Message

From: Web Ranking Team [Redacted@google.com]
Sent: 8/16/2014 1:13:23 AM
To: Redacted@google.com
Subject: [Web Ranking Team] Aug 11 -- Aug 15, 2014 was updated

Pandu Nayak updated the page Ranking Newsletters » 2014 Q3 Ranking Newsletters » Aug 11 -- Aug 15, 2014. View the changes below.

Color key: Insertion | ~~Deletion~~

Contents

This week's newsletter starts with a great article by Jangwon Seo on an in-depth analysis of the differences of user intent and behavior between mobile and desktop. As you know, we are in the process of understanding what mobile means for Ranking and Eval, and this study provides valuable data on what users are actually doing with search on mobile. Check it out and think about what this says about what we should be working on!

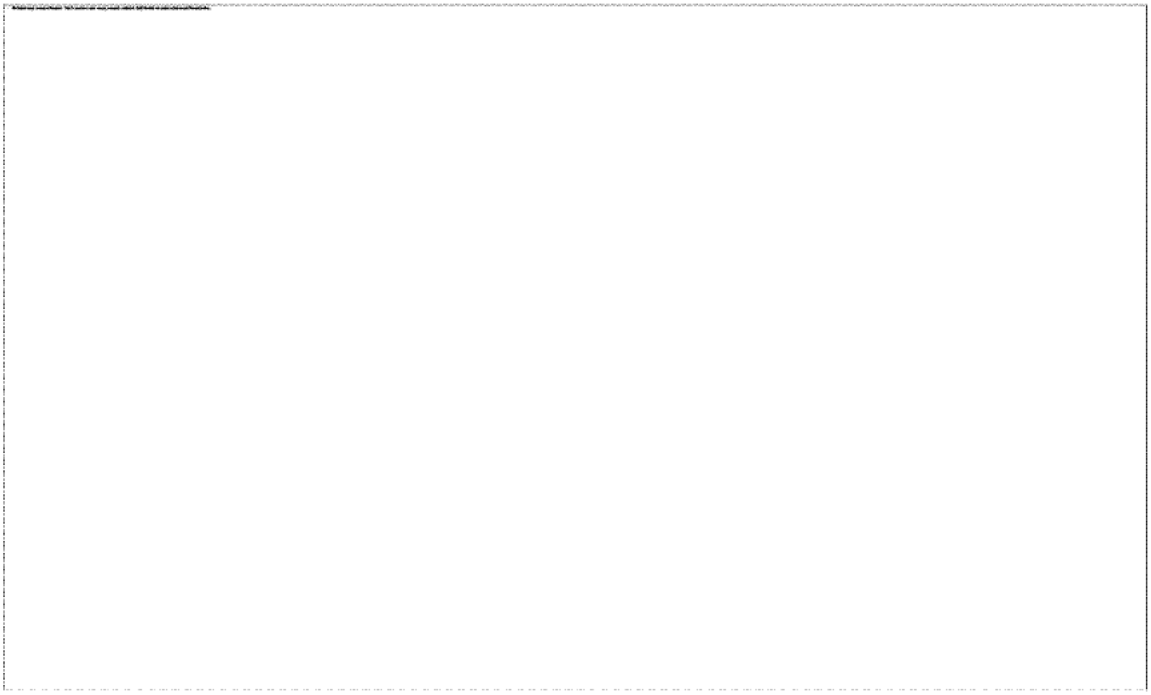
Mobile vs Desktop: User intent and behavior

In search products, mobile has often been considered a secondary interface for a long time, since historically most web traffic has come from desktop, and it was somewhat challenging to use mobile devices for search because of difficulty typing and slow network. However, as smartphones are widely used, we are seeing big volume growth of mobile search traffic. The following plot presents input query volumes of desktop and mobile in en-US over the past one year. While the desktop volume is stable, mobile volume is gradually catching up the desktop volume. Currently, mobile-to-desktop volume ratio is at least 2:3 even in a conservative sense.

