used a MAC address filter to allow only the devices of the participants to connect to its APs. The *Pub-*

licWiFi! had the same number of APs and covered the same area as the *forth public access*. To avoid introducing any bias, we did not indicate/mention the name or the presence of any network to the participants. The monitor was running on a Desktop computer and captured the network traffic of all APs. The collected data, along with the answers to the questionnaires, were analyzed and cross-validated.

The field study produced a rich dataset that includes (a) network measurements collected by the monitor and u-map, (b) users’ QoE feedback about their calls (collected by u-map), and (c) subjects’ questionnaire responses. A *variable* corresponded to a question (in the two questionnaires) and was *numerical* or *categorical* (with no numerical ordering). Most of the numerical variables measured the perception of subjects (about the quality of a certain aspect of u-map) using the Likert scale from 1 to 5, ranging from “No” to “Yes”. They were encoded as discrete numerical variables. Few numerical variables corresponded to questions about the frequency of performing an activity. Other multiple-choice questions were encoded as categorical variables. Finally, three questions required text: two questions required the names of the networks the subjects used, while the other one the characterization of the user familiarity with the WiFi configuration. The questionnaires were submitted in paper form. Prior to the analysis, we sanitized/pre-processed (when necessary) and encoded the questionnaire answers based on their type (e.g., numerical, categorical, binary, string) and domain (e.g., range).¹ Often measurement studies suffer from *missing values*. In general, the methodology for treating missing values depends on the specific objective and may introduce bias, leading to erroneous results.² The *k-nearest neighbors imputation* considers the statistical dependencies between variables to predict the missing values and is used in modeling the user satisfaction. However, the *k-nearest neighbors imputation* method is not suitable when categorical variables exist in the data. So, in the case of user profiling using clustering algorithms (where categorical variables are present), the missing values were substituted with the median of the non-missing values of the respective variable. The treatment of the “I don’t know” or “I didn’t understand the question” answers is a related issue; for the general statistics of our subject population these answers are considered valid, while missing values are ignored.

¹An example of pre-processing is for the case of a question about the significance of various factors for the use of u-map, where subjects are asked to assign a weight to each factor, with the weights summing up to 100. We normalized these weights when the sum was a different number. In another question where users had to provide the WiFi network names in a “free”-text form, only obvious spelling mistakes were corrected.

²For example, in some cases ignoring the missing values or substituting them with a mean of the corresponding variable could be a valid approach.

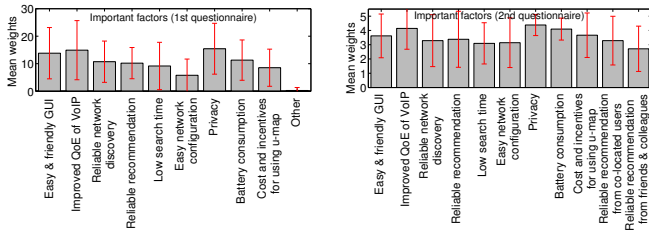


Figure 1: Factors affecting the u-map use.

To assess the statistical (in)dependence between various parameters, we employed the χ^2 test of independence, the Spearman’s rank correlation coefficient (for ordinal variables), and a G -test, based on a standard Monte-Carlo permutation procedure [26]. An advantage of this G -test is its robustness for small sample sizes.

