Computers and iPhones and Mobile Phones, oh my!

A logs-based comparison of search users on different devices.

Maryam Kamvar
Google, Inc
mkamvar@google.com

Melanie Kellar
Google, Inc
kellar@google.com

Rajan Patel
Google, Inc
rajan@google.com

Ya Xu[1]
Dept. of Statistics, Stanford University
yax@stanford.edu

ABSTRACT

We present a logs-based comparison of search patterns across three platforms: computers, iPhones and conventional mobile phones. Our goal is to understand how mobile search users differ from computer-based search users, and we focus heavily on the distribution and variability of tasks that users perform from each platform. The results suggest that search usage is much more focused for the average mobile user than for the average computer-based user. However, search behavior on high-end phones resembles computer-based search behavior more so than mobile search behavior. A wide variety of implications follow from these findings. First, there is no single search interface which is suitable for all mobile phones. We suggest that for the higher-end phones, a close integration with the standard computer-based interface (in terms of personalization and available feature set) would be beneficial for the user, since these phones seem to be treated as an extension of the users' computer. For all other phones, there is a huge opportunity for personalizing the search experience for the user's "mobile needs", as these users are likely to repeatedly search for a single type of information need on their phone.

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Search and Retrieval – *query formulation, search process.*

General Terms

Design, Human Factors

Keywords

search, mobile search, user behavior, iPhone, Google

1. INTRODUCTION

Search has become a pervasive part of life in the United States. A recent survey reported that 49% of US Internet users use a search engine on a typical day [6]. Web users are issuing queries not only from computers, but increasingly from mobile devices. A 2008 survey reported that 40% of mobile internet users find the sites they browse on their phones through search [7].

As more mobile devices support rich access to the Web, there will likely be an uptake in search from an increasing variety of devices. In order to improve the search service for all users from any sort of device, we need to understand if and how users' information needs and search patterns vary from each device. The goal of this study is to understand the differences in search patterns across platforms.

WWW 2009, April 20–24, 2009, Madrid, Spain.
ACM 978-1-60558-487-4/09/04.

The unique contributions of this paper are twofold: First, we provide a direct comparison of patterns of search users on multiple search mediums. Other studies of search patterns have generally focused on a single type of search medium, and therefore can only draw indirect comparisons to the other mediums from prior work. Secondly, we present an extensive and controlled comparison of search *users* rather than the aggregate analyses of search queries presented in prior studies.

2. RELATED WORK

There have been several large scale examinations of user search behavior through search engine logs for both computer and mobile search. The results of these analyses have been used to provide insight into areas for improvement in search interfaces.

In one of the first analyses of computer web search behavior, Jansen, Spink, Bateman, and Saracevic [9] analyzed Excite search logs and reported that users' web queries were typically short (avg. words per query = 2.35) and that users did not issue many searches within a session (67% of sessions contained only a single query). In their analysis of 1998 AltaVista search logs, Silverstein, Henzinger, Marais, and Moricz [18] reported similar query metrics; the average number of words per query was 2.35 and query refinement appeared to be even more limited (77% of sessions contained only a single query). The study also suggested that users' information needs on the Web were relatively diverse; unique queries accounted for 63.7% of all queries. Based on their findings, both authors concluded that web searchers differ significantly from users of traditional information retrieval systems.

Spink, Jansen, Wolfram, and Saracevic [20] conducted a longitudinal comparison of the query behavior of Excite web search users between 1997 and 2001. While they reported little change in query statistics (e.g., average number of query terms increased slightly from 2.4 in 1997 to 2.6 in 2001), they did observe a shift in the types of topics for which users were searching. They reported a decrease in the topics "Entertainment or recreation" and "Health or sciences" while there was an increase in the topics "Commerce, travel, employment, or economy" and "People, places, or things". They also reported an increase in the proportion of distinct queries, which accounted for 57.1% of all queries in 1997 and 61.7% in 2001.

Beitzel, Jense, Chowdhury, Grossman, and Frieder [2] conducted a large scale analysis of 2003 AOL web search logs. They reported an average query length of 2.2 words and that the most common query categories were shopping, entertainment, and

[1] Authors are ordered alphabetically and are equal contributors.

porn. Based on a temporal analysis of query characteristics, they found that time of day impacted the popularity of queries as well as the nature of users' search topics.

While most computer web search logs analysis studies were conducted in late 1990 and early 2000, analyses of mobile web search are more recent. This makes it difficult in some cases to make comparisons between the more recent mobile web search analyses and older computer-based web search analyses. There has been little direct comparison to-date of the differences between mobile and computer-based web search patterns.

Kamvar and Baluja [13] conducted one of the first large scale analyses of mobile search logs. Their analysis of two Google mobile search interfaces found that users with less sophisticated input capabilities submitted shorter queries (2.3 words per query vs. 2.7 for PDA-like devices). A topical categorization of mobile queries suggested that Adult content was the most prevalent search topic, followed by Internet & Telecom, and Entertainment. In their 2007 follow up study [14], Kamvar and Baluja reported an average query length of 2.56 words per query. Similar to their previous research, Adult content was the most common search query category (having increased proportionally from the previous study), followed by Entertainment and Internet/Telecommunications. They also reported an increase in the homogeneity of mobile queries.

Baeza-Yates, Dupret, and Velasco [1] conducted a comparison of Yahoo! mobile and computer-based search in Japan. They reported very similar query characteristics in terms of query length; the mean number of words per query was 2.29 for mobile phones and 2.25 for computers. For mobile search, the most common query topics were online shopping, sports, and health while Art, Sports, and online shopping were the most common query topics for computer-based search. Japanese mobile search is considered to be a more mature market, which may be one explanation for the decrease in Adult content.

In their large scale analysis of 2006 European mobile search logs, Church, Smyth, Cotter, and Bradley [4] reported an average query length of 2.2 words per query; however, they reported that the queries submitted to Google were much shorter in nature, averaging 1.5 words per query. An examination of search topics showed an increase in adult content, in comparison to their previous 2005 analysis of European mobile web search [3], as well as the addition of a new topic category representing user generated content.