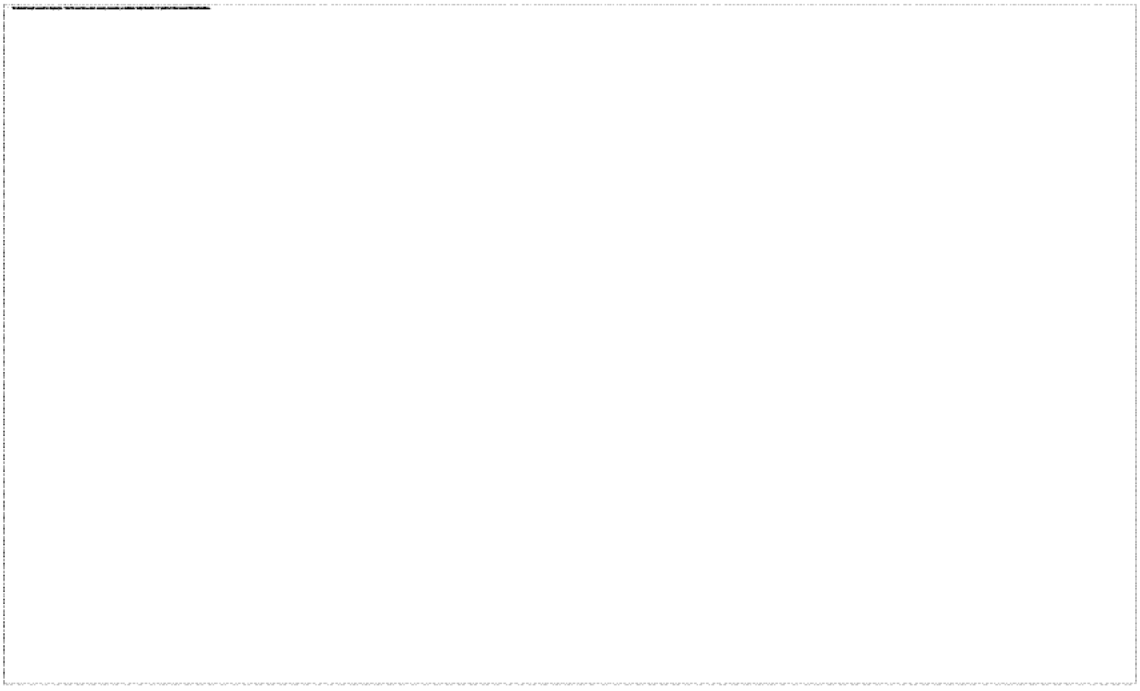
Ex. No.
UPX1087
1:20-cv-03010-APM

Redacted

GOOG-DOJ-03571718



In fact, we are even seeing mobile volume exceeding desktop volume on some days. As shown in the following graph, desktop traffic tends to be larger during week-days, while mobile traffic is larger during weekends. It shows that mobile search is already incorporated into many people's life cycle and the usage pattern is relatively independent of time.



From these trends, it is predicted that mobile search traffic becomes consistently higher than desktop volume in near future. For example, in some locales such as ja-JP, this has already

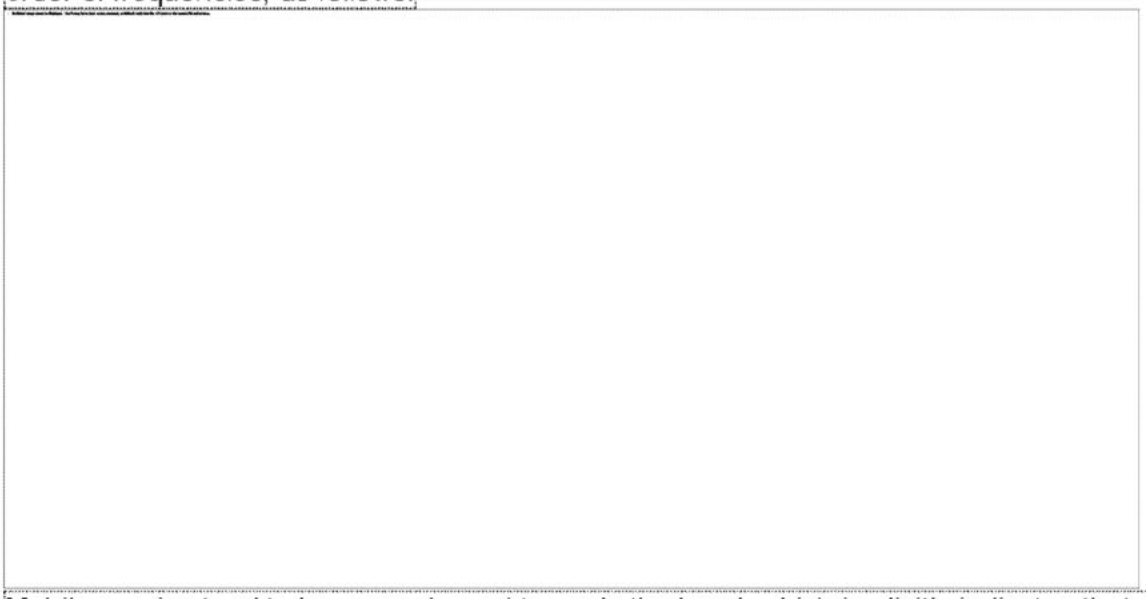
happened. Given this, a question which we need to answer, is whether mobile search is different from desktop search. If so, how are they different? To answer to this question, we conducted some analyses, focusing on users intents and behaviors.