To understand our user population, we computed some general statistics based on their feedback from the first questionnaire. Most of the subjects have a substantial level of familiarity with Android or are experts (71%). Similarly, a large percentage of users can configure their wireless interface without problems (48 %). Subjects indicated the important factors that affect the use of u-map (or a related system) with their answers in both questionnaires. The questions were phrased in a slightly different manner. Specifically, in the first questionnaire, subjects indicated the factors, specifying also their significance level (*weight*). The top-3 factors were: the “Privacy”, “Improved QoE of VoIP”, and “Easy & friendly GUI” (Fig. 1 (left)). In the 2nd questionnaire, the subjects were asked to classify these factors in a scale from 1 to 5, where 1 means “not significant” and 5 means “very significant”. There, the top-3 factors were: the “Privacy”, “Improved QoE of VoIP”, and “Battery consumption” (Fig. 1 (right)). Interestingly, the subjects had changed their opinion about the significance level of these factors: the “Battery consumption” became even more significant, reaching the third position, while the “Easy & friendly GUI” was downgraded to the fifth position. The “Privacy” steadily remained the most significant factor. The “Low search time” was not one of the important factors (Fig. 1). During the study, the load of the server was low, resulting in small response delays, which did not affect the user experience. This is consistent with the fact that users evaluated the responsiveness of u-map positively, regardless of their *responsiveness tolerance threshold*. Subjects indicated their delay/responsiveness tolerance threshold when using any Internet retrieving-based service/application at the first questionnaire. We assessed the statistical (in)dependence between the responsiveness of u-map and the overall responsiveness threshold. The χ^2 , the Spearman’s correlation, and the permutation-based test report p-values of 0.23, 0.11, and 0.25, respectively, indicating a statistical independence. Interestingly while many participants reported that they did notice an increased battery consumption (2nd questionnaire), it seems that this did not impact negatively their overall evaluation. With p-

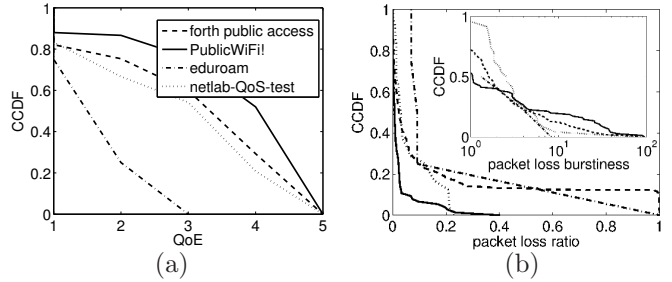


Figure 2: Performance of WiFi networks, in terms of QoE, packet loss ratio & burstiness of packet loss ((b) has the same legend as (a)).

values of 0.55 (χ^2 test), 0.33 (Spearman’s correlation), or 0.51 (permutation-based test), the two variables are considered independent. We speculate that most of the participants were able to alleviate the negative effect of the increased power consumption by performing some sort of “power management” (e.g., deactivating the WiFi interface).³

The study showed that the u-map can assist in the discovery of new networks. When a new network becomes available in a region, a u-map user can eventually discover it and connect to it.⁴ The measurements and QoE feedback collected from this (first) user were uploaded to the server, and the name of this new network was then included in the u-map recommendations for that region. From that point on, the entire community of u-map users was informed about the performance of this new network. In this study, the first Sipdroid call using the *PublicWiFi!* network took place during the fourth day of the study. Until the end of the study, 6 out of 21 users performed VoIP calls using this network. We found in various traces that low QoE scores are correlated with high packet loss ratio and burstiness. It was evident that by discovering *PublicWiFi!*, several users improved their QoE. *PublicWiFi!* performed better than the other networks, as it can be verified by the collected network measurements and QoE statistics. Fig. 2 shows the complementary CDFs (CCDFs) of the performance of all WiFi networks from which Sipdroid calls were made. Due to its light traffic, the *PublicWiFi!* network could be beneficial to all the participants. However, only some users took advantage of its presence. We asked the users to report (in a related question of the 2nd questionnaire) whether they noticed a difference in the QoE of their VoIP calls over the different networks. Then, we compared these responses with the QoE scores that the users assigned to their VoIP calls (that may have taken place in different

³Unfortunately we have not collected appropriate data to validate this hypothesis. Note that a typical user may not be able to perform such a task, and thus, an automated dynamic power management becomes important.

⁴The discovery of this new network was manual in our study but in a more general setting, it could take place through announcements or friends.

