Characterizing Users' Multi-Tasking Behavior in Web Search

Rishabh Mehrotra
Dept. of Computer Science
University College London
r.mehrotra@cs.ucl.ac.uk

Prasanta Bhattacharya
Dept. of Information Systems
National University of Singapore
prasanta@comp.nus.edu.sg

Emine Yilmaz
Dept. of Computer Science
University College London
emine.yilmaz@ucl.ac.uk

ABSTRACT

Multi-tasking within a single online search sessions is an increasingly popular phenomenon. In this work, we quantify multi-tasking behavior of web search users. Using insights from large-scale search logs, we seek to characterize user groups and search sessions with a focus on multi-task sessions. Our findings show that dual-task sessions are more prevalent than single-task sessions in online search, and that over 50% of search sessions have more than 2 tasks. Further, we provide a method to categorize users into focused, multi-taskers or supertaskers depending on their level of task-multiplicity and show that the search effort expended by these users varies across the groups. The findings from this analysis provide useful insights about task-multiplicity in an online search environment and hold potential value for search engines that wish to personalize and support search experiences of users based on their task behavior.

CCS Concepts

•**Information systems** → **Web mining; Task models;**

Keywords

Task extraction, Task multiplicity, Web search interests, User types

1. INTRODUCTION

Search engine users' information needs tend to span a broad spectrum [4]. While simple needs, such as homepage finding, can mostly be satisfied via a single query, users may also issue a series of queries to collect, filter, and synthesize information from multiple sources to solve a task. Given the inherent diversity in information needs, users engage with search systems in varied ways. Further, and as a direct result of the increasingly complex informational environment around us, users are increasingly engaged in multitasking and information task switching behaviors. Multitasking is the ability of humans to simultaneously handle the demands of multiple tasks through task switching [1, 5]. However, many interactive technologies do not provide effective support for managing multitasking behaviors of users.

Web search engines offer a typical environment where users perform multiple tasks across diverse contexts. For example, a programmer searching for solutions to a bug in his code, might take a brief hiatus to listen to some music. The two tasks described here need not be at the same level of importance for the user, nor must they be performed in parallel. While such situations are commonly observed in our daily search behavior, not much is understood about the kind of users who indulge in such multi-tasking behavior or even the extent or nature of such multi-tasking behavior in major search engines. This research gap stems mainly from the difficulty in identifying and quantifying multiple task completions from observational data. In the current study, we leverage search logs from a large-scale search engine to provide a detailed analysis of multi-tasking behavior for different user groups, and across multiple session sessions over a 30-day period. We seek to provide evidence that multi-tasking has emerged as a dominant characteristic of online search behavior and that users have varying propensities to indulge in such multi-tasking.

Different from existing studies on multi-tasking which have used topics of queries as proxies for tasks, we make use of an explicit search task extraction framework to extract the task information from web search sessions. This allows us to provide richer insights on the prevalence of task multiplicity in search sessions. Making use of real world search logs, we first quantify the extent of multi-tasking behavior in search sessions and show the existence of user groups based on multi-tasking behaviors. We go a step further in analyzing the user groups on a number of search interaction metrics and quantify the differences in these user groups based on how they interact with search systems. Our results motivate the need for designing search interfaces specifically catered to the needs and multi-tasking behavior of users.

2. RELATED WORK

Recent studies suggest that users' information search may have multiple goals or topics and occur within the broader context of their information-seeking behaviors [3]. Through an online survey, Wang *et al.* [9] show that 92% of the participants reported to have participated in online sessions where they accessed several sites, to perform between 2 to 8 tasks. In the context of web search sessions, most work on multi-tasking has been based on user studies [4, 7]. Other

CHIIR '16, March 13-17, 2016, Carrboro, NC, USA
© 2016 ACM. ISBN 978-1-4503-3751-9/16/03...$15.00
DOI: http://dx.doi.org/10.1145/2854946.2855006