3. GOOGLE SEARCH INTERFACES

This study compares user search behavior on three different types of devices: the computer, the iPhone, and the conventional mobile phone. Graphical examples of the Google interface on each of these platforms are shown in Figure 1. The user-agent sent in the HTTP request determines which interface is shown on the device.

We define searches issued on an iPhone as searches issued from any iPhone device (excluding iPods). Mobile searches are defined as searches sent from a non-iPhone mobile phone and which originate from the www.google.com/m property. Computer searches are defined as all other searches (which are predominantly from desktop and laptop computers) that were issued from the main www.google.com property. We use the terms "computer searches" and "desktop searches" interchangeably in this paper.

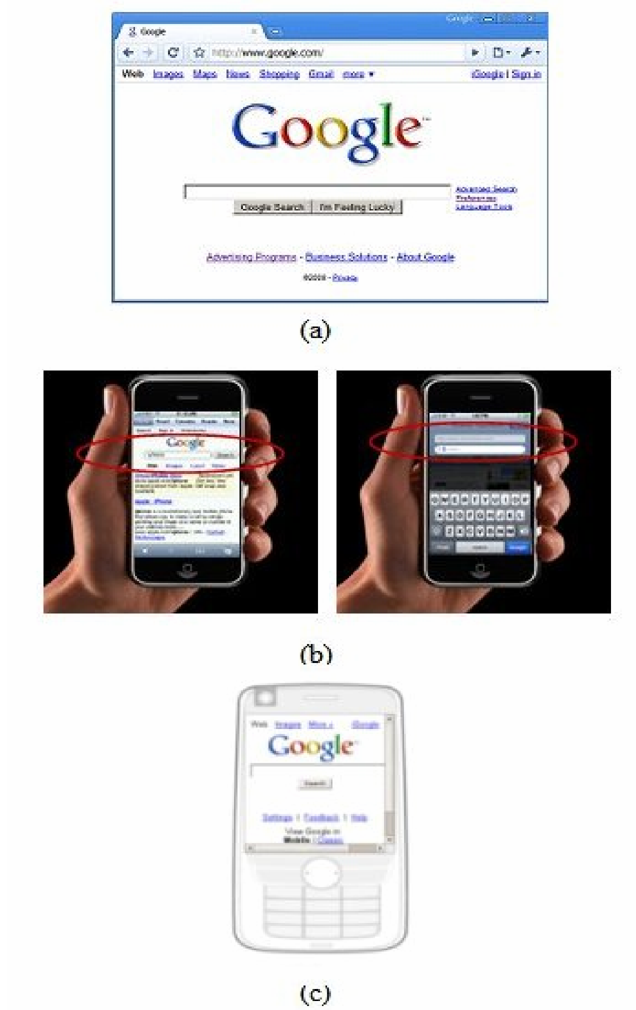

Figure 1. Examples of Google's search interfaces. (a) computer (b) iPhone (c) mobile phone

4. DATA SET

In this paper, we will present analyses of search patterns on three different Google search interfaces (which are described in the previous section). For each of these interfaces, we extracted over 100,000 queries issued by over 10,000 users during a 35-day period during the summer of 2008. The approximately 10,000 users from each platform were sampled from the Google logs by selecting a random subset of browser cookies which fell into a given numeric range. We believe an analysis conducted on this number of users is sufficiently large to draw conclusions about the differences in behavior across user populations.

We restrict the queries we analyze to all of the *Web*[2] queries made by the randomly selected users over the 35-day period. Only English queries were considered in this study. The total number of queries and users which comprise the data set for each search interface is shown in *Table 1*.

[2] Searchers are presented with the option of searching different information repositories. Besides "Web" search, other information repositories include "Images", "Maps", "News", and "Shopping".

Table 1: Dataset size.

	Computer	iPhone	Mobile
number of queries	499,999	150,000	169,448
number of users	14,209	10,184	17,201

It is important to note that we have no way of correlating a user over different devices; we can not tell if a user who issued a query on an iPhone later issued a query from their computer. All of our data is strictly anonymous; we maintain no data to match a user with an identity. For each query issued, we record a user id (generated from the request's cookie), along with the timestamp at which Google servers received the query. No other data regarding the user or the query is maintained.

4.1 Query Distributions

In this section, we provide query statistics, including query length, query classification, and query distributions, for our three interfaces described above. We show that our sample of mobile queries exhibits similar characteristics to recently published large scale studies of mobile search behavior [13, 14, 20, 23]. We update computer-based search statistics, which is valuable as the last known large-scale analysis of computer based search patterns was over seven years ago, on a set of Alta Vista queries from 2001 [8]. To our knowledge, no prior large scale analysis of search on an iPhone has been performed, but we do draw comparisons to analyses of PDA-based search in 2005 [13].

4.1.1 Query Length

As shown in Table 2, average query length is longest for computer-based search, followed by iPhone and mobile phone search. The average number of words and characters per query are approximately the same for computer-based and iPhone search, but is significantly smaller for mobile phone search.

Table 2: Average Query Length

	Computer	iPhone	Mobile phones
number of words	2.93	2.93	2.44
number of characters	18.72	18.25	15.89

For computer-based search, the average number of words per query is 2.93 (median = 3.0, standard deviation = 2.17) and the average number of characters per query is 18.72 (median = 16, standard deviation = 12.89). This is an increase from the last reported study of computer-based web search [20] where an average length of 2.6 words per query was reported. We found the average number of words per iPhone query to be the same as in computer based queries, with slightly fewer characters per query. On average an iPhone query consists of 2.93 words (median = 3, standard deviation = 1.77), and 18.25 characters (median = 16, standard deviation = 10.48). This data indicates a slight increase in query length from a study of PDA search queries in 2005 [13], which reported an average of 2.7 words and 17.5 characters per query. The length of conventional mobile phone queries is the shortest of all the mediums, with an average query consisting of 2.44 words (median = 2, standard deviation = 1.76) and 15.89 characters (median = 14, standard deviation = 9.34). That is a slight increase from the average of 2.35 words per query reported in the most recent large-scale analysis of mobile queries [23].