Mobile Search vs Desktop Search via Top-level User Interaction Metrics

Our top-level metrics measured for each interface reveal high-level ideas about how mobile search is different from desktop search. We looked at several production LE metrics over 1 month (April 2014) to see how they appear differently in mobile.

CTR	Confidential	<	Confidential
Manual Refinement		<	
Queries Per Task		<	
Query lengths (in char)		=	
Query lengths (in word)		=	
Abandonment		>	
Average Click Position		>	
Duplicates		>	

In addition, query diversity is indirectly measured, by accumulating query frequencies in reverse order of frequencies, as follows:



Mobile queries tend to be more skewed towards the head, which implicitly indicates that mobile queries are less diverse than desktop queries.

A user behavior that we can discover from those top-line metrics, is that mobile users have higher resistance to issue a refine query (lower manual refinement), therefore, tend to examine more results in a search result page (higher average click position) or abandon it more frequently (higher abandonment). This implies that mobile users want to complete their task in a fewer searches.

Diverged User Intents

A convenient way to understand what users intended by queries is looking at clicked urls for the queries and mining an intent from the urls. To take advantage of the clicked urls reflecting user intents, we leveraged Mobile Navboost data. Mobile Navboost was launched in Q1, 2014, and brought us a huge LE win. Indeed, this can be thought as an indirect clue that mobile search intents are diverged from the desktop search intents. For more in-depth analysis, we first computed a mobile-fraction.

mobile-frac = mobile-trad-denominator / total-trad-denominator

We can group queries by fraction ranges, as follows:

desktop-dominant queries	mobile-frac < 0.1	Confidential
mobile-dominant queries	mobile-frac > 0.5	
mixture queries	0.1 < mobile-frac < 0.5	

where "%" is the percentage of each group slice.

In the followings, for each group, we will present how user intents differ between mobile and desktop, with manually classified examples.

Top Intents of Mobile-Dominant Queries

We first look at the mobile dominant queries. The queries were sorted by the total trad denominators, i.e., popularity. A reason for this is that interpreting tail-ish queries is often tricky. We then manually classified intents of top 50 queries. The results are presented in the following table.

Local Business	[autozone in lubbock tx] [roxy 14 stadium santa rosa ca] [walmart excelsior springs missouri] [regal valley view grande roanoke va]	Confidential
Porn & Adult		

Redacted

Event Monitoring (sports event, stock price, traffic)	[toto 4d result malaysia today] [cirtran investorshub] [pierpass cam]	Confidential

Note that the reported percentages were computed only for top 50 queries sorted by popularity. The numbers are provided to give readers a rough idea about what types of intents exist in this slice, rather than a precise intent distribution estimation. Thus, we should not generalize that overall mobile-dominant queries would have that distribution.

As seen in the table, the top intent is local business seeking intent. This sounds obvious, considering the high portability of mobile devices. The second largest slice is porn queries. This fact has been well known from several previous studies. The reason is that a mobile device is a private device by nature, despite small screen sizes. The last intent is event monitoring queries. This intent includes various real-time event tracking, including stock price checking, or traffic monitoring on a specific route.

Top Intents of Desktop-Dominant Queries

The same procedure as done for the mobile dominant queries was applied here.

(Office- or Home-)work	[verify java] [reverse complement] [usr bin ld cannot find] [css vertical align]	Confidential
Desktop application / Download	[download internet explorer for mac] [star clip art] [form 941 schedule b] [watch how i met your mother season 8 episode 6]	
Navigation to big sites	[facebookc om] [hu u] [ebay www ebay com]	

The top intent includes queries for work. More specifically, there are many queries related to professional or industrial office work, and also queries for school homework. Among them, a type of queries that dominate, is coding-related queries. The second intent is finding desktop applications, or downloading applications, movies, songs, or images. Although the application intent and download intent are different, distinguishing one from the other is not trivial. Thus, they were classified together as the same intent. The last intent is navigating to big sites, such as facebook.com or hulu.com. These queries' population in mobile tends to be smaller because these sites are often accessed by their dedicated mobile apps in mobile.

Divergence in Intent for Common Mobile/Desktop Queries

This group is probably the most interesting slice, because mobile users sometimes have

different intents even for the same queries. First of all, we need to measure how much the intents are diverged. To do this, we computed the Kullback-Leibler divergence between the mobile trad distribution and the desktop trad distribution for each query, assuming that each distribution forms a multinomial distribution.
As a result, Confidential of queries in this slice have KL-divergence values greater than 0.5, i.e., substantially diverged. Since we are interested in only queries having different intents, we sorted the queries by their KL-divergences, and manually classified top 100 selected queries. Notice that all the percentage numbers reported below are based on the top 100.

