Mobile Search Pattern Evolution: The Trend and the Impact of Voice Queries

Jeonghee Yi* Farzin Maghoul
Yahoo! Inc.
701 First Ave, Sunnyvale, CA, 94089, USA
{jeonghee, fmaghoul}@yahoo-inc.com

ABSTRACT

In this paper we study the characteristics of search queries submitted from mobile devices using `Yahoo! Search for Mobile` during a 2 months period in early of 2010, and compare the results with a similar study conducted in late 2007. The major findings include 1) mobile search queries have become much more diverse, and 2) user interest and information needs have been substantially changed at least in some areas of search topics, including adult and local intent queries. In addition we investigate the impact of voice query search interface offered by Yahoo!'s mobile search service. We examine how unstructured spoken queries differ from conventional search queries.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; H.3.0 [**Information Storage and Retrieval**]: Genearl—*Web search*; A.1 [**General Literature**]: introductory and survey

Keywords

mobile search query analysis, mobile search, query log analysis, mobile queries, voice queries, query categorization

1. INTRODUCTION

Wireless mobile subscribers have grown explosively and wireless services have increasingly become a part of our daily lives. The number of mobile phones in the US has reached to over 285M in 2009[1], and 4.6B worldwide[2].

We have studied search query log data of various Yahoo! mobile search applications, in order to understand users' information needs and demands for mobile data access and to better serve them. While mobile search is still in its infancy, it has been going through rapid evolution over the last couple of years.

The objectives of this study include:

- investigating the current mobile search query patterns and user interests to find the trend of change
- testing the hypothesis that mobile search queries are following the similar evolutionary trajectory of the PC_ based web search queries in its early days
- studying voice search query patterns on mobile devices, and investigating their impact on the evolution of mobile search.

In the next section, we present the change of mobile search query patterns and compare with the evolutionary pattern of web search queries in their early days. Section 3 presents the analysis results on voice queries and the result is compared with overall mobile search queries and web search queries from PCs. We provide an overview of related work in section 4, and conclude with discussion in section 5.

Table 1: US Mobile Query Categorization

	2010 Samples		2007 Samples	
Categories	% of unique queries	% of all queries	% of unique queries	% of all queries
Arts & Humanities	<1%	<1%	<1%	<1%
Automotive	4%	2%	2%	1%
Consumer Goods	5%	3%	2%	2%
Entertainment	19%	13%	44%	51%
Finance	3%	2%	1%	1%
Government&Politics	<1%	<1%	1%	<1%
Health & Pharma	5%	3%	2%	1%
Hobbies	<1%	<1%	<1%	<1%
International Interest	<1%	<1%	<1%	<1%
Life Stages	3%	2%	2%	1%
Local	14%	15%	7%	7%
Miscellaneous	3%	2%	2%	2%
News	3%	2%	2%	2%
People	3%	2%	3%	5%
Reference	2%	1%	1%	<1%
Religion	1%	<1%	1%	1%
Retail	9%	6%	5%	4%
Science	<1%	<1%	1%	1%
Small Business	3%	1%	2%	1%
Sports	5%	3%	3%	3%
Electronic Gadgets	3%	16%	6%	7%
Telecommunications	<1%	<1%	2%	2%
Uncategorized	12%	23%	12%	9%

2. MOBILE SEARCH QUERY TREND

Sample Data Set: 20 million non-empty English US mobile search queries were randomly selected, with 10 million samples each from March and April 2010, from user search queries submitted to `http://m.yahoo.com` and voice queries from Yahoo!'s mobile voice search interface application.

The query log data we used in the study is strictly anonymized: there is no personal identity data available to match with the request's cookie.

Query pattern and Trend: Mobile search looks to be still in its infancy and the shift of search topics has happened quite rapidly for the last three years of the study period. First, the number of unique queries has increased significantly from 4.49M (2007) to 6M (2010). This indicates mobile queries are becoming more diverse at quite a fast rate.

Second, the average number of words and median character length have decreased slightly (from 2.35 to 2.23, and

*The contact author

[1]CTIA's semi-annual wireless industry survey: http://files.ctia.org/pdf/CTIA_Survey_Midyear_2009_Graphics.pdf

[2]ITU Corporate Annual Report 2009: http://www.itu.int/dms_pub/itu-s/opb/conf/S-CONF-AREP-2008-E06-PDF-E.pdf

WWW 2011, March 28–April 1, 2011, Hyderabad, India.
ACM 978-1-4503-0637-9/11/03.

from 13 to 12, respectively), while the average character length is increased (from 13.73 to 14.66). This means that there are more *head queries* (queries with high frequency of repetitions) with small number of words, while long queries are also increasing on *torso* (queries with smaller number of repetitions) and *tail queries* (queries with very small number of repetitions).

When only unique queries are considered (i.e., with no weight by their repetition), the query lengths and variance both increased quite substantially. The average number of words and characters per unique query is 3.6 and 22.16, respectively, for the 2010 samples. The corresponding numbers were 3.05 and 18.48, respectively, in 2007. The increased lengths of unique queries imply that many head queries are shorter than average, and longer queries lie on tail of the spectrum. Again, this supports the previous hypothesis that mobile query spectrum in terms of their query length is becoming wider.