We attribute the difference in query length to two factors: the disparity in ease of text-entry on each platform, and the disparity in the types of queries made on each platform. Perhaps users are modifying their search behavior due to the more difficult text entry conditions. We see a significant decrease (p-value < 0.0001) (in terms of average and median) in query length for queries issued on mobile phones, which have keypads that are sub-optimal for text entry [15][16]. However, there is little difference in the length of queries issued from iPhones and Computers, both of which have QWERTY keyboards. We examine the disparity in the types of queries made on each platform in the next section, where we again see that iPhone and Computer behavior closely mimics each other, and mobile behavior is different from the two.

4.1.2 Query Classification

In order to determine the types of queries issued from the three different interfaces, we classified queries into 30 different categories using the same categorization tool described by Kamvar and Baluja [13].

The distribution of categories for each device is shown in Figure 2; iPhone categories closely mimic Computer based categories, however there are significant differences between the category distribution of queries issued from mobile phones and the category distribution of queries issued on iPhones and computers.

Because the differences in category distributions on iPhones and desktop computers are small (on average there is a .70% difference across each category), we assert that the content for which people query on iPhones and computers is generally the same. The biggest difference between these two platforms is in the "Computers & Electronics" category, where there are 2.3% more queries issued in this category from computers than from iPhones. Some example queries which were classified under "Computers & Electronics" include queries "*apple*" and *"best buy*". The second biggest difference between iPhone and computer query distributions is in the "Telecommunications" category, where there are 2.2% more queries issued in this category from iPhones than from Computers. Example "Telecommunications" queries include queries *"comcast" and "iphone"*. There are no other differences between iPhone and computer category distributions which exceed 2.0%.

Perhaps the most surprising finding is that there is only 1.7% more local queries issued from an iPhone than from a computer. This finding refutes the prevailing hypothesis that mobile web search begets a significant percentage of local searches. One explanation for the low percentage of local search from iPhones may be that users use the Maps application, rather than web search to ascertain local information. To confirm this hypothesis, we conducted a small study of user search behavior on Google Map properties on different devices[3]. Users of iPhone Maps used the application 1.6 times more days per week than maps.google.com users, and issued approximately 1.3 times more queries per day than maps.google.com users. We believe that mobile users will continue to search for a higher proportion of

[3] In this study, we compared users who issued English queries on maps.google.com to users who issued English queries on the iPhones' Google Mobile Maps Application during a one-week period during the summer of 2008.

local content than computer users, but may look for this information within an application that can provide a richer experience than what a browser can provide.

The category distribution for queries issued from mobile phones is significantly different from that of iPhones and computers. On average, there is a 1.9% difference across each category. The most drastic shift is an 11.6% increase in Adult content in mobile web search. This finding and hypotheses surrounding it have been previously discussed by Kamvar and Baluja [13][14]. Other notable differences between mobile and computer category distributions are in the "Telecommunications" category (there is a 5.7% increase in these queries on mobile phones as compared to computers), the "Computers & Electronics" category (a 4.8% decrease), the "Online Communities" category (a 4.15% increase) and the "Internet" category (a 4.07% increase). Between mobile phones and computers, there are a total of 6 categories where the difference in percentage of queries issued in a category is greater that 2.0%. This is double the number of categories where the difference in percentage of queries issued is greater than 2.0% between iPhones and computers.

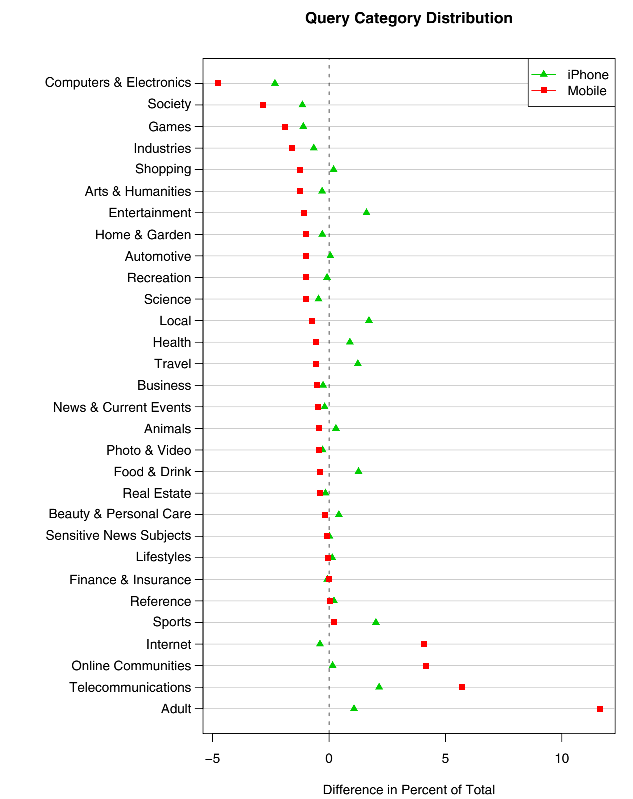

Figure 2: The difference in the distribution of query categories between PCs and iPhones (shown in green) and between PCs and conventional mobile phones (shown in red).

4.1.3 *Query Diversity*

In addition to looking at the distribution of the query categories, we also measure the distribution of the individual queries. Query distribution is another measure of the diversity of each query set. We used the query distribution to compute the diversity of the query set in two ways. First, we simply computed the number of unique queries in the query set. Secondly, in order to get a deeper understanding of the repetition pattern in the query set, we examined what percentage of the total query volume is accounted for by the top 1000 unique queries. The "long tail"[4] of web search is an often-referenced phenomenon [9, 10, 11, 21, 22], and this metric allows us to compare the "tail" of web search on each medium.

Again, following the results from the prior sections, we found that for both cases, iPhone query diversity resembles Computer query diversity more closely than mobile query diversity. The number of unique queries accounts for 69.6% of computer based queries, 61.6% of iPhone queries and 45.4% of mobile phone queries. A higher percentage of unique queries indicates that the query set is more diverse.

