

Understanding Portal-Based Mobile Search: a Case Study

Karen Church and Nuria Oliver
Telefonica Research

Torre Telefonica - Diagonal 00, Pl. Ernest Lluch i Marti 5, 08019 Barcelona, Spain
karen@tid.es, nuriao@tid.es

ABSTRACT

In this paper we present a real-world case study aimed at understanding mobile search behaviour in the context of an operators portal-based mobile search service. We describe key results of a large-scale logs-based analysis of the queries issued via the operator's mobile portal. Using these results, we highlight implications for the design of portal-based mobile search with a view to improving the search experiences of mobile customers. We conclude we a number of insights into the challenges of dealing with large-scale logs analysis as a means of understanding real-world mobile behaviours.

Author Keywords

Mobile search, mobile Web, query-log analysis, portal-based search, user behaviour, search behaviour

ACM Classification Keywords

H.3.3 Information Storage and Retrieval: Information Search and Retrieval.

General Terms

Design, Experimentation, Human Factors

INTRODUCTION

The mobile world is continuing to grow and evolve at a phenomenal rate. Thanks to significant advances in handset technology, improvements in mobile networks and increased popularity of social mobile services, an increasing number of mobile users are accessing and searching for online content while on-the-move¹. In terms of mobile Internet access, the operators portal is the gateway to mobile content for millions of users, in particular for users of less sophisticated mobile handsets with more limited browsing and interaction capabilities. The focus of this paper is to showcase how analyzing and understanding the search habits of these average mobile Internet users can lead to a set of viable implications for design that may be used to develop richer mobile Web

¹comScore. The 2010 Europe Digital Year in Review. See: <http://bit.ly/hfuCFG>

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. To copy otherwise, to publish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Large'11, September 18, 2011, Beijing, China.

Copyright 2011 ACM 978-1-4503-0924-0/11/09...\$10.00.

experiences tailored to the unique needs of portal-based mobile search users.

To this end, we carried out an analysis of a set of query logs generated by users of an operators mobile Internet portal. Specifically, we analyzed approx. 1.4 million Spanish queries issued by approx. 235,000 unique European mobile subscribers in Spain over a 3-month period in late 2009. Our analysis focuses on query patterns, search topics and query intent as well as the devices used to access the mobile portal. Our results help shed light on new trends in portal-based mobile search behaviours and help to inform a set of implications that we believe could help enrich the online experiences of portal-based mobile search users. We conclude with a discussion of open research questions and challenges brought forth by the large-scale query log analysis.

RELATED WORK

Recently, there has been significant effort to understand the behaviours and motivations of mobile users and their Web accesses [8, 9]. Specifically, query log analysis has been used to investigate how users access and use mobile search services. One such body of work concerns commercial mobile search engines such as Google and Yahoo!. Results of this research has shown that mobile users tend to execute short queries, across a set of niche topics such as adult content and they tend to search for similar items to one another [5, 6, 1, 10]. The most recent Google study shows that mobile search is evolving in particular for users of advanced handsets such as Apples iPhone [7].

The focus of this paper is on portal-based mobile search behaviour and previous work in this area has shown that users of operator-specific mobile Web portals exhibit more stringent search behaviours. These studies focused mainly on Europe and show that portal-based queries are shorter than queries issued from mobile search engines, interest in niche content tends to be heightened and there are low volumes of diversity within the queries executed by end-users [4, 3].

PORTAL-BASED MOBILE SEARCH

A carrier-owned (i.e. operator's) mobile portal is an entry point for accessing content/services on the Internet via a mobile device. Comparable to early Web portals (e.g. Yahoo!), mobile portals provide a gateway to content, communications services and software applications. In the early days of the mobile Web, operators mobile portals were the dominant form of information access for mobile users. However, with the advent of more advanced handsets with more so-

phisticated Internet-capabilities, users began to venture beyond mobile portals and towards general mobile Web content accessed directly through the browser (in a similar way to PC based web access). However despite this trend, feature phones still dominate in many parts of the world. For example a recent report by comScore shows that feature phone usage in Europe was almost 70% in 2010. For many of these users, carrier-owned operator portals still remain the gateway to the mobile Web. Given the scale of this user group, we think it's important to understand their search behaviours so that we may devise new services and solutions to enrich their mobile Web experiences. Figure 1 shows a sample of the operator's portal in 2009 with options for the user to search for on-portal and off-portal (i.e. general mobile Web) content using one of two search boxes.² Users of these portals issue queries and receive a set of on-portal results as well as external search results (in this case extracted from Yahoo!).