- Local Business Seeking Intent Confidential

[norton pub] Desktop: 1(0.500): http://www.wwnorton.com/ 2(0.192): https://www.wwnorton.com/ 3(0.080): https://plus.google.com/107868540749299955940/about?gl=us&hl=en Mobile: 1(0.412): http://local.yahoo.com/info-11764169-norton-s-pub-rochester 2(0.192): https://plus.google.com/112597814646379876636/about?gl=us&hl=en 3(0.133): http://www.yelp.com/biz/nortons-pub-rochester
[galaxy nyc] Desktop: 1(0.649): http://schools.nyc.gov/Offices/DBOR/Galaxy/ 2(0.163): http://www.menupages.com/restaurants/galaxy-diner/ 3(0.043): http://mygalaxy.nycenet.edu/ Mobile: 1(0.502): http://www.menupages.com/restaurants/galaxy-diner/ 2(0.161): http://www.galaxyarmynavy.com/army_navy_stores_nyc....asp 3(0.112): http://www.yelp.com/biz/galaxy-global-eatery-new-york

Similar to the mobile-dominant queries, local business seeking queries are prominent in this slice also. In the first example [norton pub], desktop users meant "Norton Publishing" company, whereas mobile users wanted to find a pub named "norton". For [galaxy nyc], desktop users sought a site for the Department of Education in New York City, while mobile users wanted to find Galaxy Diner in New York City, or other similar named local stores.
A finding from the first example, is that it may be crucial that mobile ranking signals model fine-granularity locales properly, because local queries are abundant and these local queries are very sensitive to specific local areas. For example, in Mobile Navboost,

the top and third clicked results for [norton pub] are about a pub in Rochester, NY. For people in Rochester, the recommend results look good, but people not in the Rochester, NY area are not likely to be satisfied with this. People outside Rochester would want other "Norton Pub" in their regions (if any), or prefer the "Norton Publishing" site even when searching on a mobile phone.

- Dictionary Lookup Intent Confidential

[black] Desktop: 1(0.223): http://cdn.howtogeek.com/wp-content/.../Plain-Black-Wallpaper.png 2(0.219): http://www.ledr.com/colours/black.jpg 3(0.149): http://en.wikipedia.org/wiki/Black Mobile: 1(0.544): http://en.wikipedia.org/wiki/Black 2(0.126): http://en.wikipedia.org/wiki/Black_people 3(0.038): http://dictionary.reference.com/browse/black
[vent] Desktop: 1(0.612): http://www.ventrilo.com/ 2(0.208): http://www.ventrilo.com/download.php 3(0.059): http://www.thefreedictionary.com/vent Mobile: 1(0.666): http://www.thefreedictionary.com/vent 2(0.071): http://www.merriam-webster.com/dictionary/vent 3(0.064): http://www.ventrilo.com/

Another common intent type causing a discrepancy is dictionary or encyclopedia lookup. In the above examples, for [black], desktop users wanted to find images, while mobile users wanted to find wikipedia articles. For [vent], it meant an application with name "ventrilo" in desktop, but mobile users actually wanted to know the definition of "vent". In addition to these examples, many mobile queries are asking "what" or "who".

- Actionable Item Finding Intent Confidential

[sacramento penthouses] Desktop: 1(0.528): http://www.sacramentopenthouses.com/

Redacted

2(0.081): http://www.lstreetlofts.com/
3(0.072): http://sf.curbed.com/...penthouse_at_1170_sacramento.php

Mobile:
1(0.546): http://www.sacramentopenthouses.com/contact.htm
2(0.149): http://www.sacramentopenthouses.com/leasetoown.htm
3(0.135): http://downtownsac.org/...penthouses-at-capitol-park-2/

Mobile users have a tendency to find direct access to actionable items via fewer transitions. Accordingly, in many cases, this intent often appears that sub-pages in a site are preferred to the top gateway page. In the above example, mobile users clicked pages where they are able to perform some actions, such as "../contact.htm" or "../leasetoown.htm", over the gateway page. It makes sense in that browsing via page transitions in mobile devices are often costly.

- Other Interesting Differences

There are also other interesting minor intent differences. First, sometimes mobile users have different site preferences from desktop users even when the intents are almost identical. For query [shortcuts not working], the Apple forum is preferred to Microsoft forums in mobile. That is, some preferences depend on which operating system is more widely used in the interface.

Second, mobile queries tend to have common or casual interpretations, while desktop queries sometimes have industrial or professional interpretations. For query [high temp], the desktop query meant "high temperature metal", while the mobile query meant "high fever". It may be because mobile devices are seldom used in working environments, or because the user spectrum of mobile devices is wider that the intent is less biased towards a few demographic groups having special interests.