The second way we visualized query diversity was by examining what percentage of the total query volume is accounted for by the top N unique queries (independent of case and spacing). We took a random sampling of 50,000 queries from computer, iPhone and mobile searches. Figure 3 shows the percentage of query volume that the most popular 1 to 1000 unique queries account for. A higher percentage indicates lower diversity in the query set; as this percentage increases the "head" of the query frequency graph is getting bigger while the "tail" of the graph is getting smaller. We find that there is an increasing percentage of query volume accounted for by the top 1000 queries in computer, iPhone, and mobile searches, respectively. In other words, the "tail" is shorter for mobile web search than for iPhone search. The "tail" of computer-based web search is the longest of all mediums.

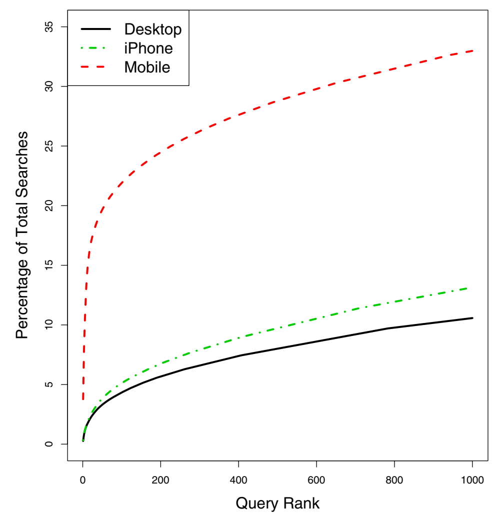

Figure 3: Cumulative percentage of total searches accounted for by the top 1000 queries.

Our finding mimics the Kamvar and Baluja's 2006 findings [13] which discuss various hypotheses for the decreasing diversity in queries across devices, such as user demographic, browser capabilities and ease of text entry. However, what is interesting to note is that the gap in diversity between high-end phones and computers seems to be shrinking, as compared to the finding in 2005. Table 3 enumerates the values at the beginning and end of the cumulative percentage graph. The difference between the

[4] *Long tail queries* are rare queries which are issued only by few users, but the aggregate of these queries account for a large percentage of unique queries.

cumulative percentage of the top 1000 computer-based and iPhone queries is a mere 2.5%, which much smaller than the approximately 7% gap in diversity reported between computers and PDAs in 2005 [13].

Table 3: Percentage of total searches accounted for by top queries

	Computer	iPhone	Mobile
percentage of 50,000 query sample accounted for by the top query	0.30	0.31	3.75
percentage of 50,000 query sample accounted for by the top 1000 queries	10.53	13.05	32.78

In our first order analysis of web search across computers, iPhones, and mobile phones, we have consistently found that search patterns on an iPhone closely mimic search patterns on computers, but that mobile search behavior is distinctly different. We hypothesize that this is due to the easier text entry and more advanced browser capabilities on an iPhone than on mobile phones. Thus we predict that as mobile devices become more advanced, users will treat mobile search as an extension of computer-based search, rather than approaching mobile search as a tool for a distinct subset of information needs.

Our goals in reporting these first-order statistics are to provide a direct comparison of search patterns across three mediums, and to show that our dataset, while small, is a representative sample. The mobile queries in our dataset resemble recently published statistics for mobile search. The iPhone query statistics that we report confirm the finding that search from high-end phones are continuing along the trajectory of meeting a more diverse set of information needs than is met for search from conventional mobile phones.

In the next section we will shift our attention to a comparison of the patterns that search *users* exhibit on each of these interfaces rather than aggregate analysis of search *queries* issued from each interface, to see if the same trends hold.

5. USER-BASED QUERY PATTERNS

In this section, we will focus on user-based search behavior across each medium. We will first look into what comprises an average search session (a search session is defined as series of queries made by a single user in succession). Next, we analyze those users who had multiple search sessions over the 35-day period, and explore to what extent past sessions are determinate of future sessions. Finally, we analyze those users who did *not* return more than once in the 35-day period to see if there were any commonalities in their search experience that would indicate areas for improvement.

5.1 A Users' Search Session

We start with analyzing a single search session. We take the definition of a search session from [18] as "a series of queries by a single user made within a small range of time". We will refer to this range of time as the session delta. Following [13, 18, 20], we will user a session delta of 5 minutes - if no interaction happens within 5 minutes of the previous interaction, a user's session is deemed closed. The next interaction is considered a separate session. Table 4 shows the number of search sessions which comprised our dataset for each platform.

Table 4: Search Session Statistics

	Computer	iPhone	Mobile
number of search sessions	257,163	82,043	99,649
average number of queries per search sessions	1.94	1.82	1.70

On average, there were more queries per session on computers than on iPhones and conventional mobile phones. There are an average of 1.94 queries per session on a computer (median=1, standard deviation=2.07), 1.82 queries per session on an iPhone (median=1, standard deviation=1.67), and 1.70 queries per session from conventional mobile phones (median=1, standard deviation=1.91). This trend may be indicative of the following:

- The depth of users' information needs. Perhaps users on mobile devices are more likely to query topics which have a "quick answer" available. (For example, weather and stock queries are often known to trigger a pre-formatted result, and users may not need to explore any further for their desired information). Taking this hypothesis a step further, we suggest that perhaps users are simply unwilling to explore topics in depth as the barriers to exploration (text entry, network latency) increases. This idea has been explored by Jones et al [12] in what is called a "laid-back" approach to search on mobile devices. This approach implies that users enter queries because they are of interest to their current situation, but have no urgency in iterating or deeply exploring their query in real time. This would suggest that on mobile devices, topical information relating to the query should be summarized on the page which contains the list of search results. Increasing the size of the snippet, or aggregating common information across the web pages in the search results may benefit users on mobile devices. Furthermore, an integrated search experience across a user's computer and mobile device may be beneficial to users in that it would allow for an easier follow-up search when the user has more time, and better computing resources (e.g. network bandwidth, keyboard size).
- While previous research suggests that computer-based searchers prefer to refine their queries in place of browsing through results [8], the converse may be true for mobile search. In particular, as text entry gets more difficult, users may be willing to spend more time browsing the list of search results rather than refining their query in order to find the desired information. A future study is planned to look at the click position of results on each three mediums to verify this hypothesis. Following this hypothesis, we would expect clicks in more positions from iPhones and conventional mobile phones, than from Computers (where we would expect the result in the first position to dominate the clicked results).