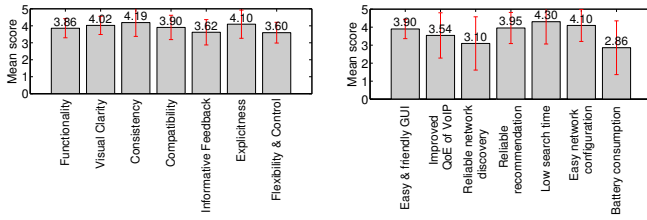


Figure 3: Evaluation GUI (left) & scores of user satisfaction criteria (right).

networks). Indeed, the users who indicated that they perceived a better quality of VoIP when using a specific network had also provided higher scores for the QoE of the VoIP calls made from *PublicWiFi!*, compared to the calls made from *forth public access*. Similarly, the stochastic difference between the QoE scores assigned to the calls performed through different networks by the users who did not perceive any difference on the QoE of their calls is less prominent. Although several participants agreed that the improvement of QoE is important, they evaluated the “reliable network discovery” and “easy network configuration” as less significant factors (Fig. 1). Users were interested in increasing the quality of their experience through u-map without having to perform the network discovery and configuration themselves. This phenomenon was also consistent with what was observed via the u-map traces: there were users that despite their low QoE did not change network. This could be due to their unwillingness to “navigate” through the network discovery and configuration GUI or their privacy concerns about a new and unknown network (e.g., *PublicWiFi!*). This is an example of the trade-off between privacy protection and the automation of the network discovery and configuration process (i.e., reducing the user intervention by making the calls an almost “1-click” operation). The evaluation of the GUI focused on the *functionality, visual clarity, consistency, compatibility, informative feedback, explicitness, and flexibility & control* of the u-map. The users evaluated these aspects in the 2nd questionnaire using the Likert scale. The mean overall score for the GUI is 3.90 out of 5, while the mean score for each aspect is shown in Fig. 3 (a).

User profiling: To distinguish different user groups with respect to their gender, demand, expectations about u-map etc, we applied an extended version of the K-means clustering algorithm [27] on a subset of the collected data, which better describes the user profile. This subset contains answers to the questionnaires and network traffic statistics. The specific clustering algorithm handles *mixed categorical and numerical* datasets.⁵ In both cases, the “weight” indicates the significance in the

⁵Apart from these parameters, the frequency of VoIP and GSM calls, the usage of recommendation systems, and the gender were also included in the clustering. However, the analysis did not report any interesting trend, and thus these metrics were omitted.

determination of user profiles. The impact of the overall evaluation of u-map and the traffic demand are of significant importance. Also, the preference between QoE and price was not found to be significant, since all used WiFi networks were free of charge. Under different conditions, this factor should have had higher importance. The clustering algorithm reports three clusters of 4, 6, and 11, users, respectively. With respect to the evaluation of u-map, users of cluster 1 gave the highest score and were more willing to adopt u-map compared to users of other clusters. They also assigned the highest score to the provider selection capability. This is in accordance with the high score for u-map, as provider selection is one of the key features of u-map. Furthermore, users of clusters 1 and 3 expected that the u-map would have a positive impact on the quality of VoIP calls, while users of cluster 2 were more skeptical. Users of cluster 3 evaluated poorly the quality of VoIP calls on Android. This belief motivated them more to adopt u-map. On the other hand, users of cluster 2 gave a relatively high score to the quality of VoIP calls on Android. This belief in conjunction with their expectation that u-map would not significantly improve the quality of VoIP calls explains why they provided the lowest score for u-map compared to other clusters. Another interesting result for users of cluster 2 is that their majority has a higher level of familiarity with Android compared to users of other clusters. It seems that users with low familiarity with Android provide a higher overall score compared to users with high familiarity. We speculate that the u-map can improve the experience of users that are not very familiar with new technologies. Finally, users of clusters 1 and 3 have the highest and lowest levels of traffic demand, respectively, while users of cluster 2 are characterized by a medium level of demand.