Figure 1. An example of the mobile operator portal in 2009

DATA & ANALYSIS

The data used in this study corresponds to approx. 1.4 million search requests issued by approx. 235,000 unique European mobile subscribers in Spain³ over a 3 month period in late 2009. These searches were issued via the operators portal and the search results were obtained from an operator-specific search service. Each search request is made up of the following fields:

²Note that the screenshot does not show the entire mobile browser page. That is only a screenshot of the mobile portal is shown. We cannot see the browsers URL input box for example.

³Given that the queries originate from Spain, the queries were issued in Spanish.

1. *User ID*: A unique (anonymous) identifier for each user.
2. *Timestamp*: The date/time (measured in hours, minutes, seconds and milliseconds) for each search request.
3. *Query string*: The actual query/search keywords issued by the user.
4. *User agent string*: The user agent string which corresponds to the device used by the users to issue the query.

A session is defined as a sequence of actions performed by the same user within some continuous period of time. In our study, as with similar mobile search studies, we identify sessions as a sequence of search requests made by a unique user delimited by a 5 minute period of inactivity. Using this approach, we identified over 800,000 individual search sessions within the dataset.

KEY QUERY-LOG ANALYSIS RESULTS

In this section we present key results including the nature of mobile search queries, shifts in search topics, classification of query intent and an analysis of the devices used during these portal-based mobile search sessions.

Query Length

We find the average number of words per query is 1.6 (min=1, max=54, standard deviation=1.12), while the average number of characters per query is just 10.6 (min=1, max=256, standard deviation=6.9). This is significantly lower than the query lengths found in previous mobile studies of commercial mobile search engines, which typically range from approx. 2.3-2.6 terms per query [5, 6, 7, 1, 10]. Furthermore, we found even shorter query lengths when compared to previous mobile operator studies (2.1-2.3 terms) [4, 3]. When we only consider *unique queries*, the average query lengths increase significantly. For example, the average length per unique query is 2.5 (standard deviation=1.5), while the average number of characters per unique query is 16.8 (standard deviation=8.5). This sharp increase in query lengths over unique queries implies that popular queries, i.e. queries with the highest frequency, are shorter than average in their length when compared to less popular queries, i.e. the queries that lie in the *long tail*.

Search Topics

To carry out our analysis of search topics, we manually classified the top 1000 queries⁴ using the set of categories outlined in previous portal-based mobile search studies [4, 3]. Our results show significant shift in search topics within our dataset when compared to previous mobile studies. For example, searches for adult content have decreased significantly. We find almost 14% of the top 1000 queries falling into this category. Church et al. previously found that adult-content was much more prevalent in portal-based mobile search, representing in excess of 50% of top 500 queries [3]. While the most recent Google study highlighted an 11.6% increase in adult content in mobile web search [7].

The most popular search topic is “*Socializing/Dating*” with over 40% of the top 1000 queries falling into this category.

⁴The top 1000 queries account for approx. 55% of total queries in this study.

The vast majority of these queries include searches for social networking websites, e.g. “facebook” and “hi5.com”. If we compare our results with the previous mobile operator studies we find much higher percentages of “socializing/dating” queries in our dataset. This may be due to the fact that social networking services, in particular mobile access to such services, has increased dramatically in the past 2 years. For example, a report by comScore shows that social networking was the fastest growing mobile content category in 2010⁵. Overall we find that searches issued via operators portals have shifted away from adult content and towards social networking services.

Category	% Top 1000 Queries
Socializing/Dating	40.3
Search/Finding things	14.4
Adult	13.9
Email/IM	8.0
Multimedia	6.1

Table 1. Top 5 search categories showing the percentage of the top 1000 most popular queries associated with each category.

