Research Paper

Identification and Analysis of Multi-tasking Product Information Search Sessions with Query Logs

Xiang Zhou, Pengyi Zhang[†] & Jun Wang

Department of Information Management, Peking University, Beijing 100871, China

Citation: Xiang Zhou, Pengyi Zhang & Jun Wang (2016). Identification and Analysis of Multi-tasking Product Information Search Sessions with Query Logs.

Received: Mar. 16, 2016
Revised: May 24, 2016
Accepted: Jun. 6, 2016

Abstract

Purpose: This research aims to identify product search tasks in online shopping and analyze the characteristics of consumer multi-tasking search sessions.

Design/methodology/approach: The experimental dataset contains 8,949 queries of 582 users from 3,483 search sessions. A sequential comparison of the Jaccard similarity coefficient between two adjacent search queries and hierarchical clustering of queries is used to identify search tasks.

Findings: (1) Users issued a similar number of queries (1.43 to 1.47) with similar lengths (7.3–7.6 characters) per task in mono-tasking and multi-tasking sessions, and (2) Users spent more time on average in sessions with more tasks, but spent less time for each task when the number of tasks increased in a session.

Research limitations: The task identification method that relies only on query terms does not completely reflect the complex nature of consumer shopping behavior.

Practical implications: These results provide an exploratory understanding of the relationships among multiple shopping tasks, and can be useful for product recommendation and shopping task prediction.

Originality/value: The originality of this research is its use of query clustering with online shopping task identification and analysis, and the analysis of product search session characteristics.

Keywords Product search; Shopping task identification; Shopping task analysis; Multi-tasking session

† Corresponding author: Pengyi Zhang (E-mail: pengyi@pku.edu.cn).

JDIS
Journal of Data and Information Science
Vol. 1 No. 3, 2016
pp 79–94
DOI: 10.20309/jdis.201621
http://www.jdis.org

1 Introduction

Online shopping has gained great popularity among consumers because it is fast, convenient, and unrestricted in terms of time of day and product locale. The Internet has greatly lowered the cost and increased the efficiency of shopping compared to in-person searches, especially for alternative or substitute products, as it enables consumers to quickly collect more information about a wide range of products, brands, and sellers before they make purchasing decisions. Information search, identified by consumer behavior research as the first stage in the buying process (Rowley, 2000), thus becomes more important in online shopping than in traditional retailing. Online shopping is more "information intensive," meaning that the e-commerce websites intended for transactions become increasingly vital and comprehensive information sources (Fortune, 1998).

According to *The Research Report of Online Shopping Market in China, 2014* (CNNIC, 2015), online retail transactions reached a revenue of 2.79 trillion Yuan with a yearly growth of 49.7%. Online shopping is undergoing a rapid growth in China, but research on consumers' online search behavior is rather limited. More study is needed to better understand the characteristics of online consumer search behavior to improve e-commerce sites, consumer services, and sales.

Identifying the specific patterns related to how consumers seek information has always been critical for understanding consumer buying behavior trends (Bhatnagar & Ghose, 2004), and has important implications for decision-making tasks such as purchasing a product. Research has found that multi-tasking is quite common in Web search (Ye & Wilson, 2014). For example, Spink, Ozmutlu, and Ozmutlu (2002) found that 11.4% of 1,000 randomly extracted sessions involved multi-tasking. Spink et al. (2006) found that in sessions with more than three queries, more than 90% included multi-tasking. It is common for online shoppers to search multiple product categories simultaneously when making multiple purchases. Very little research has been done on product information searches, however, to identify and analyze the characteristics of multi-tasking product search, which are different from standard Web search queries. This research aims to bridge this gap.

The availability of clickstream data has contributed greatly to information seeking research for many tasks, including online shopping. In this paper, we analyze query terms from click-through logs to identify consumers' shopping tasks, and to discover the characteristics of their multi-tasking product searches.

Definitions of the important concepts of session and shopping task in this study are:

1) A *session* is a series of queries by a single user made within a small range of time, which is meant to capture a single user's attempt to fill a single information need. In this research, we use the heuristic that queries for single

information queries become clustered over time, followed by a gap of up to 45 minutes before the user returns to that search engine (Moorthy & Talukdar, 1995).

2) A *shopping task* is a set of activities that a consumer conducts in order to purchase a product. A multi-tasking session refers to the consumer product search conduct for multiple shopping tasks.

This paper first reviews the related literature, followed by a description of the methodology and findings on characteristics of multi-tasking product search. We conclude with an analysis of the results, and discuss the limitation and implications of the research as well as future study suggestions.