Third, there has been substantial shift of user search query intent as compared in Table 1. The most interesting change can be found in the drastic decrease of entertainment queries, where most of decrease in entertainment queries is attributed by the decrease of adult queries. Some other notable findings include a substantial increase of local queries (at least twofold) and overall commerce queries (consumer goods, retail, electronic gadgets combined).

Note that these observations are quite similar and consistent to the finding by Spink *et. al* on the shift of web search query topics in its infancy [6]. They found that the most popular search topics in the beginning of the test period (entertainment, recreation, and pornography) decreased a lot in their popularity during the four years of test period, and they were replaced by queries about commerce, travel, employment, economy, people, places, or things. This similarity in search topic shift may indicate that search interest of mobile users are following similar evolutionary path the web searchers have gone through.

3. VOICE QUERIES

We have collected over 79K voice queries at random from the same period of study from queries submitted using the voice query interface. These queries were compared with typed mobile search and PC search queries.

As expected voice queries are considerably longer. Voice queries have more than one word longer than typed mobile queries on average (3.41, 2.23, and 2.78 words, and 21.21, 14.66, and 18.27 characters for voice, typed mobile, and PC search queries, respectively). Interestingly, the maximum number of words or characters are not much different. This is probably because extremely long search queries are often a cut-and-pasted text from elsewhere, which is not quite possible with voice search.

The high rate of increase of query length may not be directly imply the voice queries are proportionally richer, however. First, with voice interface, users tend to pose queries that resemble natural language with many function words than just a sequence of keywords: 'starbucks in chicago' vs 'starbucks chicago'. In fact, some popular prepositions (like '`in`' and '`at`') appear twice as frequently in the voice sample data set as in the other samples. Nevertheless, ease of input seems to make voice queries a lot more descriptive than typed queries.

In terms of search query category, voice queries increased on retail (5.8%), local (2.9%), automotive (2.6%), and finance (1.2%) categories, while decreasing on (electronic gadgets) significantly, in comparison to typed mobile queries. Compared to computer web queries, voice queries increased on local (11.1%) and retail(1.6%), while decreasing on electronic gadgets (4.2%) and Health (1%) categories.

We hypothesize this difference can be explained by the user behavior that mobile users may use voice queries when they have to, yet are distracted and can not type. Some examples of such situation could be where they are lost while driving, or where they need to find and call a place quickly. The increase of local and retail type queries probably reflects such use cases. For electronic and technology gadgets, or health related queries are more of a research type queries. For those types of queries, users often need an environment where they can focus on the research, and in those environment people still seem to prefer to type.

4. RELATED WORK

There have been several large scale studies in the near past on mobile query log analysis for deciphering mobile search query patterns. Kamvar and Baluja studied mobile search query patterns using Google query log data [2, 3, 4]. Yi *et. al.* conducted a large scale mobile search pattern analysis using Yahoo! mobile query logs including comparison between US and international queries, and between various mobile application interfaces [7]. Baeza-Yates *et. al.* reported a large-scale mobile query log analysis on Yahoo! Japan's search logs [1].

Silverstein *et. al.* conducted a large scale query log analysis on web search queries [5] that provides an invaluable reference information for comparison and gives us insight on how the mobile web search might evolve in the near future. Spink *et. al.* investigated Excite search engine queries to find the evolution pattern of web search qeuries.

5. CONCLUSION

Mobile search seems still in its early stage of development and is evolving rapidly. It currently is undergoing major shift of the content type in demand, and query diversity. It appears to be following the evolutionary path of web search queries in the early days where the initial interests in entertainment and adult content decreased drastically and the demand on commerce, travel, economy and business related content increased substantially within a few years of use.

6. REFERENCES

[1] R. Baeza-Yates, G. Dupret, and J. Velasco. A study of mobile search queries in japan. In *WWW*, 2007.

[2] M. Kamvar and S. Baluja. A large scale study of wireless search behavior: Google mobile search. In *ACM SIGCHI*, 2006.

[3] M. Kamvar and S. Baluja. Deciphering trends in mobile search. *Computer*, 40(8):58–62, 2007.

[4] M. Kamvar, M. Kellar, R. Patel, and Y. Xu. Computers and iphones and mobile phones, oh my! In *WWW*, 2009.

[5] C. Silverstein, M. Henzinger, H. Marais, and M. Moricz. Analysis of a very large web search engine query log. *SIGIR Forum*, pages 6–12, 1999.

[6] A. Spink, B. Jansen, D. Wolfram, and T. saracevic. From e-sex to e-commerce: Web search changes. *Computer*, pages 107–109, Mar 2002.

[7] J. Yi, F. Maghoul, and J. Pedersen. Deciphering mobile search patterns: a study of yahoo! mobile search queries. In *WWW*, 2008.