The last interesting minor intent is related to songs or media. Indeed, a considerable portion of queries with diverged intents are related to songs, including lyrics of, streaming, or downloading them. For [my pony download], the mobile query meant "ride my pony mp3 download". On the other hand, the desktop query is for downloading a desktop application "mipony".

Different User Behaviors On Direct Answers

From what we've learned earlier, mobiles users appear to favor fewer transitional interactions, e.g., more direct access to answers. Our one-boxes are considered as containers that provide direct answers, especially well corresponding to mobile user intents presented earlier. For one-boxes, will mobile users have different interactions than desktop users? To see this, we conducted some analysis on one-month long session data (5/16/2014-5/22/2014).

Metrics

The first metric of interest is the abandonment rate. A hypothesis is that users will abandon the

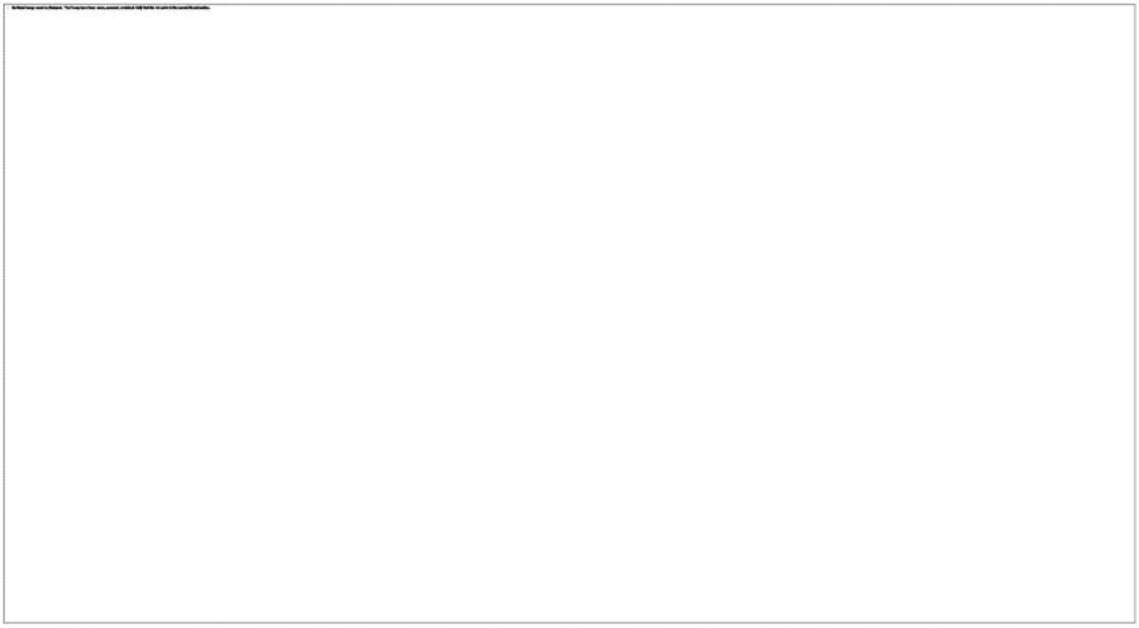
query if they are immediately satisfied with returned direct answers, thus user satisfaction upon direct answers may be measured in a form of abandonment. However, the abandonment is an ambiguous metric. For example, if a user issues multiple refined queries consecutively and abandons the last query, it is indeed more like a give-up, that is, a sign of total frustration with search results. Intuitively, abandonment can be either a positive or negative signal, depending on the context.

To focus on the positive side, we use singleton abandonment, which counts abandonment only in singleton queries, where singleton queries roughly mean first page search queries neither following nor followed by any refinement. A previous Google study shows that non-singleton abandonment is effective to measure dissatisfaction. Although singleton abandonment is not a clearly positive metric solely, we assume that it is so in the context of returned direct answers.

We computed the metric for a variety of interesting slices. The considered slices are as follows:

General	All queries
WebAnswer...	Queries triggering a WebAnswers or PorkyPig one-box
KC	Queries triggering a Knowledge Card or Panel
Weather	Queries triggering a Weather one-box
Finance+C...	Queries triggering a Finance or Currency one-box
Dictionary	Queries triggering a Dictionary one-box
IU	Queries triggering an Image Universal
LU	Queries triggering an Local Universal
PopularQuery	Top 50 queries common both in desktop and mobile
OrgRQuery	Top 100 queries common both in desktop and mobile, satisfying all of 1) NOT trigger any one-box 2) NOT have any well-known mobile apps
-KC	Queries not triggering a Knowledge Card or Panel