2 Literature Review

2.1 Task Identification in Web Search

Previous research has identified two types of approaches for task identification: time splitting and query clustering (Lucchese et al., 2013). Query clustering is based on the content of the queries while time splitting uses contextual cues. Content-based methods to identify search tasks in Web search include comparisons of (1) similarities of two search queries, (2) URLs that the Web search engine returns (Glance, 2000), and (3) documents that the Web search engine returns (Raghavan & Sever, 1995). Similarity scores are calculated based on these three indexes to decide whether two queries belong to the same search task.

The two major methods used herein for comparing the relevance of these two search queries are (1) identifying word similarities in the queries and extracting the sets of the search terms from these two queries. Some useful indexes for this task include the Jaccard distance (Järvelin, Järvelin, & Järvelin, 2007), which calculates the ratio of the intersection and the union of the two search-term set and the Levenstein distance (Jones & Klinkner, 2008), and (2) comparison of the semantic relevance of the search terms by using the idea of vector space (Salton & Mcgill, 1986). For example, utilizing the semantic relation from Wiktionary and Wikipedia, Lucchese et al. (2011) calculated similarities between each search term and each source in the semantic network, and created a search term vector composed of the similarities between a search term and each source in the semantic network.

Usually the angle (cosine similarity) between two search query vectors is calculated as the index of the similarity between these two search queries. Lucchese et al. (2011) first processed the search log, including the removal of empty log records and stop words, as well as stemming and deleting sessions that last too long or include too many queries, which indicates it is likely produced by machines. Then they calculated the word and semantic similarities between queries using two methods to calculate the final similarity index. The first method is a weighted average of the word similarity and the semantic similarity, whereas the second method is to use a threshold. When the word similarity score is above the threshold, the final similarity index score equals the word similarity; when the word similarity score is lower than the threshold, the final similarity index is the greater value of the word similarity and the semantic similarity.

2.2 Multi-tasking Web Search

Information users often demonstrate multi-tasking behaviors in Web search. Spink et al. (2006) suggested that users generally produce multi-tasking sessions for two reasons. The first reason is that a user may have several search topics at the beginning of the search process, and the second reason is that although users may have only one search topic in the beginning of the search process, they may discover new search topics in relation to information needs while searching.

Numerous studies have examined the characteristics of multi-tasking search sessions, including the time involved in queries. For example, Spink, Ozmutlu, and Ozmutlu (2002) found that the length of search queries and the time costs in multi-tasking sessions are longer than those in mono-tasking sessions. Lin and Belkin (2005) also confirmed that the average number of search queries used in multi-tasking sessions is more than that in mono-tasking sessions. When Lucchese et al. (2011) analyzed the search logs of 307 search sessions and 1,424 queries from American On Line (AOL), they found that the average duration of each search session was about 15 minutes. The shortest session lasted for less than one minute and had only one or two queries, while the longest session lasted for about two and a half hours. There were on average 4.49 queries in one search session, where half of the sessions had fewer than five queries. The logs were divided into 554 search tasks, and the average number of queries per task was 2.57. On average, a session included 1.8 tasks. Within the total 307 sessions, there were 162 (52.8%) with only one search task, while the rest (47.2%) were multi-tasking search sessions. The number of queries in the multi-tasking sessions was 1,046, which accounts for 74.0% of total queries.

In another study of AltaVista (Spink et al., 2006), researchers found that among the 254 two-query sessions, 206 (81.1%) involved more than one task. There were 254 sessions that included two queries, 206 of which (81.1%) were multi-tasking sessions. There were 483 sessions that included more than two queries, 441 of which (91.3%) were multi-tasking sessions. In the multi-tasking sessions, there were on average 3.2 tasks per session.

Wang et al. (2013) analyzed the search logs collected from Bing.com, a dataset that includes 7,628 users, 37,547 sessions, and 114,723 queries. On average a user participated in 4.9 sessions and made 15.1 queries. There were 8,044 (77.9%) tasks that included only one query, 2,283 tasks (22.1%) that included more than one query, and 1,307 multi-session tasks. The average amount of tasks that a user performed was 7.2. Tasks that generally involved more than one query consisted of 2.8 sessions and 6.6 queries, where the task needed 491.1 minutes to finish.

3 Experiment

3.1 Data Collection and Preprocessing

In order to identify and analyze the characteristics of consumer multi-tasking product Web search, we performed a series of experiments on large-scale product search log records from taobao.com. The whole dataset includes browser click-through logs of 4,285 users with 81,759 sessions from taobao.com during the month of May, 2013. The whole dataset contains 1,410,960 records from 81,759 sessions (Yuan, 2014). Each record contains the following fields:

- Uid: a uniqueuser code assigned to identify a user;
- IP address: the IP address from which a click is made;
- URL: the URL of the Web page a user visited;
- Date and time: the starting time a user opened a certain URL in a browser window;
- Staytime: the duration in seconds a user stayed active on a Web page;
- Query terms: queries as entered by a user (if any);
- Sessionid: a unique session identifier marking the session a record belongs to.