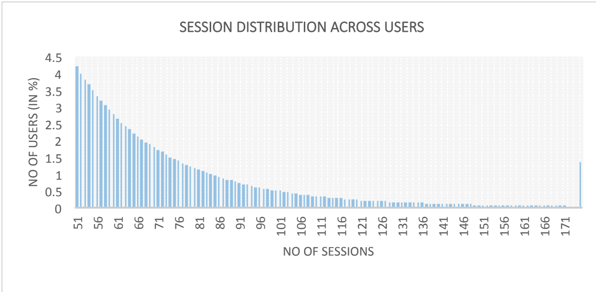

Figure 1: The variations of number of sessions per user across the user population.

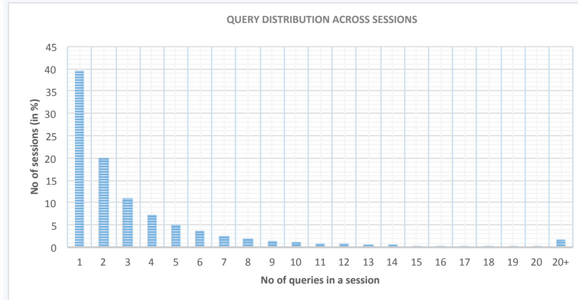

Figure 2: The variations of the number of queries across sessions.

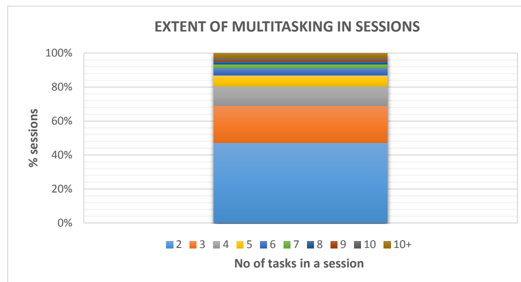

Figure 3: Quantifying the extent of multi-tasking in search sessions: .

studies do not explicitly refer to online multi-tasking, but provide useful insights. For instance, users access different sites during a session [2] and a large proportion of pages are visited more than once. In addition, the frequency at which a page is revisited differs depending on user habits and the type of website [2, 6], or in other words, the web tasks a user accomplishes on the site.

All these provide a strong evidence that multi-tasking during online sessions exists and depends on the web tasks. We depart from existing work that rely on using topics as a proxy to identify tasks by making use of an explicit task identification algorithm to identify users' search tasks. This approach enables us to provide estimates and qualitative insights on the extent of multi-tasking behavior in web search tasks.

3. CHARACTERISING TASK BEHAVIOR ACROSS SEARCH SESSIONS

In this study, we seek to characterize user behavior in online search sessions based on task specificity and multiplicity. While users generally perform a single task in a single search session, the task process might get interrupted by other competing tasks that become salient in the particular context. For example, a programmer attempting to search for bug solutions on the search engine, might choose to take a short break to listen to music from her favorite musician or band. The search session, in this case, would transform into a multi-task session. Given this backdrop, we contend that it is imperative for the search engine to understand the type of users who might be more prone to multi-tasking within a single session, and also the type of tasks that might be more susceptible to interleaving or interference by competing tasks. In the current study, We formulate the following 3 research questions, and offer preliminary evidence from a large scale observational dataset on search behavior to answer these questions.

RQ1: *Quantifying the Extent of Multi-tasking in Online Search Sessions.*: While it is well known that online search sessions often tend have interleaving of multiple tasks, but what has been largely ignored are the heterogeneities at a user level and/or a search session level. Specifically, we aim at quantifying, first, the prevalence of multi-tasking behavior in online search sessions (i.e. *how common is multi-tasking?*), and second, the extent of multi-tasking behavior in multi-task sessions (i.e. *how many tasks on average are there in multi-task search sessions?*).