We believe each search session represents a single "information need", or "task". To understand how the nature of "information needs" differs on each device, we analyze the distribution in *task* categories. This is similar to our analysis of query categories, but on an aggregate level. We represent the category of each "information need" by categorizing one query in the search session (using the same tool described in section 4.1.2). We feel that this is a representative description of a search session, since a user is likely to query within the same topic for the duration of their search session. The average number of categories per session is 1.3 for computers (median = 1, standard deviation =0.76), 1.2 for iPhones (median = 1, standard deviation = 0.54) and 1.2 for conventional mobile phones (median = 1 and standard deviation = 0.49).

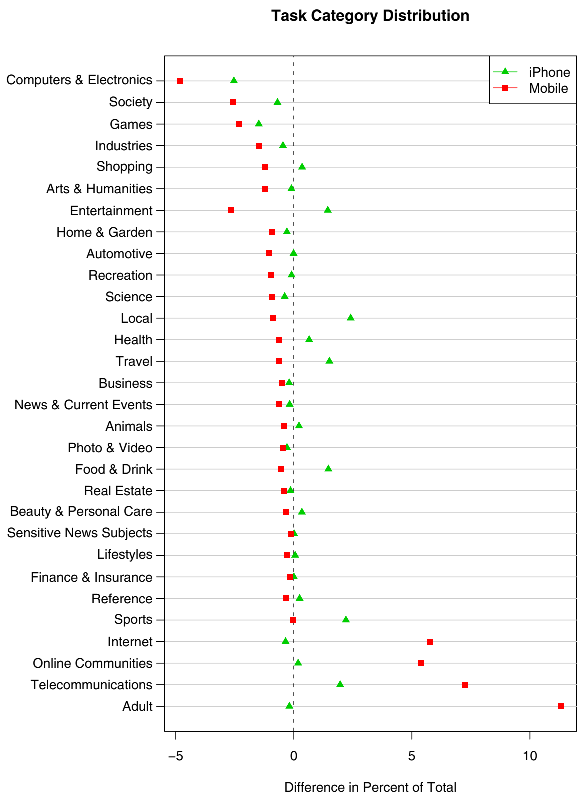

Figure 4: The difference in the distribution of task categories between PCs and iPhones (shown in green) and between PCs and conventional mobile phones (shown in red).

In Figure 4 we show the distribution of *task* categories. Overall, the task distribution indicates a further convergence between iPhone and Computer based search patterns. There is on average a 0.68% difference in iPhone and Computer *task category* distribution, as opposed to an 0.70% difference in the iPhone and Computer *query category* distribution.

When we measure search categories by task rather than by query, there are a few notable changes. For example, although the percentage of Adult queries on iPhones exceeds the percentage of Adult queries on computers, the percentage of adult-oriented tasks on iPhones is slightly lower than on Computers. This indicates that looking for Adult content is a less prominent task on iPhones, but that the users who do choose to query adult content have longer search sessions[5] on an iPhone than they do on computers. The same trend follows with Entertainment queries; those users who query for entertainment queries on an iPhone have longer search sessions than those who query entertainment queries on a computer. However, the trend of longer search session on an iPhone is not prevalent. There are few categories which inspire longer sessions on an iPhone; as Table 4 shows, on average the length of iPhone search sessions is shorter than Computer based queries.

The most prominent example of shorter search sessions on iPhones occurs in the "Local" category. iPhone users have 1.7% more local queries than computer users, but 2.4% more local tasks than computer users. This implies that Local "information needs" on iPhones are comprised of fewer queries than on computers. This may be due to the fact that users are redirected to the iPhone Maps application if they click on a business listing. Any queries subsequent to a click on a business listing are not counted in the web session.

In the next section, we analyze the task patterns of individual users to understand to what extent past search behavior predicts future search behavior.

5.2 Frequent Users

In this section, we analyze those users who return to Google for more than one information need, during the 35-day period. As shown in Table 5, computers are used for more than twice as many tasks as iPhones and mobile phones. The number of tasks per user is the biggest difference between iPhone and computer based search behavior. By this metric, the iPhone is more like a mobile phone than a computer, in that it seems to be a secondary mode of searching.

Table 5: Repeat User Statistics

	Computer	iPhone	Mobile
average number search sessions per user during the 35-day time frame.	18.10	8.06	5.79

We wanted to determine if there was a pattern which emerged for frequent searchers. We define our frequent searches as those individuals who returned to Google for at least 10 tasks during the 35-day period analyzed. There were 4227 computer based users, 2224 iPhone users, and 1839 mobile users, who had engaged in 10 or more tasks. To normalize our analysis, we randomly sampled 10 tasks from each of these users and categorized each task using the same method described in section 2. For the rest of this section, we will restrict our analysis to those users, and their randomly sampled 10 tasks.

Maintaining a parallel to our computation of query diversity, we measured the diversity of users' information needs in two ways. The first measure of diversity was to see what percent of users used their device for a single type of information need. Only 0.5% of computer users had all 10 tasks concentrated in a single category, but iPhone users were even less likely to have

[5] By "longer" search sessions we are referring to the number of queries issued in the session, not the absolute time duration of the session.

concentrated interests; 0.13% of iPhone users had all 10 tasks fall into one category. On the other extreme, 9.8% percent of mobile users had all 10 tasks concentrated in one topic. Furthermore, the percentage of users' whose tasks were exactly the same (eg they issued the same query) was 0.2%, 0.0% and 6.6% for computers, iPhones and conventional mobile phones, respectively. This suggests that iPhone searchers have a slightly *more diverse* set of information needs than computer-based users, and both are a lot more diverse than the mobile users.

One explanation for the diversity of information needs exhibited by iPhone searchers may be the contextual nature of mobile search. Based on a diary study of mobile information needs, Sohn, Li, Griswold and Hollan [19] reported that 75% of mobile information needs were prompted by contextual factors, which consist of either *activity* at the time, current *location* and related artifacts, *time* when the need arose, or con*versations* occurring with others. These factors may result in a varied set of information needs. However, previous studies [5, 19] suggest that many information needs that arise while users are mobile are not answered through mobile search. We hypothesize that the iPhone, with its relatively easier mode of text entry, faster connection speeds, and increased screen size, may encourage users to answer a higher proportion of their mobile information needs via their iPhone, thereby resulting in a more diverse set of information needs than either desktop or mobile search.