Query Intent

Broder [2] previously classified Web queries according to intent using three taxonomies: (1) *Navigational* queries refer to a class of queries where the immediate intent is to reach a particular site, e.g. queries that contain domain suffixes (.com, .net, etc), queries that refer to single business or service names, e.g. facebook. (2) *Informational* queries are those that involve the user attempting to find information online, e.g. the user is probably interested in reading some information online but no further interaction is expected. (3) *Transactional* queries refer to queries where the purpose is to visit a site to carry out additional interactions, e.g. shopping, downloading files, gaming, etc⁶. Our results show that query intents have shifted considerably. Over 70% of the top 1000 queries were classified as navigational, that is, the users intent was to reach a single website or point for interest (e.g. facebook, google.es, etc). Previous mobile search studies report much lower occurrences of navigational queries (5.0-29.4%)[5, 10, 4].

Query Intent	% Top 1000 Queries
Navigational	71.1
Transactional	20.8
Informational	2.2
Unknown	5.9

Table 2. Query intent among the top 1000 most popular queries.

Query Diversity

Query diversity refers to how much difference exists between the various queries issues by mobile users. If we find high levels of query diversity this shows that users have very different needs and they execute different queries to address those needs. If we find low levels of query diversity this shows that users are quite similar to each other in that they

⁵See comScore report, <http://bit.ly/b3pdwJ>, last checked June 2011

⁶We group all adult-related queries into this category.

execute similar queries. To measure query diversity we extract the top N queries and calculate what percentage of the total query volume these N most popular queries represent (independent of case). (See Table 3) Our results show low levels of query diversity, i.e. our users appear to search for very similar items to one another. The top 1000 queries account for >55% of all queries, while the top query accounts for a very significant 18.4% of all queries analyzed. Previous mobile studies have shown more diversity in their datasets. For example, the top query in the latest Google study accounts for only 3.8% of the top 1000 queries.

Study	Top Query	Top 1K Queries
Google 2006[5]	1.2%	22%
Google 2007[6]	0.8%	17%
Google 2009[7]	3.8%	32.8%
European operator 2007[4]	2%	26%
<i>Our study</i>	18.4%	55.2%

Table 3. Comparison with previous studies showing the percentage of the total queries accounted for by top queries.

A Note on Mobile Devices and the UI

Although we have not focused explicitly on this point, factors that are likely to impact on end user behaviour relate to (1) the handsets used to access mobile content and (2) the UI provided to support end-user interaction. On analyzing the top 10 mobile handsets in our dataset (which generated 34% of all queries), we find that only 3 support touch-based interaction, with only one of these handsets also providing a full physical QWERTY keyboard. This implies that portal-based mobile search users (most of whom do not use advanced smartphones) are more likely to have more difficulties in accessing and interacting with online content than users of more advanced handsets (given the inherent input and interaction limitations of their handsets). It would be interesting to gain some subjective user input regarding end-user experiences with the existing portal design/layout to determine what improvements can be made and to what extent the UI has impacted or influenced the user behaviour we have witnessed. This is something that we could not explore on this occasion, but it is something we would like to investigate as part of future work.

IMPLICATIONS FOR DESIGN

In this section we highlight two implications in designing future portal-based mobile search services.

Personalization of on-portal mobile search

Our results show that our users search for very similar items to each other, execute the same queries repeatedly (almost 80% repeat queries in our dataset) and execute high volumes of navigational queries. As such, there is an opportunity for personalizing the search experience for mobile users. Simple improvements at the interface level, such as including links to top queries on the main portal page, would probably yield increased customer satisfaction. We could automatically analyze query string input to assess the users intent and take appropriate actions. For example, if the query is navigational,

the user could be redirected to the intended website. Aggregated search histories could also be exploited to provide enhanced query recommendation. Note that even though some of these enhancements have already been included (at least in part) in existing mobile search engines (e.g. Google), they are not part of standard mobile operators portals.

Enriching the Global Mobile Experience

Another interesting implication of our findings lies in the potential to use the search history of individuals and groups to enrich the global mobile experience. For example, the queries issued via the mobile portal could be used to proactively recommend novel and interesting native mobile applications to the end-user, thus shifting behaviours towards custom native applications designed for a specific purpose. The high frequency of navigational queries and of repeat queries indicates that many mobile users use their operators portal as a stepping-stone to content of interest. Recommending native applications that satisfy user needs based on large scale data sources such as query logs could really enhance the overall mobile experiences of our users.