Modeling the user satisfaction about u-map:

In general, the user satisfaction about a system can be parameterized by the perception of the user about the quality of its services and functionalities, given his/her requirements. Requirements exist with respect to the battery consumption, responsiveness, GUI, security, fault-tolerance, and scalability. In this work, we focus on the battery consumption, responsiveness and GUI. In the 2nd questionnaire, the participants were asked to assess the u-map services, namely the *improvement in the QoE of VoIP, reliability of network discovery, reliability in network recommendation, and easiness in network configuration*. Also, the subjects assessed their satisfaction about u-map, providing an overall score and an indication about their intention for future adoption of u-map (Figs. 3 (b) & (d)). We model the user satisfaction as a utility function based on the above parameters and their significance level (following the approach by Grigoroudis *et al.* [28]). The user satisfaction is modeled as a weighted sum of criteria based on user feedback. These criteria, with their respective score and significance weight, are presented in Fig. 3 (b) and

Fig. 1 (right). Fig. 3 (c) shows the mean opinion score of users, per criterion. 33 questions are involved in the generation of the utility: 20 of them evaluate the GUI, 6 of them the remaining criteria, and 7 others provide the criteria weights. For the evaluation of the GUI and the assessment of the user satisfaction, the missing values needed to be filled, and the *k-nearest neighbors imputation* method was used. The predicted user satisfaction of u-map is consistent with the overall score provided by users. The mean user satisfaction is 3.65 (out of 5) compared to the mean overall score of 3.58. The proposed model accurately predicts the overall score: for 10 out of 21 users, the model estimates the exact score, while for 7 other users the absolute error is 1 (in a range from 1 to 5). An extended performance analysis and more detailed discussion of its results can be found in [22]. The questionnaires are available at [23].

4. CONCLUSIONS AND FUTURE WORK

The outcome of the analysis is encouraging. Most of the participants indicated that u-map can improve their experience, helps in the discovery of networks that offer improved QoS. During the study, 30% of users discovered the best available network, improving substantially their experience. Users were also satisfied with the GUI of u-map. The presence of users that despite the low QoE of their calls did not change network motivates us to jointly consider the problem of improving the automation of the network discovery and configuration process, addressing the privacy concern in joining a new network. Finally, the users evaluated positively the overall performance as well as specific features of u-map. Based on user feedback, we modeled the user satisfaction about u-map as a utility function. The model can predict the overall score of the user satisfaction about u-map accurately. A larger scale field study will be performed soon and the prototype will become available to the community. We have been also developing utility functions that incorporate various techno-economic terms, such as price, willingness to pay, preference of price over QoE, and intrinsic indicators towards a service provider (e.g., its brand name, perceived value and reliability). Furthermore, we plan to “cross-validate” the feedback that a client uploads in the database with respect to erroneous and missing values before using it for the u-map recommendations. This will enhance the reliability of the recommendations, and thus, the value of its service. Finally, this work sets a methodological basis for evaluating such recommendation systems in access markets.

References

[1] “OpenSignal: 3G & 4GLTE Cell Map,” <http://opensignal.com/>.
 [2] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, “Senseless: A database-driven white spaces network,” in *IEEE DySPAN*, 2011.
 [3] M. Mishra and A. Sahai, “How much white space is there?” EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-3.