Figure 1 shows some sample log records.

uid	ip	url	date	staytime	url_kw		sessionid
12131355551097960974	121.10.248.46	http://login.taobao.com/member/loginByIm.do?_input_charset=utf-8	2013-05-14 10:34:02	0	NULL		122899
12131355551097960974	121.10.248.46	http://item.taobao.com/item.htm?spm=a1z10.1.w7344075-17592727432.15.ZJ6O2L&id=14193849774	2013-05-14 10:38:38	5	NULL		122899
12131355551097960974	121.10.248.46	http://item.taobao.com/item.htm?spm=a1z10.1.w3-18358782537.15.ZJ6O2L&id=17820737079&	2013-05-14 10:42:42	13	NULL		122899
12131355551097960974	121.10.248.46	http://item.taobao.com/item.htm?spm=2013.1.w1978670555.54.c2trZ4&id=16595276905&	2013-05-14 10:43:33	6	NULL		122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?q=%B6%CC%D0%E4%D1%A9%B7%C4%C1%AC%D2%C2%C8%B9%CF%C4&searcy_type=item&s_from...	2013-05-14 10:44:51	31	短袖雪纺连衣裙夏		122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?app=api&m=custom_bottom&id=10203735101%2C17815366911%2C14458882985%2C17909185983...	2013-05-14 10:45:22	12	短袖雪纺连衣裙夏		122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?q=%B6%CC%D0%E4%D1%A9%B7%C4%C1%AC%D2%C2%C8%B9%CF%C4&searcy_type=item&s_from...	2013-05-14 10:45:34	37	短袖雪纺连衣裙夏	40	122899
12131355551097960974	121.10.248.46	http://item.taobao.com/item.htm?spm=a230r.1.14.242.7zaohw&id=24417004372&_u=erelsdtb5a6	2013-05-14 10:46:11	42	NULL		122899
12131355551097960974	121.10.248.46	http://item.taobao.com/item.htm?spm=a230r.1.14.356.7zaohw&id=18126461135&_u=erelsdt4c68	2013-05-14 10:46:53	6	NULL		122899
12131355551097960974	121.10.248.46	http://rate.taobao.com/userTagInfo.htm?userNumId=1679034624&_ksTS=1368506190781_1733&callback=_ajaxUserInfo	2013-05-14 10:47:04	3	NULL		122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?q=%B6%CC%D0%E4%D1%A9%B7%C4%C1%AC%D2%C2%C8%B9%CF%C4&searcy_type=item&s_from...	2013-05-14 10:48:16	35	短袖雪纺连衣裙夏	80	122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?q=%B6%CC%D0%E4%D1%A9%B7%C4%C1%AC%D2%C2%C8%B9%CF%C4&searcy_type=item&s_from...	2013-05-14 10:48:51	33	短袖雪纺连衣裙夏	120	122899
12131355551097960974	121.10.248.46	http://s.click.taobao.com/t?e=zGU34CA7K%2BPkqB07S4%2FK0CFcRfH0EI%2BpZcoXo5axlZqc9H5ZEYWPS6XLnfhLc6ceR7sP...	2013-05-14 10:49:34	0	NULL		122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?q=%B6%CC%D0%E4%D1%A9%B7%C4%C1%AC%D2%C2%C8%B9%CF%C4&searcy_type=item&s_from...	2013-05-14 10:50:57	33	短袖雪纺连衣裙夏	160	122899
12131355551097960974	121.10.248.46	http://s.taobao.com/search?q=%B6%CC%D0%E4%D1%A9%B7%C4%C1%AC%D2%C2%C8%B9%CF%C4&searcy_type=item&s_from...	2013-05-14 10:51:30	292	短袖雪纺连衣裙夏	200	122899

Figure 1. Sample log records.

The log data contains click-through activities of both consumers and shop owners, but we are only interested in the search and browsing activities of consumers. Since shop owners tend to be a lot more active in making purchases than average consumers, we removed users who had too many sessions as belonging to businesses. Figure 2 shows the distribution of the users by the number of sessions.

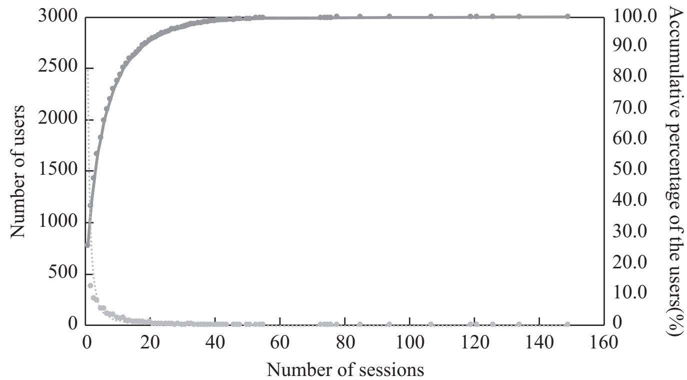

Figure 2. Distribution of users by number of sessions.