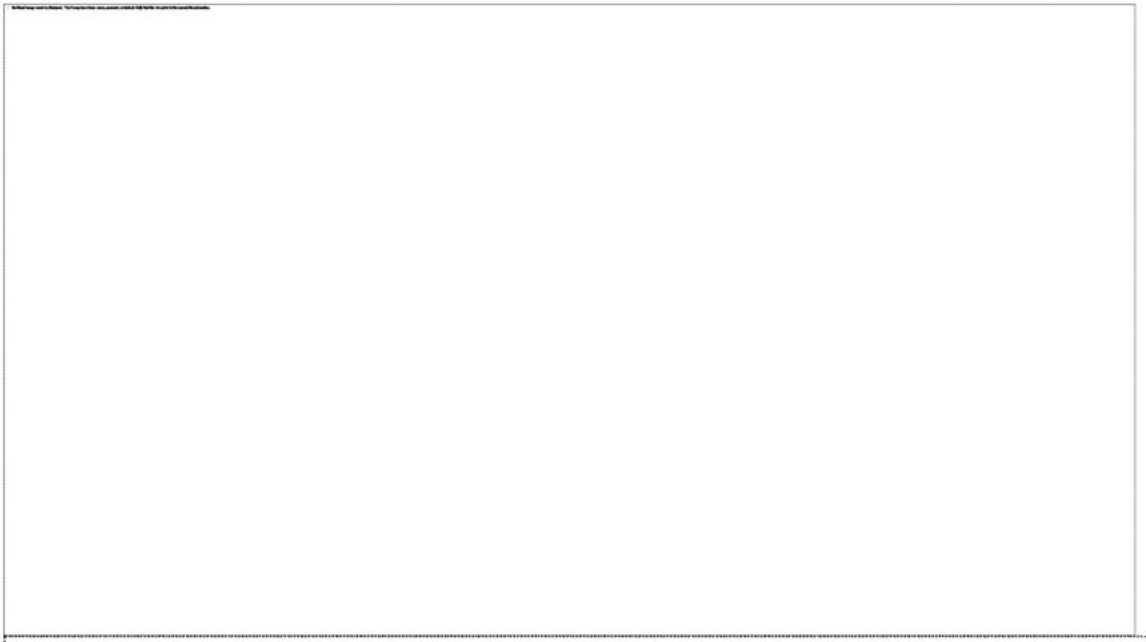
"PopularQuery" is included because we want to see the impact by interfaces solely, suppressing the difference in query streams. "OrgRQuery" is added to see user behavioral differences not affected by providing information for direct consumption. Furthermore, we also wanted to exclude the impact by existence of mobile apps, because we suspected that it may distort the metric.



As shown in the above, the abandonment rates are high for the one-boxes providing more direct answers, such as WebAnswers, Weather, Finance or Dictionary. We may assume higher satisfaction for those types of queries. Also, the metric values are even higher in mobile. An interesting question is why mobile abandonment is higher than desktop for the same queries which are mostly navigational queries, as appeared for "PopularQuery". For those interested in this, please see an analysis result that may answer to the question, in Section 4.3.

Satisfaction appears not only in the form of abandonment, but also of Confidential which is widely used as a clearly positive signal at Google. We tried to roughly estimate a query satisfaction level by summing sington abandonment and Confidential here is defined in a query level (or SERP level) rather than an individual result level or an one-box, because we want to model query satisfaction in the same level as abandonment.