RQ2: *Uncovering User-level Heterogeneities in Multi-tasking Behavior*: Through the current study, we also seek to uncover the presence of user-level idiosyncrasies in multi-tasking behavior in search sessions. Specifically, we attempt to understand the proportion of sessions per user that are single tasked vs. multi-tasked. Consequently, we seek to uncover any underlying categorizations among the users based on the extent of their multi-tasking behavior (i.e. *Can we identify and classify groups of users who demonstrate similar proportions of multi-tasking behavior?*. Uncovering such user groups would pave the way for the search engine to provide better personalized search assistance based on the group-level features and characteristics.

RQ3: *Characterizing Task Effort across User Groups*:The presence of competing or interfering tasks within a single session could accentuate or attenuate the search effort expended by the users. Specifically, we wish to understand the relationship between task multiplicity and total effort expended by the users (i.e. *do users who multitask more(less) expend more effort than users who multitask less(more)?*. In the context of the current study, we operationalize search effort using the query time, the average length of queries etc.

4. DATA CONTEXT

We use backend search logs for users of a major US-based search engine for a period of 30 days from May 1, 2015 to May 31, 2015 and choose a random sample of over 2 million users where each user is identified by a unique IP address. Over the 30-day period, the users participated in a total of over 200 million search sessions comprising one or more search queries, as illustrated in Fig. 2. We also observe that most users participate in 50 to 100 sessions in the 30-day period as clear from Fig. 1. We filter out inactive users from our dataset who participate in <50 sessions, and focus instead on the more active user population.

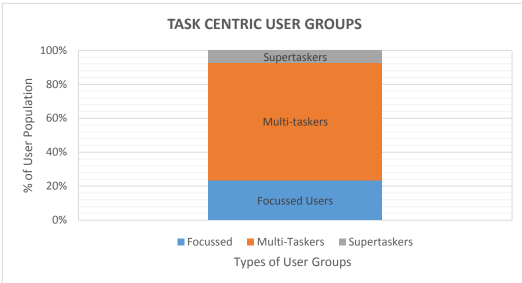

Figure 4: User groups based on multi-tasking behaviors.

User Type	Example Queries from a Typical Session
Focused User	"test guide.com", "CNA Practice Test", "CNA State Board Exam", "CNA Testing Schedule and Locations", "CNA State Board Practice Test", "CNA Practice Test 2014", CNA 50 Questions Test", "Free GED Practice Test 2014"
Multi-Tasking User	"Gravity FSX 2.0", "Full Suspension Mountain Bikes", "Walmart Cards", "Walmart Instant Card Application", "Gravity FSX 2.0 price", "full suspension bike sale"
Supertasking User	"hairstyles for women over 50", "thin wavy hairstyles for women", "facebook", "fb sign in", "pulled pork crock pot recipe easy", "Slow-Roasted Pulled Pork", "barefoot contessa", "miley cyrus hair styles", "hairstyler.com"

Table 1: Example query sessions from different types of users.

5. EXPERIMENT SETUP

We next describe our experimental setup wherein we first extract search tasks from search logs and then proceed to analyze the multi-tasking behavior of searchers.

5.1 Task Extraction

Search tasks, comprising a series of search queries serving the same information need, have recently been recognized as an accurate atomic unit for modeling user search intent. For our analysis, we make use of the Latent Structural SVM framework [8] for task identification. Given query sequences within sessions, search tasks are identified by clustering queries into tasks by find the strongest link between a candidate query and queries in the target cluster (*bestlink*). This is achieved by making use of a structural learning method with latent variables, i.e., latent structural SVMs, to utilize the hidden structure of query inter-dependencies to explore the dependency among queries within the same task.

Given a query sequence $Q = q_1, q_2, ..., q_M$, a feature vector for the task partition y is specified by the hidden best-link structure h as $\psi(Q, y, h)$. Based on $\psi(Q, y, h)$, the bestlink SVM is a linear model parameterized by w, and predicts the task partition by,

$$(y*, h*) = argmax_{y,h} w^T \psi(Q, y, h) \quad (1)$$

where Y and H represent the sets of possible structures of y and h respectively. $y*$ becomes the output for cross-session tasks and $h*$ is the inferred latent structure. A detailed overview of the approach can be found in Want *et al.* [8].