The *x*-axis in Figure 2 is the number of sessions, and the *y*-axis shows the number of users who had a particular number of sessions. The secondary axis of *y*-axis shows the accumulative percentage of the users. We remove users who belong to the upper 2.5% (with more than 33 sessions), those who were likely to be shop owners rather than regular consumers. The remaining log records include 2,910 users with 18,102 sessions and 47,387 queries. We use one fifth of this dataset for our experiment, which contains 582 users with 3,483 sessions and 8,949 queries. Table 1 shows sample records from these queries.

Table 1. Sample query records.

User ID	Sid	Query terms	Query terms (translation)
1028433974716967148	1973	丰胸仪	Breast augmentation instrument
1028433974716967148	1973	优格格丰乳仪	Yougege breast augmentation instrument
1028433974716967148	1975	北京茶月饼	Beijing tea mooncake
1028433974716967148	1975	金凤呈祥	Jinfengchengxiang
1028433974716967148	1975	金凤呈祥 200	Jinfengchengxiang 200
1028433974716967148	1976	美优食品	Meiyou food
1028433974716967148	1977	XQB38-83皮带	XQB38-83 belt
1028433974716967148	1978	味多美卡	Meiduomei gift card
1028433974716967148	1978	Laver丰胸精油	Laver breast augmentation oil
1028433974716967148	1978	AOC 拉莫圣日尔曼干红葡萄酒 750ml	AOC Saint Germain Rameau claret 750ml
1028433974716967148	1978	AOC银奖圣玛杰庄园干红葡萄酒 750ml	AOC silver award Domaine Saint Majan claret 750ml
1028433974716967148	1978	圣玛杰庄园干红葡萄酒 750ml	Domaine Saint Majan claret 750ml
1028433974716967148	1978	红绳	Red rope
1028433974716967148	1978	红绳批发	Red rope wholesale
1028433974716967148	1978	项链挂绳编织	Necklace rope woven

3.2 Task Identification

We use Rwordseg (Li, 2013) as the default dictionary and an additional dictionary containing terms from the Product Catalog acquired from Taobao API[①] for query term segmentation. Then we calculate the pairwise Jaccard index (Järvelin et al., 2007) of queries that belong to a same user, and construct a similarity matrix based on the Jaccard values. We employ the following four methods to identify whether queries belong to the same task:

1) Rule-based sequential comparison, where for each query q_i, we calculate its Jaccard similarity score s_{ij} with all previously labeled queries q_j, $j \in \{1,\ldots,i\text{-}1\}$; if s_{ij} is greater than a given threshold t, it is assigned a task label of T_j;
2) Clustering that uses the average Jaccard value as the Jaccard index between the new cluster and other clusters (clustering-avg); and
3) Clustering that uses the maximum Jaccard value as the Jaccard index between new cluster and other clusters (clustering-max).

Hierarchical clustering stops, however, when the Jaccard indexes between the two clusters are lower than a given threshold. For each method (with the two dictionaries of default and product catalog), we experiment with threshold values ranging from 0.2–0.6 and plot the F-score results as discussed in Session 3.3. Figure 3 shows the results.

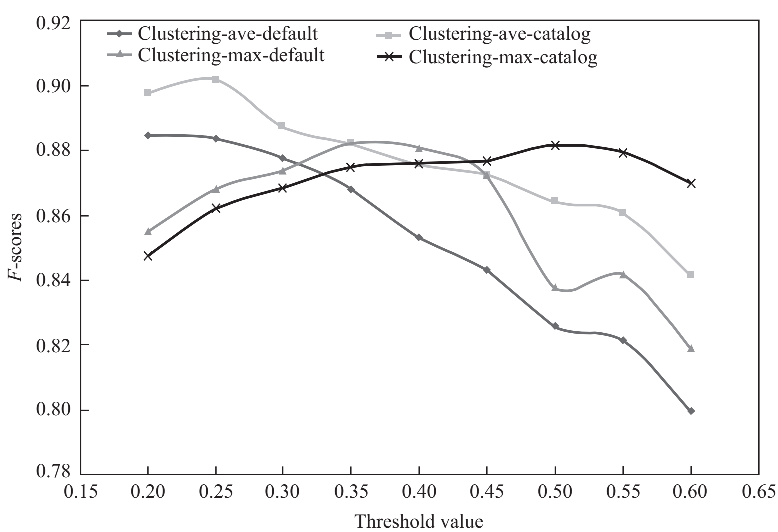

Figure 3. Thresholds and F-scores of the clustering approaches.