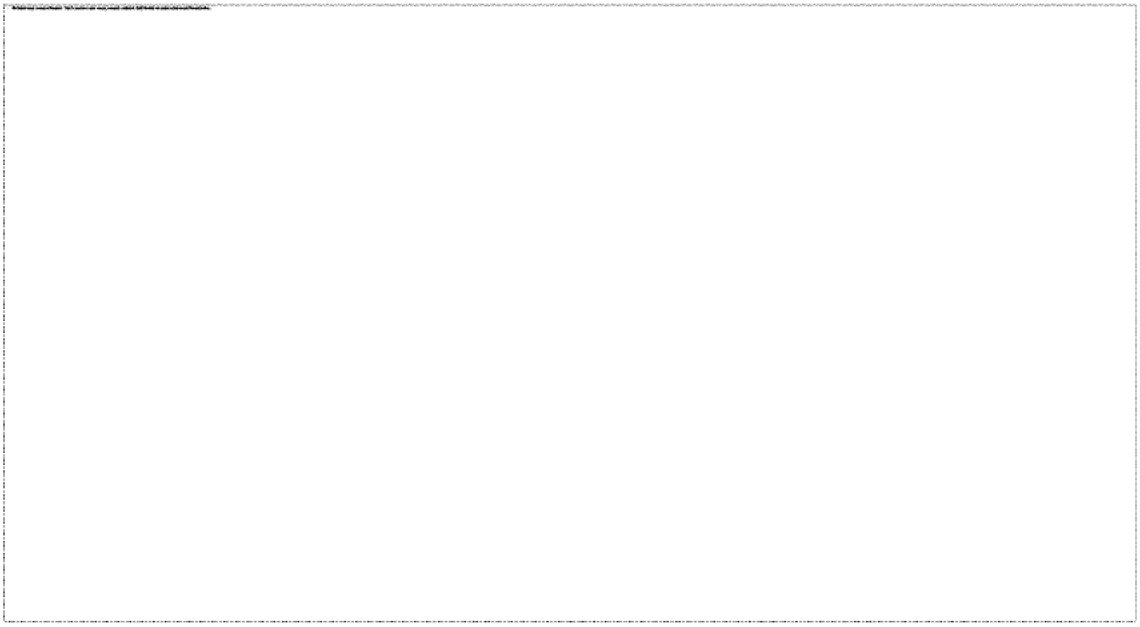
Redacted



As shown in the combined metric, the satisfaction levels in mobile and desktop are very similar for most slices. This is as predicted, because we don't expect different satisfaction levels depending on interfaces for general queries. Though, WebAnswers, Weather and Dictionary make exceptions. For those, mobile looks like a clear win. We assume that mobile users have stronger intents for these types of queries, which are being satisfied well, although it is not easy to be conclusive at this moment. Indeed, evaluating how satisfactory Google search is, would be beyond the scope of this study, thus we want readers to focus more on the tendencies.

We can derive another metric combining these to see another aspect of the diverged user behaviors. The ratio of singleton abandonment to (singleton abandonment **Confidential** provides us with how much users favor direct information consumption (appearing in a form of abandonment) over **Confidential**

Redacted



For WebAnswers, Weather, Finance and Dictionary, direction information consumption is shown to be favored. This makes sense because these types provide more direct answers. Also, in mobile, this tendency for WebAnswers, Weather, and Finance (not for Dictionary here because presumably Dictionary's abandonment is already too high to make any difference), is even stronger compared to desktop. Hence, we think that mobile users have stronger intent to consume direct answers.

Divergence in Query Streams

The differences in user behaviors on direct answers can be partially due to diverged query streams, that is, different queries are issued in mobile and desktop. We tried to quantify the difference by computing Jaccard similarity between the sets of top 100 queries selected for each slice in mobile and desktop. As references, the similarities between browsers in the same interface are also computed.

As seen in the above, even "General" (indeed, this is a popular navigational stream, because top 100 of all queries were selected), has much lower similarity than references. The similarities for WebAnswers and Dictionary are even lower. That is, mobile queries are quite different from desktop queries, and the difference is even bigger given query types related to direct answers. In the followings, we present the diverged patterns with examples, according to query types.

- Differences in Popular Query Streams

For evaluating query streams, we selected and present top mobile-only queries and desktop-only queries that don't belong to the intersection of the top 100 queries.

[maps]	[backpage]
[pandora]	[worldstarhiphop]
[cool math games]	[youjizz]
[yahoo.com]	[porn hub]
[chase]	[xvideo]
[you tube]	[youporn]
[linkedin]	[pizza hut]
[capital one]	[lebron james]
[google drive]	[nba]
[zillow]	[iphone 6]
[aol.com]	[freeporn]

Even popular query streams look substantially different. Most desktop queries are typical navigational queries to big sites. Some mobile queries are also navigational, but the navigation target sites of the queries tend not to have well-known mobile apps. Porn intent is prominent in mobile, again. Lastly, while all desktop queries are related to sites, we can

see more entity queries in mobile, for example, [lebron james] or [iphone 6].

- Differences in Question Answering Query Streams

The same procedure was applied to queries triggering WebAnswer one-box.

[how to take a screenshot on a mac] [copyright symbol] [social security benefits] [how to take a screenshot] [social security disability] [how much is a stamp] [facebook cover photo size] [paypal fees] [ein number] [area code 201] [how to calculate gpa]	[mcdonalds breakfast hours] [fifa meaning] [how long is a soccer field] [how to delete instagram account] [what does thot mean] [std symptoms] [what does fifa mean] [jay z net worth] [who won the last world cup] [ios 8 release date] [bill gates net worth]

Some desktop queries have instruction seeking intent, e.g., beginning with "how to". In desktop, work intent is prominent, e.g., [copyright symbol]. Many desktop queries have somewhat unclear answer seeking intent. For example, for queries such as [social security disability] or [area code 201], the answer seeking intent is not obvious, although we may provide broadly relevant answer sets.

On the other hand, many mobile queries look like factoid questions. The answer seeking intent is much clearer, as many queries begin with "how", "what" or "who". Again, local business intent is observed here, such as [mcdonalds breakfast hours].