[4] J. van de Beek, J. Riihijarvi, A. Achtzehn, and P. Mahonen, “UHF white space in Europe - a quantitative study into the potential of the 470 - 790 MHz band,” in *IEEE DySPAN*, 2011.
 [5] A. Min, K.-H. Kim, and K. Shin, “Robust cooperative sensing via state estimation in cognitive radio networks,” in *IEEE DySPAN*, 2011.
 [6] D. Gurney, G. Buchwald, L. Ecklund, S. Kuffner, and J. Grosspietsch, “Geo-location database techniques for incumbent protection in the TV white space,” in *IEEE DySPAN*, 2008.
 [7] S. Gaonkar, J. Li, R. R. Choudhury, L. Cox, and A. Schmidt, “Micro-blog: sharing and querying content through mobile phones and social participation,” in *ACM MobiSys*, 2008.
 [8] J. Froehlich, M. Y. Chen, S. Consolvo, B. Harrison, and J. A. Landay, “Myexperience: a system for in situ tracing and capturing of user feedback on mobile phones,” in *ACM MobiSys*, 2007.
 [9] “Sensorly: Coverage Maps,” <http://www.sensorly.com/>.
 [10] A. Farshad, M. K. Marina, and F. Garcia, “Urban wifi characterization via mobile crowdsensing,” in *IEEE NOMS*. IEEE, 2014.
 [11] E. Gregori, L. Lenzini, V. Luconi, and A. Vecchio, “Sensing the internet through crowdsourcing,” in *IEEE Per-Moby*, 2013.
 [12] S. Sen, J. Yoon, J. Hare, J. Ormont, and S. Banerjee, “Can they hear me now?: a case for a client-assisted approach to monitoring wide-area wireless networks,” in *ACM IMC*, 2011.
 [13] S. Rosen, S.-J. Lee, J. Lee, P. Congdon, Z. Morley Mao, and K. Burden, “MCNet: Crowdsourcing wireless performance measurements through the eyes of mobile devices,” *IEEE Commun. Mag.*, 2014.
 [14] “WiFi,” <http://www.wifi.com/>.
 [15] X. Wei, L. Gomez, I. Neamtiu, and M. Faloutsos, “Profiledroid: Multi-layer profiling of android applications,” in *ACM MobiCom*, 2012.
 [16] H. Falaki, R. Mahajan, and D. Estrin, “Systemsens: a tool for monitoring usage in smartphone research deployments,” in *ACM MobiArch*, 2011.
 [17] W. Enck, P. Gilbert, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. Sheth, “Taintdroid: An information-flow tracking system for realtime privacy monitoring on smartphones,” in *OSDI*, 2010.
 [18] C. Shepard, A. Rahmati, C. Tossell, L. Zhong, and P. Kortum, “Livelab: measuring wireless networks and smartphone users in the field,” *ACM SIGMETRICS Perf. Eval. Review*, 2011.
 [19] F. Qian, Z. Wang, A. Gerber, Z. Mao, S. Sen, and O. Spatscheck, “Profiling resource usage for mobile applications: a cross-layer approach,” in *ACM MobiSys*, 2011.
 [20] G. Fortetsanakis, M. Katsarakis, M. Plakia, N. Syntychakis, and M. Papadopoulou, “Supporting wireless access markets with a user-centric qoe-based geo-database,” in *ACM MobiArch*, 2012.
 [21] “Who is afraid of active users? the u-map as a catalyst of wireless access markets,” <https://www.youtube.com/watch?v=k34S4okFMao>, in *Student Video Contest, DySPAN 2012*.
 [22] M. Katsarakis, V. Theodosiadis, and M. Papadopoulou, “On the evaluation of a user-centric qoe-based recommendation tool for wireless access,” *ICS-FORTH, Heraklion, Crete, Greece, tech. rep.*, vol. 445, 2014.
 [23] “u-map,” <http://www.ics.forth.gr/mobile/umap/>.
 [24] B. A. Kitchenham, S. L. Pfleeger, L. M. Pickard, P. W. Jones, D. C. Hoaglin, K. El Emam, and J. Rosenberg, “Preliminary guidelines for empirical research in software engineering,” *IEEE TSE*, 2002.
 [25] L. Wilkinson, “Statistical methods in psychology journals: guidelines and explanations,” *American psychologist*, 1999.
 [26] I. Tsamardinos and G. Borboudakis, “Permutation testing improves bayesian network learning,” in *ECML PKDD*, 2010.
 [27] A. Ahmad and L. Dey, “A k-mean clustering algorithm for mixed numeric and categorical data,” *Data & Knowledge Engry*, 2007.
 [28] E. Grigoroudis and Y. Siskos, “Musa: A decision support system for evaluating and analysing customer satisfaction,” in *PCI*, 2003.