① http://open.taobao.com/

As Figure 3 shows, the performances of the clustering methods are most stable between thresholds 0.3 and 0.4. Therefore, we chose the following three thresholds for our later experiments: 0.3, 0.35, and 0.4.

3.3 Assessment

To identify which combination of dictionary, method, and threshold works best for task identification, we created a gold standard with 10% of the experiment data (1,015 search log records) chosen at random. Two human coders examined the query terms and identified product search tasks separately. The coders were instructed to assign a task number to each query in a sequence, where the same task numbers are assigned to queries that belong to the same task. Table 2 presents part of the human task identification results.

Table 2. Sample of human task identifications.

	Sid	Query Terms (original)	Query Terms (translation)	Coder #1	Coder #2
1	1973	丰胸仪	Breast augmentation instrument	1	1
2	1973	优格格丰乳仪	Yougege breast augmentation instrument	1	1
3	1975	北京茶月饼	Beijing tea mooncake	2	2
4	1975	金凤呈祥	Jinfengchengxiang	3	3
5	1975	金凤呈祥 200	Jinfengchengxiang 200	3	3
6	1976	美优食品	Meiyou food	4	4
7	1977	XQB38-83皮带	XQB38-83 belt	5	5
8	1978	味多美卡	Meiduomei gift card	6	3
9	1978	Laver丰胸精油	Laver breast augmentation oil	7	6
10	1978	AOC 拉莫圣日尔曼干红葡萄酒 750ml	AOC Saint Germain Rameau claret 750ml	8	7
11	1978	AOC银奖圣玛杰庄园干红葡萄酒 750ml	AOC silver award Domaine Saint Majan claret 750ml	8	7
12	1978	圣玛杰庄园干红葡萄酒 750ml	Domaine Saint Majan claret 750ml	8	7
13	1978	红绳	Red rope	9	8
14	1978	红绳批发	Red rope wholesale	9	8
15	1978	项链挂绳编织	Necklace rope woven	9	8

As noted in Table 2, the two coders agreed on most of the queries, but for record #8 (gift card), Coder 1 considered it as a separate task than task #3 (mooncake), whereas Coder 2 considered it as the same task as task #3, making the agreement level for these two human identification results 91.97%. For the records that the two coders did not initially agree on, we asked the two coders to discuss and resolve their different interpretations. We then used the agreed-on identification result as the gold standard to assess different task identification methods used in this paper.

For each identification approach, we calculated standard recall and precision. Recall (R) is the percentage of correctly identified records in all manually identified tasks, and precision (P) is the percentage of correctly identified records in all identified records. Then we calculated the F-measure ($F = \frac{2PR}{P+R}$) to assess each task identification approach.

4 Findings

4.1 Task Identification Results

We experimented with several combinations of task identification methods, dictionaries, and thresholds. The results are shown in Table 3.

Table 3. Task identification results.

Method	Dictionary	Threshold	*R*	*P*	*F*
Rule based	Default	0.3	0.8995	0.8542	0.8763
Rule based	Default	0.35	0.8670	0.8946	**0.8806**
Rule based	Default	0.4	0.8414	0.9232	0.8804
Rule based	Default + Pro-Catalog	0.3	0.8837	0.8552	0.8692
Rule based	Default + Pro-Catalog	0.35	0.8453	0.8926	0.8683
Rule based	Default + Pro-Catalog	0.4	0.8266	0.9054	0.8642
Clustering-avg	Default	0.3	0.8394	0.9192	0.8775
Clustering-avg	Default	0.35	0.8079	0.9379	0.8681
Clustering-avg	Default	0.4	0.7724	**0.9527**	0.8531
Clustering-avg	Default + Pro-Catalog	0.3	0.8246	0.9232	0.8711
Clustering-avg	Default + Pro-Catalog	0.35	0.7892	0.9379	0.8571
Clustering-avg	Default + Pro-Catalog	0.4	0.7557	**0.9537**	0.8432
Clustering-max	Default	0.3	**0.9123**	0.8384	0.8738
Clustering-max	Default	0.35	0.8867	0.8778	**0.8822***
Clustering-max	Default	0.4	0.8512	0.9123	**0.8807**
Clustering-max	Default + Pro-Catalog	0.3	**0.9015**	0.8433	0.8714
Clustering-max	Default + Pro-Catalog	0.35	0.8650	0.8788	0.8719
Clustering-max	Default + Pro-Catalog	0.4	0.8138	0.9143	0.8611

Note. *This approach yields the highest *F*-score and is used to perform task identification for the rest of the dataset.