CHALLENGES AND OPEN RESEARCH QUESTIONS

Our query log analysis provided us with a large-scale trace of real user behaviour, that is the behaviour was not recalled nor was it subjective. One of the benefits of query log analysis is that it provides a complete and accurate picture of user behaviour. The queries we analyzed were actual queries executed by real users of a live, fully functional operators mobile portal. As such we were able to shed significant light on the behaviours of our users. However, it should be noted that such analyses also bring forth some drawbacks. We may be able to analyze *what* our users search for via the portal, but we have no idea *why* they executed those searches. That is, nothing is known about the actual intent or motivations of the end user. We also know very little about the users context. And despite efforts to analysis the dataset and report on results, it's easy for researchers to make assumptions about why certain behaviours or trends emerge.

We believe the answer to this dilemma is to combine these large-scale data sources with more subjective input from the end-user (either explicit or implicit) to help shed further light on the motivations and contexts surrounding mobile search patterns. Involving the user in the loop either through diary studies, contextual inquiries, or experience sampling and combining these findings with an analysis of their actual search patterns (via query-log analysis) would result in a more complete picture of mobile search patterns. We believe the next challenge facing researchers in this space is approaching how we can design such studies (that combine large-scale logs analysis with subjective input from end-users) and how we analyze the data generated in an efficient and unbiased manner.

CONCLUSIONS

The focus of this work has been on understanding the search patterns of portal-based mobile search users, many of whom are constrained by the inherent limitations of their mobile handsets. We highlighted key results of a query log analysis using a large dataset generated in Spain. Furthermore

we discussed a number of implications in the design of future mobile Internet experiences and discussed a number of open research questions and challenges emerging from this work. As part of future work we would like to explore the combination of datasets from multiple, different portals, to determine if we can generalize and validate our findings for portal-based mobile search. In addition we would like to combine large-scale logs analysis with subjective input from end-users to determine if we can answer more about *why* mobile users search in this manner.

ACKNOWLEDGMENTS

This work is funded as part of a Marie Curie Intra European Fellowship for Career Development (IEF) award held by Karen Church.

REFERENCES

1. R. Baeza-Yates and J. Velasco. A study of mobile search queries in japan. In *Query Log Analysis: Social and Technological Challenges, WWW 2007 Workshop*, 2007.
2. A. Broder. A taxonomy of web search. *SIGIR Forum*, 36(2):3–10, 2002.
3. K. Church, B. Smyth, K. Bradley, and P. Cotter. A large scale study of european mobile search behaviour. In *Proceedings of MobileHCI '08*, pages 13–22. ACM, 2008.
4. K. Church, B. Smyth, P. Cotter, and K. Bradley. Mobile information access: A study of emerging search behavior on the mobile internet. *ACM Transactions on the Web*, 1(1):4, 2007.
5. M. Kamvar and S. Baluja. A large scale study of wireless search behavior: Google mobile search. In *Proceedings of CHI'06*, pages 701–709. ACM, 2006.
6. M. Kamvar and S. Baluja. Deciphering trends in mobile search. *Computer*, 40(8):58–62, 2007.
7. M. Kamvar, M. Kellar, R. Patel, and Y. Xu. Computers and iphones and mobile phones, oh my!: a logs-based comparison of search users on different devices. In *Proceedings of WWW'09*, pages 801–810. ACM, 2009.
8. S. Nylander, T. Lundquist, and A. Brännström. At home and with computer access: why and where people use cell phones to access the internet. In *Proceedings of CHI'09*, pages 1639–1642. ACM, 2009.
9. C. A. Taylor, O. Anicello, S. Somohano, N. Samuels, L. Whitaker, and J. A. Ramey. A framework for understanding mobile internet motivations and behaviors. In *Proceedings of CHI'08 extended abstracts*, pages 2679–2684. ACM, 2008.
10. J. Yi, F. Maghoul, and J. Pedersen. Deciphering mobile search patterns: a study of yahoo! mobile search queries. In *Proceedings of WWW'08*, pages 257–266. ACM, 2008.