Next, we measured the distribution of information needs for those users. Figure 5 illustrates a heat map of task category. The columns represent each user, and each row represents a category. Each cell is a number between 0 and 10, which is the number of tasks a user issued in that category. The darker cells represent numbers closer to 10, and the lighter cells indicate cells closer to 0. These heat maps not only confirm the finding that iPhone users have the most diverse interests, and mobile users have the least diverse interests, but it gives us insight into the search patterns of individual users. We estimate that 45% of frequent mobile users can be classified in a behavioral "bucket", meaning that they are likely to query within a single category from their device. Interestingly, the categories which are most popular (such as Online Communities) are those which have content "optimized" for mobile devices. This finding furthers the hypothesis that mobile users approach search with a specific set of information needs.

In the next section, we propose a metric for quantifying the variability of information needs for any frequent user.

5.2.1.1 *Entro-percent: a metric for user variability*

In this section, we propose a metric to quantify the variability of information needs for a specific user. The goal is to compare quantitatively whether mobile users tend to have a narrower set of information needs than computer users or iPhone users. Different from the earlier analysis that are restricted to users who have at least 10 tasks and the 10 sampled tasks per user, the metric proposed here applies to any user with more than one task. This allows us to extend our findings on the restricted set of users and tasks to general frequent users.

We first motivate our metric with a few desired properties. We denote $\{C_1,\ldots,C_K\}$ to be the K possible categories that a task can be classified into.

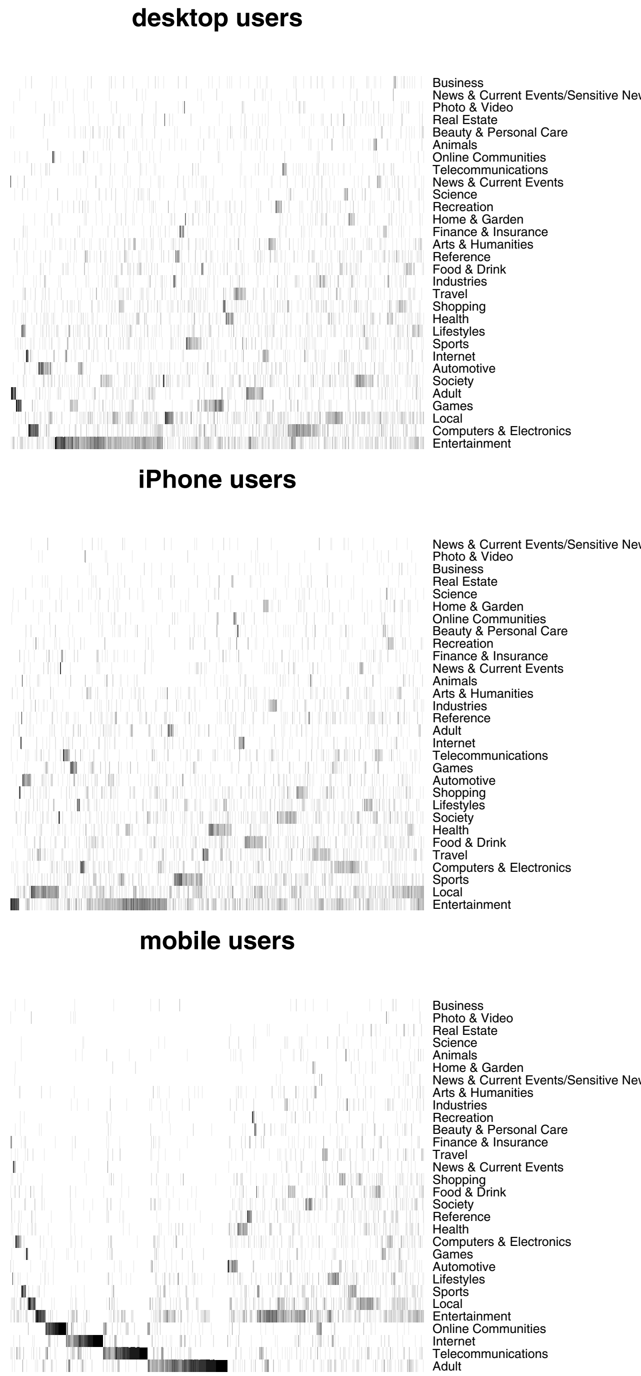

Figure 5: Heat maps depicting the variability of task category per user

We also denote p_k as the corresponding percentage of the user's tasks that are in category C_k. Each user is then associated with a percentage vector $\boldsymbol{p}$, the abbreviation for $(p_1,\ldots,p_k)$. To quantify the variability of information needs for a user, the metric $M(\boldsymbol{p})$ should have the following desired properties:

- $M(\boldsymbol{p})$ should return values between 0 and 1. A higher value indicates higher variability.
- The percentage vector that has the maximum metric score should be $(1/K,\ 1/K,\ \ldots,\ 1/K)$. This is because a user should be considered most variable if she has the same

number of tasks in each category. In other words, we do not have any information about which categories this user prefers.

- The percentage vector that has the minimum metric score consists of one and only one "1" and the rest are all zeros. This is because a user is the least variable if all his/her tasks are in one category. We have the most information that this user prefers that category.

The category of a task by a user is a categorical random variable that takes values in $\{C_1,\ldots,C_K\}$. The variability of such random variable can be quantified through its information entropy [17]. Entro-percent is a normalized entropy metric such that it has the properties described above and is defined as:

$$M(p) = \frac{-\sum_{k \in K} p_k \log(p_k)}{\min(\log(n), \log(K))}$$

where n is the number of search tasks from the user.

5.2.1.2 Comparing entro-percent across search interfaces

We now compare the average entro-percent scores on the three platforms. Because entro-percent is defined for any frequent users, we are no longer limited to 10 tasks. In fact, we can bucket the users by their number of tasks and make comparison within each bucket.