Results show that the combination of the clustering method with the maximum similarity score, default dictionary, and threshold 0.35 yields the highest *F*-measure. So we used this combination with all 8,949 queries in the dataset and identified 6,189 shopping tasks associated with these queries.

Then we analyzed the task characteristics based on the task identification results. Basic characteristics of the sessions and tasks are shown in Table 4.

Table 4. Basic characteristics of sessions and tasks.

Item	Basic characteristics
Average number of queries per session	2.57
Highest number of queries in a session	21
Average number of tasks per session	1.78
Highest number of tasks in a session	41
Average number of queries per task	1.45
Highest number of queries in a task	15

On average, users issued 1.45 queries per task, with a maximum of 15 queries in one task. The average number of tasks is 1.78 per session, with a maximum of 41 tasks. The distribution of the sessions according to the number of task included in each session is shown in Table 5.

Table 5. Distribution of the sessions according to the number of task per session.

Number of task in a session	Freq.	Percent (%)	Cumulative percent (%)
1	2140	61.4	61.4
2	748	21.5	82.9
3	292	8.4	91.3
4	132	3.8	95.1
5	73	2.1	97.2
6	37	1.1	98.2
7	23	0.7	98.9
8	14	0.4	99.3
9	11	0.3	99.6
10	5	0.1	99.8
11 and more	8	0.2	100

Of the 3,483 sessions, 2,140 (61.4%) contain only one task, and 38.6% are multi-tasking sessions. There are 748 (21.5%) two-task sessions and 292 (8.4%) three-task sessions. Only 98 (2.8%) sessions contain more than five tasks.

4.2 Search Characteristics in Mono-tasking and Multi-tasking Sessions

4.2.1 Number of Queries

We compared the number of queries per session with mono-tasking and multi-tasking sessions. Table 6 shows the results.

Table 6. Average number of queries per session and per task.

Session type	Number of queries per session	Number of queries per task
One task	1.45	1.45
Two tasks	2.93	1.47
Three or more tasks	6.14	1.43

Table 6 shows that users issued more queries in multi-tasking sessions. Mono-tasking sessions contain 1.45 queries on average, whereas two-task sessions contain 2.93 sessions, and sessions dealing with three or more tasks contain 6.14 queries. The average number of queries issued per task is about the same, however, regardless of the number of tasks included in a session. An independent-sample *T*-test shows that there is no significant difference in the number of queries per task in mono-tasking and multi-tasking sessions. On average, users issue 1.45 queries per task.

4.2.2 Query Length

We analyzed the length of the queries (i.e. number of characters included in a query) in one-task sessions, two-task sessions, and three-or-more-task sessions. Table 7 shows the results.

Table 7. Average query length.

Session type	Average query length in characters
One task	7.56
Two tasks	7.28
Three or more tasks	7.32

The average length of queries in all session is 7.39 while the average length of queries in one-task sessions is higher and the average length of queries in two-task and three-or-more-tasks is slightly shorter. The mean length of queries used in each task is quite similar to each other regardless of the number of tasks included in a session. An independent-sample *T*-test analysis shows that there is no significant difference in the mean length of queries in mono-tasking and multi-tasking sessions. The length of user queries is similar in the mono-tasking and multi-tasking sessions.

4.2.3 Session Duration

We examined duration of the sessions and compared their durations by session type (one task, two tasks, and three or more tasks). The results are shown in Tables 8 and 9.

Table 8. Session duration.

Item	Session duration
Average session duration	49 minutes 3 seconds
Average task duration	27 minutes 36 seconds
Longest session	14 hours 56 minutes 22 seconds

Table 9. Average session duration.

Session type	Average session duration
One task	36 minutes 9 seconds
Two tasks	54 minutes 19 seconds
Three or more tasks	1 hour 22 minutes 22 seconds

The correlation analysis between the number of tasks and the session duration results in the correlation coefficient of 0.3458 ($p < 0.01$). The duration of a session is positively related to the number of tasks a user is dealing with in that session. The average duration of two-task sessions is 1.5 times the average duration of one-task sessions, and the average duration of three-or-more-task sessions is 2.3 times the average duration of one-task sessions. The average duration of a task in one-task sessions, two-task sessions, and three-or-more-task sessions is shown in Table 10.

Table 10. Average duration of tasks.

Session type	Average duration of tasks
One task	37 minutes 10 seconds
Two tasks	27 minutes 26 seconds
Three or more tasks	19 minutes 44 seconds

Table 10 shows that as the number of tasks in a session increases, users spend less time on each task on average. The average duration of tasks in mono-tasking sessions is 37 minutes 10 seconds, while the average duration of tasks in multi-tasking sessions is 22 minutes 35 seconds (including two-task sessions and sessions with more than three tasks). *T*-test results suggest that there is a significant difference in the average task durations between mono-tasking sessions and multi-tasking sessions (F-value = 794.32, $p < 0.01$). The average duration of a task in mono-tasking sessions is significantly longer than that in multi-tasking sessions.