Given the diverged query stream, we wondered how much abandonment difference was caused by the stream difference. We tried to measure abandonment rates for the top 100 common WebAnswers queries. Mobile abandonment for the common queries is Confidential higher than desktop, which is smaller compared to that mobile has Confidential higher abandonment for all WebAnswers queries. It proves that a considerable portion of the abandonment difference is due to the query stream difference but the user behavioral differences still affect the metric difference a lot.

- Differences in Dictionary Lookup Query Streams

Lastly, the difference in dictionary lookup query streams is investigated.

[colleague] [unfortunately] [miscellaneous] [fiance]	[goof] [what does smh mean] [what is the olecranon] [cancel]

[available]	[what is anthrax]
[convenience]	[what does hmu mean]
[verbiage]	[what does wcw mean]
[opportunity]	[so]
[liaison]	[mediocre]
[discrepancy]	[what is lupus]
[privilege]	[optimistic]

Most desktop queries are single-word queries. Indeed, it looks as if users issue the queries for the purpose of spell-checking. On the other hand, many mobile queries clearly request the definitions, using natural language. Also, an intent finding what an abbreviation or acronym stand for is prominent in mobile.

Abandonment Type Classification

We did a simple analysis answering to a question why abandonment in mobile is higher even for the same navigational queries. Query [amazon] was chosen as a testing query because this popular query is mostly navigational and has somewhat homogeneous shopping intent that we can easily figure out. We looked into at 150 sessions having abandonment for [amazon] and manually classified them according to abandonment types.

[amazon]

The y-axis represents the actual abandonment value, thus each color in the stack corresponds to how much actual abandonment can be explained by each type. "SHOP" means shopping intent, when the previous or next queries have clear product search or shopping site intent, e.g., [ebay]. This abandonment type may be assumed as hesitation, for example, between amazon or ebay for a product. "INTERRUPT" was assigned to a presumable interruption when the duration between [amazon] and the next query is long. We fairly imagine a number of interruption scenarios in mobile, such as receiving a call, or opening an messenger app upon a notification. "DUP" is for duplications of [amazon]. In mobile, refetching the previous query frequently occurs because of software or hardware reasons, such as small browser cache. "?" is a type that we cannot infer the context of the abandonment. For example, the previous query

1 minute ago is a celebrity name and the next is a navigational query without shopping intent in 1 minute, where no duplication is found. The user may check the amazon stock price or look for company information presented in Knowledge Panel. However, it is not possible to verify these conjectures by looking at sessions. "UNINTENDED" means queries by Psychic pause actions. This abandonment can be assumed that [amazon] was not intended but input by Psychic.

As shown in the graph, "DUP" and "INTERRUPT" together explain a big chunk of abandonment. Given that "UNINTENDED" cannot exist in mobile because Psychic has been enabled for a very small traffic in mobile, the other key difference comes from "?". Unfortunately, we don't have very good explanations for this right now. We will leave this for any future analysis.

Conclusions

Our analysis shows that mobile queries sometimes have different user intents. More specifically, local search, dictionary or encyclopedia lookup, direct answer seeking, actionable item finding intents are major sources causing the divergence. These intents imply mobile search as a personal assistance who can immediately provide information or knowledge given a situation or context.

We also showed that user behaviors in mobile are often quite different. A fundamental difference is that users favor less transitional interactions, such as result clicks. Moreover, in mobile, people tend to want direct information consumption.

Based on these findings, we may need to consider the following:

- More features allowing fewer transitional interactions and more direct information consumption. To allow more users to pleasantly consume direct information for various information needs in mobile, the traditional search result presentation containing about 10 page links with short snippets doesn't sound enough. Our one-boxes do the job, and more one-boxes may be better. Though, any revolutionary features or UXs can be effective ultimately. The stream UI in development can be an example.

- Separate mobile ranking signals or evaluation reflecting different intents. Mobile queries often have different intents, and we may need to incorporate additional or supplementary signals reflecting these intents into our ranking framework. As discussed earlier, it is desirable that these signals handle local-level breakdowns properly. Also, these diverged intents may need to be reflected on evaluation processes.

Finally, mobile search is still in growth mode and dynamically evolving. What we presented here may be only a few aspects from the current time slice. Therefore, mobile search should be regularly studied in various angles to update our understandings.

Ranking Leads meeting

At this week's leads meeting Alex Kojoukhov and Lina Martensson presented their work on BScorer. Their work was divided into four major parts:

Confidential

Launch approvals

Confidential

Metrics and launches

Confidential

Redacted

Tool Time: Retrieval debugging

Confidential

Confidential

Go to page: **Ranking Newsletters » 2014 Q3 Ranking Newsletters » Aug 11 -- Aug 15, 2014**

You requested this notification from Google Sites. You can unsubscribe at any time.
Don"t want to get notification of your own changes? Change your settings.

Redacted