Figure 6 plots the average entro-percent scores across the number of tasks. The most distinctive trend in the graph, is that the curve for mobile phone users is completely below the desktop computer curve and the iPhone curve, indicating that on average, the mobile users are indeed 15-50% less variable. On the other hand, the computer and iPhone users are very similar, as their curves trace each other very closely, particularly for smaller number of tasks. iPhone users have slightly higher scores when the number of tasks is greater than 10, which confirms our findings earlier. Notice that the comparison shown in Figure 6 is controlled for the total number of tasks the user performed, and hence the conclusion we have drawn here is more general.

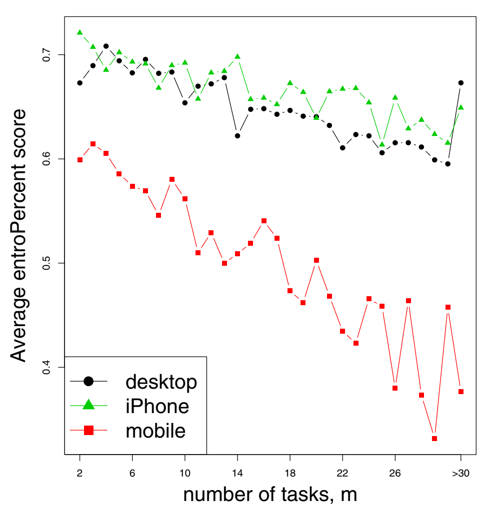

Figure 6: Variability of search sessions for users who issued *m* tasks during the 35 day period

Finally, as a concrete example, in Figure 7 we present two examples of sessions that resulted in entro-percent scores of .2 and .3 respectively. Two random users, who each had engaged in 10 tasks over the 35-day period, were selected. Even though both users query sessions consisted of queries in only two categories, the first user has a lower entro-percent score because she is more focused in one category.

Query	**Category**
free tones	Telecommunications
message in a bottle free ring tone	Telecommunications
free music downloads	Entertainment
free tones	Telecommunications
free beastie boys ringtones	Telecommunications
free beastie boys ringtones	Telecommunications
free beastie boys	Entertainment
free tones	Telecommunications
free tones	Telecommunications
free tones	Telecommunications

(a)

Query	**Category**
mike murphy baseball	Sports
sf giants	Sports
mike murphy baseball	Sports
lost in space wavs	Entertainment
lost in space downloads	Entertainment
dr smith wavs	Entertainment
lost in space	Entertainment
lost in space wavs	Entertainment
mike murphy baseball	Sports
mike murphy baseball	Sports

(b)

Figure 7: Two examples of user search sessions over the 35 day period (the session's representative query shown on the left, with the associated task category on the right). User (a)'s search sessions results in an entro-percent score of .2 and user (b)'s search sessions result in an entro-percent score of .3

5.3 Infrequent Users

We have discussed search patterns for frequent users (those who return for at least 10 information needs), however the analysis of *in*frequent users is also valuable to understand where search on each medium may be failing. In this section, we look at the search patterns for users who engaged in only one search session over the 35-day period we analyzed.

As shown in Table 6, there are far more of these single-session users on mobile phones than any other medium. This may indicate that mobile search is not as "ubiquitous" as the devices themselves are. However, we see that there are fewer iPhone users who only search once over the 35-day period than computer users who search only once. This trend indicates that phones *can be* a more ubiquitous entry point for search.

However, the reason for why users don't return is elusive. No consistent trends were found in the analysis of queries per session, query length, or task distribution for infrequent users. A follow-up diary study is planned across all three platforms to investigate this question with data that is *not* captured in the logs, such as physical circumstances surrounding a query, the users' perceived experience of their search session, and exact latency metrics for search sessions for those users who return and those who don't. The only clear trend apparent from studying infrequent users is that there are more of them on conventional mobile devices than any other search medium.

Table 6: Single-Session User Statistics

	Computer	iPhone	Mobile
percent of users who engaged in one search session over the 35-day period	29.4[6]	22.89	42.6
average number of queries per search sessions	1.88	1.89	1.74
average characters per query	18.00	16.04	15.86
average words per query	2.795	2.589	2.489

6. DISCUSSION AND FUTURE WORK

We have presented the results of a two-part analysis of Google search patterns across three separate device types: computer, iPhone, and conventional mobile phones. We first conducted a first-order analysis of the query stream, encompassing query length, topical query classification, and query diversity. The second piece of analysis focused on user-based query patterns, that is, the diversity of information needs on a per-user basis and the patterns of frequent and infrequent searchers. We also presented a new metric, *entro-percent*, for quantifying the variability of a user's search intentions across time.

We have consistently found that search patterns on an iPhone closely mimic search patterns on computers, but that mobile search behavior is distinctly different.

Our findings can be summarized as follows:

- Query length is very similar between computer and iPhone search, but is significantly shorter for mobile phone search. We hypothesize this may be due to ease of text entry on each type of device.
- The distribution of query categories was similar between iPhone and computer searches. The category distribution from mobile search is decidedly less diverse than those from both iPhone and computer search; both computer and iPhone queries had a much higher percentage of unique queries.
- Surprisingly, there was no significant difference in the percentage of local search on the iPhone and on the computer. However, this does not imply that local content is less important to high-end phone users than conventional mobile phone users. We find that users search for local content within an application that can provide a richer experience (such as the iPhone maps application) if it is available. In the absence of a dedicated maps application on mobile phones, we see an increase in queries for local information, relative to computer-based search.
- We observed that the proportion of adult content from iPhone searches was similar to that from computer-based searches and had significantly decreased from the proportion of adult queries on conventional mobile phones. This decrease in adult content on high-end devices is in line with the hypotheses discussed in a 2006 study of mobile search behavior [14][15].
- The diversity of information needs *per user* was greatest for iPhone searchers. Conventional mobile phone users had the least diverse information needs, such that we estimate that 45% of these users could potentially be classified into single topic area of interest.
- On a per search session basis, computer users had the greatest number of queries per session, followed by iPhone, and then conventional mobile phones. We hypothesize that this may be indicative of the nature of the information needs exhibited by users on different devices (e.g., users may be more likely to search for quick factual information on mobile phones). For conventional mobile phone users, the difficulty of text entry may also discourage them from issuing more queries. Users on mobile phones may be more likely to browse multiple results in place of issuing query refinement.
- The biggest difference discovered between computer and iPhone users was that frequent computer-based searchers had a much higher rate of return than frequent iPhone or mobile phone searchers. We hypothesize that search on any mobile device is still considered to be a secondary mode of searching in the US.