5.2 Analysis of Multi-task Search Behavior

Fig. 3 illustrates the prevalence and extent of multi-tasking behavior in our data. In our dataset, we find close to 90 million search sessions which have 2 or more completed tasks. Among these 90 million search sessions, there is a varied distribution of task multiplicity as described by Fig. 3. Specifically, we observe that while over 60% of the sessions have 2 or 3 tasks, about 20% of the sessions have 5 or more tasks. On an average, however, each user participates in 76 sessions in which she performs an average of 2 tasks.

Next, we look at the user-level heterogeneities in search sessions. Specifically, we categorize users based on the average number of tasks completed by the user across all sessions in the 30-day period. We observe that a sizable number of users (>20%) perform just a single task on average across their session history. We call such users *focused users*, as their search behavior is focused on a specific task, free from interference of competing tasks. On the other extreme, we also find a small group of users who perform 4 or more tasks on average across their session history. We call such users *supertaskers* who perform several tasks within a single session. We categorize all the other users as *multi-taskers* who completed between 2 and 3 tasks on average in their session history. The density of users across each of the three groups has been better depicted in Fig. 4. Interestingly, from Fig. 4, we observe that most users are not focused in their search behavior, and tend to complete at least 2 tasks within a session. This is not entirely unsurprising, given that one of the tasks could be the primary (e.g. search for solution to a programming bug on the Internet) or important task, while the others might be ancillary tasks (e.g. listen to music, check weather updates). Table 1 provides a list of sample queries executed by users across the three user groups.

Finally, we analyze search effort across the three user groups identified in the study. We characterize search effort using 4 different metrics viz. *(i) Time to First Click (TTFC)*, *(ii) Time to Last Click (TTLC)*, *(iii) Page Click Count (PCC)*, *(i) Pagination Click Count (PgCC)*. The TTFC metric measures the time elapsed before the user clicks the first link on the query result page. A longer TTFC is an indication of user surprise or confusion with the search results, and hints at a more extensive cognitive elaboration process as the user decides which link to click on. The TTLC metric, on the other hand, measures the time elapsed before the user clicks the final link in her search session. The TTLC is a more direct measure of search effort expended by the user within a particular session. A higher TTLC could indicate that the user dissatisfaction with the early results provided by the query results, or a heightened motivation on part of the user to search more about the particular topic of interest.

The PCC metric measures the total number of clicks made by the user on the query result page, while the PgCC metric measures the number of times the search page was incremented or decremented by the users. Both these metrics are direct measures of the search intention of the user, and are hence good proxies to capture the different facets of search effort. We compute each of these metrics across each of the user groups as defined earlier, and highlight our findings in Figure 5.

Our analysis indicates that Supertasking users have a much higher TTFC and TTLC scores, but a lower PCC score than the Focused and Multi-Tasking groups. This supports our conjecture that most supertaskers perform multiple tasks in a master-slave fashion, where they focus bulk of their attention on a focal task, while being periodically distracted by ancillary tasks (e.g. music, weather updates). This periodic distraction causes a decrease in attention span on the focal

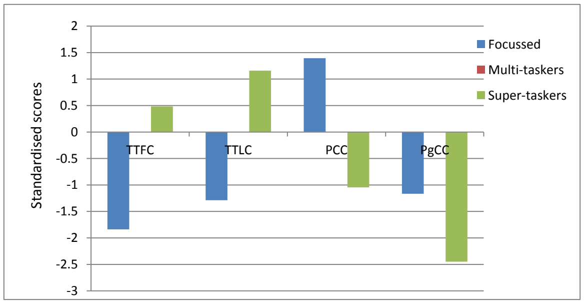

Figure 5: Differences in user groups quantified via effort based metrics. The scores reported are deviations from the Multi-taskers group which is held as baseline. All numbers are standard scores (Z-scores).

task, which is manifested by a decrease in click count, and an increase in click delays on the focal task. We do not, however, find any noticeable difference in the PgCC scores across the groups.