4.3 Task Relationship in Multi-tasking Sessions

4.3.1 Two-task Sessions

We examined the relationships between tasks in multi-tasking sessions using exploratory qualitative analysis. For example, Table 11 shows an example two-task session with two tasks that are related. The first two queries belong to Task 1 and the third query belongs to Task 2. The user searched for men's shirts in Task 1 and men's shorts in Task 1. The user was likely to search for men's summer outfits (short-sleeves shirts and shorts), which resulted in two sub-tasks that are related.

Table 11. Session with related search tasks.

SID	Time	Query terms	Query terms (translation)
1985	2013/5/20 20:21:34	休闲衬衫 男 短袖	Casual shirt male short sleeve
1985	2013/5/20 20:22:10	休闲衬衫 男 绿	Casual shirt male green
1985	2013/5/20 20:23:43	短裤 男	Shorts male

Table 12 shows a sample two-task session with two unrelated search tasks. The user searched for a 16G memory card (first two queries) in Task 1, and a water cup (third query) in Task 2, a multi-tasking session with two seemingly unrelated items.

Table 12. Session with unrelated search tasks.

SID	Time	Query terms	Query terms (translation)
67804	2013-05-17 10:51:33	内存卡16g正品包邮	Memory card 16g free delivery
67804	2013-05-17 10:53:35	vip内存卡16g正品包邮	Vip memory card 16g free delivery
67804	2013-05-17 11:12:08	水杯	Water cup

4.3.2 Sessions with Three or More Tasks

Similar to two-task sessions, we observed both related and unrelated tasks in sessions with three or more related tasks. For example, Table 13 shows an example session with three different tasks that are related. Each task includes one query search for different types of shoes.

Table 13. Three related search tasks.

SID	Time	Query terms	Query terms (translation)
879	2013-05-04 12:06:25	增高鞋真皮休闲	Hidden heel shoes leather leisure
879	2013-05-04 12:20:25	夏季潮男洞洞鞋牛皮	Summer male leather crocs
879	2013-05-04 13:05:58	万斯低帮豹纹	Vance leopard print low-cut

While some tasks were closely related, perhaps with purchasing intentions of products that belong to the same category, there were sessions with seemingly unrelated tasks. For example, Table 14 shows a search session with search tasks for sea-lion oil, a mobile phone card, and a lip balm.

Table 14. Three unrelated search tasks.

SID	Time	Query terms	Query terms (translation)
13527	2013-05-29 18:27:16	海狮油	Sea lions oil
13527	2013-05-29 18:35:24	上海移动100元快充	Shanghai Mobile 100 yuan recharge
13527	2013-05-29 18:35:48	上海移动10元	Shanghai Mobile 10 yuan
13527	2013-05-29 18:35:58	上海移动100元	Shanghai Mobile 100 yuan
13527	2013-05-29 19:05:44	澄糖滋润护唇膏玫瑰粉红	Sugar moist lip balm rose pink

5 Conclusion and Discussion

Further analysis is needed to better identify the relationships among tasks in the same session and how users cope with or manage different types of multi-tasking sessions. Understanding users' search tasks is a complex challenge. Sometimes search tasks span multiple sessions while other users deal with multiple tasks in one session. After identifying and analyzing multi-tasking online product search sessions, study results show that 38.6% of all search sessions are multi-tasking sessions, where users deal with two or more tasks at the same time, 3.4 times more than Web search (11.4% reported by Spink, Ozmutlu, & Ozmutlu, 2002). This may be due to the differences in the nature of Web search, where queries generally involve concepts and more extensive data, and product search, where data generally describe the products.

Comparing mono-tasking sessions and multi-tasking sessions, we found that (1) users issued a similar number of queries (ranging from 1.43 to 1.47) with similar lengths per task (7.3 to 7.6 characters) in mono-tasking and multi-tasking sessions, and (2) users spent more time in sessions with more tasks, which is similar to Web search, but spent less time on average for each task when the number of tasks increases in a session. The length of search queries in multi-tasking sessions for Web search are longer than those in mono-tasking sessions, which is not the case in product search.

The relationships between sessions and tasks are complex due to the myriad types of online search technology and variation in consumer behavior and intentions. Research has found that people may be involved in off-topic tasks while working on one-topic tasks (Feild & Allan, 2013), where search is a changing process that combines keyword search, browsing, and serendipity or unintentional discovery (Jiang, He, & Allan, 2014), in addition to impulse purchasing triggered by advertisement banners and promotions that are common in product search activities.