Based on our findings, we offer the following suggestions for improving the search experience across mobile devices:

- For conventional mobile phones, we suggest that search engines use the relatively low diversity of queries to improve the service. For example, since the "tail" of search is much smaller on these devices, possibilities such as prefetching likely queries and search results would yield a much higher target rate than on other devices. This target rate can be further improved by considering the narrow scope of an individual user's interests based on past queries made from this device. Search interfaces that are targeted to the users' primary interest (such as a sports-themed front page) may also improve the user experience and increase user return rates.
- For high end phones, we suggest search be a highly integrated experience with computer-based search interfaces. The consistent similarity in search behavior between computer and iPhone based search suggests that users may begin to treat mobile phone search as an extension of computer-based search. For example, content that was searched for on a computer should be easily accessible through mobile search (through bookmarks, search summaries), and vice versa. Most importantly, this similarity in queries indicates that we can use the vast wealth of knowledge amassed about conventional computer based search patterns, and apply

[6] In their analysis of Yahoo! query logs, Wedig and Madani [22] reported that 25% of cookies had only one search within a six month period.

it to the emerging high-end phone search market, to quickly gain improvements in search quality and user experience.

As in all logs-based studies, one limitation of our research is that it is difficult to infer the context surrounding users' search activity. For example, we can not discern from the logs what drives users to return, or to abandon search from a particular medium. A follow-up diary study is planned across all three platforms to investigate several of the questions raised by this study, such as the reasons behind users' return rate (we hypothesize this may be impacted by factors such as network latency during their first search session and ease of accessing the browser on the device), factors that dictate which medium is used to answer an information need, and success and satisfaction across the different mediums.

7. ACKNOWLEDGMENTS
We would like to thank Shumeet Baluja for his comments and feedback on our research.

8. REFERENCES
[1] Baeza-Yates, R., Dupret, G., & Velasco, J. 2007. A study of mobile search queries in Japan. Query Log Analysis: Social and Technological Challenges. WWW 2007 Workshop.

[2] Beitzel, S., Jense, E.C., Chowdhury, A., Grossman, D., & Frieder, O. 2004. Hourly analysis of a very large topically categorized web query log. SIGIR 2004. 321-328.

[3] Church, K., Smyth, B., Cotter, P., & Bradley, K. 2007. Mobile information access: A study of emerging search behavior on the mobile Internet. ACM Transactions on the Web, 1(1).

[4] Church, K., Smyth, B., Bradley, K., & Cotter, P. 2008. A large scale study of European mobile search behaviour. Mobile HCI 2008. 13-22.

[5] Dearman, D., Kellar, M., & Truong, K. 2008. An Examination of Daily Information Needs and Sharing Opportunities. CSCW 2008. 679-688.

[6] Fox, S. 2008. Almost half of all internet users now use search on a typical day. http://www.pewinternet.org/pdfs/PIP_Search_ Aug08.pdf. 2008. Pew/Internet.

[7] Covey, N. 2008. *Critical Mass: The Worldwide State of the Mobile Web*. Nielsen Mobile. http://www.nielsenmobile.com/documents/CriticalMass.pdf.

[8] Jansen, B.J., & Spink, A. 2005. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Information Processing and Management, 52(2006): 249 – 263.

[9] Jansen, B. J., Spink, A., Bateman, J., & Saracevic, T. 1998. Real life information retrieval: A study of user queries on the web. SIGIR Forum, 32(1): 5-17.

[10] Jansen, B. J., Spink, A., & Saracevic, T. 1998. Real life information retrieval: A Study of User Queries on the Web. SIGIR Forum, 32(1): S-17

[11] Jansen, B., Spink, A., & Saracevic, T. 2000. Real life, real users, and real needs: A study and analysis of user queries on the web. Information Processing and Management, 36(2): 207–227.

[12] Jones, M., Jain, P., Buchanan, G., & Marsden, G. 2003. Using a mobile device to vary the pace of search. Mobile HCI 2003. 90-394.

[13] Kamvar, M., Baluja, S. 2006. A Large Scale Study of Wireless Search Behavior: Google Mobile Search. CHI 2006. 701 – 709.

[14] Kamvar, M. & Baluja, S. 2007. Deciphering Trends in Mobile Search. Computer, 40(8): 58-62.

[15] Kamvar, M. & Baluja, S. 2007. The Role of Context in Query Input: Using contextual signals to complete queries on mobile devices. Mobile HCI 2007. 121-128.

[16] MacKenzie, I.S. 2002. KSPC (keystrokes per character) as a characteristic of text entry techniques. Mobile HCI 2002. 195-210.

[17] Shannon, C.E. 1948. A Mathematical Theory of Communication. Bell System Technical Journal, 27: 379-423.

[18] Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. 1999. Analysis of a Very Large Web Search Engine Query Log. SIGIR Forum, 33(1): 6 -12.

[19] Sohn, T., Li, K.A.,Griswold, W. & Hollan, J.D. 2008. A diary study of mobile information needs. CHI 2008. 433-442.

[20] Spink, A., Jansen, B., Wolfram, D., & Saracevic, T. 2002. From E-Sex to E-Commerce: Web search changes. IEEE Computer, 35(3): 107-109.

[21] Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. 2001. Searching the web: The public and their queries. Journal of the American Society for Information Science and Technology, 52(3), 226-234.

[22] Wedig, S. & Madani, O. 2006. A large-scale analysis of query logs for assessing personalization opportunities. SIGKDD 2006. 742 – 747.

[23] Yi, J., Maghoul, F., & Pendersen, J. 2008. Deciphering Mobile Search Patterns: A Study of Yahoo! Mobile Search Queries. WWW 2008. 257-266.