6. CONCLUSION AND FUTURE WORK

This study is among the first to analyze task behavior in web search sessions using large-scale and objective search logs. Specifically, we emphasize that, contrary to popular understanding, most users on search engines are multi-task users performing 2 or more tasks within a single search session. We also provide evidence of "Supertaskers" who perform onwards of 4 tasks within a single session. This widespread prevalence of task multiplicity makes it imperative for search engines to refocus their personalization and recommendation strategy towards a task-oriented view. For example, if search engines can fully identify and characterize the number and types of tasks performed by a given population of users on the engine, they could potentially optimize the session to better fit specific user- and task-based needs, while also making potential task-recommendations to reduce the search effort, as quantified in this paper.

Yet another finding we wish to highlight through our study is the characterizing of multiple tasks into a combination of a single primary and multiple ancillary tasks. Our task effort scores provide preliminary evidence to suggest that such categorization of multiple tasks into a task-hierarchy might indeed be plausible. Such insights are useful for search engines in that they could reduce task-transition delays and make design improvements to reduce cognitive loads in such multi-task sessions.

In future work, we wish to perform a deeper task-level analysis to uncover possible topical differences among tasks that occur frequently in multi-task sessions, versus those that occur frequently in single-task sessions (i.e. *Are tasks in single-task sessions qualitatively different than tasks prevalent in multi-task sessions?*). Understanding such task-level differences would provide a stepping stone towards understanding specific task-characteristics that might make it more or less susceptible to interference and/or distraction. Yet another direction of future work is about making more accurate task-predictions within a single search session, as well as across multiple sessions for a given user. With a steadily improving understanding of task and search behavior online, we envision a day when the search engine would be able to infer user-,task- as well as session- level characteristics based on just the first query issued by the user, and personalize the search experience accordingly.

7. REFERENCES

[1] M. A. Just, P. A. Carpenter, T. A. Keller, L. Emery, H. Zajac, and K. R. Thulborn. Interdependence of nonoverlapping cortical systems in dual cognitive tasks. *NeuroImage*, 14(2):417–426, 2001.

[2] R. Kumar and A. Tomkins. A characterization of online browsing behavior. In *Proceedings of the 19th international conference on World wide web*, pages 561–570. ACM, 2010.

[3] J. Lehmann, M. Lalmas, G. Dupret, and R. Baeza-Yates. Online multitasking and user engagement. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, pages 519–528. ACM, 2013.

[4] C. Lucchese, S. Orlando, R. Perego, F. Silvestri, and G. Tolomei. Identifying task-based sessions in search engine query logs. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 277–286. ACM, 2011.

[5] M. Miwa. User situations and multiple levels of user goals in information problem solving processes of askeric users. In *Proceedings of the ASIST Annual Meeting*, volume 38, pages 355–71. ERIC, 2001.

[6] H. Obendorf, H. Weinreich, E. Herder, and M. Mayer. Web page revisitation revisited: implications of a long-term click-stream study of browser usage. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 597–606. ACM, 2007.

[7] A. Spink, M. Park, B. J. Jansen, and J. Pedersen. Multitasking during web search sessions. *Information Processing & Management*, 42(1):264–275, 2006.

[8] H. Wang, Y. Song, M.-W. Chang, X. He, R. W. White, and W. Chu. Learning to extract cross-session search tasks. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1353–1364. International World Wide Web Conferences Steering Committee, 2013.

[9] Q. Wang and H. Chang. Multitasking bar: prototype and evaluation of introducing the task concept into a browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 103–112. ACM, 2010.