One limitation of this study is that our methods only consider query terms, which may not completely reflect the complex nature of consumer shopping behaviors. In future research, the identification of search tasks may take clues from click-through logs, which yield data on sites and items visited, mouse movement sequences, and so on. The identification of search tasks may also yield better results if the items viewed can be taken into consideration. Other measurements that help to measure the semantic similarity of queries instead of term similarity could also be used in further study. As understanding consumer behavior is a key aspect of many business enterprises, and the Internet and social media have become increasingly powerful consumer tools, this study contributes to the literature on online shopping trends. Gaining insights on information search activities within the Internet buying processes is thus an essential step to enhance awareness of consumer behavior for industry and providing better product search and recommendation services to consumers.

Acknowledgements

This research is supported by the National Science Foundation of China (NSFC) Grant (No. 71373015).

Author Contributions

X. Zhou (zhouxiang@pku.edu.cn) carried out data cleaning and analysis, and wrote the first draft. P.Y. Zhang (pengyi@pku.edu.cn, corresponding author) was in charge of research design and revised the paper. J. Wang (junwang@pku.edu.cn) participated in research design and provided feedback throughout the research.

References

Bhatnagar, A., & Ghose, S. (2004). Online information search termination patterns across product categories and consumer demographics. Journal of Retailing, 80(3), 221–228.

China Internet Network Information Center. (2015). 2014 China Online Shopping Market Research Report (in Chinese). Retrieved from https://www.cnnic.net.cn/hlwfzyj/hlwxzbg/dzswbg/201509/P020150909354828731159.pdf

Feild, H., & Allan, J. (2013). Task-aware query recommendation. In Proceedings of the 36th international ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 83–92). Dublin, Ireland.

Fortune. (1998). Net profits: Making the Internet work for you and your business. Technology Buyer's Guide Supplement, Summer, 240–243.

Glance, N.S. (2000). Community search sssistant. Artificial intelligence for web search. Menlo Park, CA: Association for the Advancement of Artificial Intelligence Press, 91–96.

Järvelin, A., Järvelin, A., & Järvelin, K. (2007). S-grams: Defining generalized n-grams for information retrieval. Information Processing & Management, 43(4), 1005–1019.

Jiang, J., He, D., & Allan, J. (2014). Searching, browsing, and clicking in a search session: Changes in user behavior by task and over time. In Proceedings of the 37th international ACM SIGIR conference on Research & Development in Information Retrieval (pp. 607–616). Queensland, Australia.

Jones, R., & Klinkner, K.L. (2008). Beyond the session timeout: Automatic hierarchical segmentation of search topics in query logs. In Proceedings of the 17th ACM Conference on Information and Knowledge Management (pp. 699–708). Napa Valley, California, USA.

Li, J. (2013). Rwordseg: Chinese word segmentation. Retrieved from http://R-Forge.R-project.org/projects/rweibo/.

Lin, S.J., & Belkin, N. (2005). Validation of a model of information seeking over multiple search sessions. Journal of the American Society for Information Science and Technology 56(4), 393–415.

Lucchese, C., Orlando, S., Perego, R., Silvestri, F., & Tolomei, G. (2011). Identifying task-based sessions in search engine query logs. In Proceedings of the 4th ACM International Conference on Web Search and Data Mining (pp. 277–286). Hongkong, China.

Lucchese, C., Orlando, S., Perego, R., Silvestri, F., & Tolomei, G. (2013). Discovering tasks from search engine query logs. ACM Transactions on Information Systems, 31(3), 1–43.

Moorthy, S., & Talukdar, D. (1995). Consumer information search revisited: Theory and empirical analysis. Journal of Consumer Research, 23(4), 263–277.

Raghavan, V.V., & Sever, H. (1995). On the reuse of past optimal queries. In Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 344–350). Seattle, Washington, USA.

Rowley, J. (2000). Product search in e-shopping: A review and research propositions. Journal of Consumer Marketing, 17(1), 20–35.

Salton, G., & Mcgill, M.J. (1986). Introduction to modern information retrieval. New York, NY: McGraw-Hill.

Spink, A., Ozmutlu, H.C., & Ozmutlu, S. (2002). Multitasking information seeking and searching processes. Journal of the American Society for Information Science and Technology, 53(8), 639–652.

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

Wang, H., Song, Y., Chang, M.W., He, X., White, R.W., & Chu, W. (2013). Learning to extract cross-session search tasks. In Proceedings of the 22nd International Conference on World Wide Web (pp. 1353–1364). Rio de Janeiro, Brazil.

Ye, C., & Wilson, M.L. (2014). A user defined taxonomy of factors that divide online information retrieval sessions. In Proceedings of the 5th Information Interaction in Context Symposium (pp. 48–54). Regensburg, Germany.

Yuan, X. (2014). Modeling user behavior on e-commerce websites. Unpublished Master Thesis, Peking University.