Multi-task Recognition and Analysis in Product Retrieval (Postprint)

Authors: Zhang Pengyi, Zhou Xiang, Wang Jun

Date: 2017-10-11T00:00:00+00:00

Abstract

【Purpose】To identify shopping tasks within product search and analyze behavioral characteristics of multi-task sessions. 【Method】Shopping task identification was conducted based on search queries in product search, leveraging Taobao' s product classification system and a self-constructed product vocabulary. The dataset comprised 19,704 search sessions from 2,754 users. 【Results】Factors affecting the number of search queries utilized per shopping task include product facets, quantity, and description difficulty; in multi-task sessions featuring primary and secondary tasks, inter-task relationships exhibit greater cohesion. 【Limitations】The shopping task identification methodology requires further refinement, and restricting the research object to search queries alone cannot comprehensively capture user behavioral characteristics. 【Conclusion】This study enhances understanding of product search behavior in shopping contexts and provides a basis for designing improved product recommendation algorithms and predicting users' shopping processes and behaviors.

Full Text

Preamble

ChinaXiv Collaborative Journal, Issue 268, 2016, No. 3
Multi-task Identification and Analysis in Product Search
Zhang Pengyi, Zhou Xiang, Wang Jun
(Department of Information Management, Peking University, Beijing 100871, China)

Abstract

[**Objective**] This study aims to identify shopping tasks in product search and analyze the behavioral characteristics of multi-task sessions. [**Methods**] Using Taobao' s product classification system and a self-constructed product lexicon, we identified shopping tasks based on search queries from 19,704 search sessions by 2,754 users. [**Results**] Factors influencing the number of queries per shopping task include product facets, quantity, and description difficulty. In multi-task sessions with primary and secondary tasks, the relationships among tasks are more closely connected. [**Limitations**] The task identification method needs improvement, and focusing solely on search queries cannot fully capture user behavioral characteristics. [**Conclusions**] This research helps understand product search behavior in online shopping and provides a basis for designing better product recommendation algorithms and predicting user shopping processes and behaviors.

Keywords: Product search, Shopping task identification, Shopping task analysis, Multi-task session

Classification Number: G358

1. Introduction

In recent years, online shopping has become increasingly popular among users due to its convenience, efficiency, and freedom from spatial and temporal constraints. Leveraging the high interactivity, convenience, transparency, and personalization of the Internet, online shopping has enhanced service quality and enabled real-time and precision marketing through data mining and machine learning technologies [1]. According to the "2014 China Online Shopping Market Research Report" published by the China Internet Network Information Center (CNNIC) [2], online retail sales reached 2.79 trillion yuan in 2014, a year-over-year increase of 49.7%. Given this rapid growth, analyzing online shopping user behavior to identify characteristic patterns and optimize e-commerce systems has become a significant research concern.

A common phenomenon in online shopping is the simultaneous execution of multiple shopping tasks, yet few studies have addressed the identification of online shopping tasks and search characteristics in multi-task scenarios. Multitasking is also prevalent in general web search. For instance, Spink et al. [3] found that multi-task information retrieval is a common user behavior, with 11.4% of sessions in their random sample of 1,000 sessions involving multi-task searches. Spink et al. [4] also discovered that over 90% of sessions containing three or more queries were multi-task sessions. However, among studies analyzing online shopping user behavior, few have examined search queries to identify user tasks. This research aims to fill this gap by analyzing search queries used during online shopping to identify shopping tasks and investigate relationships and search characteristics among multiple shopping tasks.

Key definitions for this study are as follows: A session is an uninterrupted sequence of request-response interactions between a client and server, representing a user' s online shopping process [5]. In this study, if a user performs no actions for 45 consecutive minutes after completing an action, the session is considered ended, with the user's next action marking the start of a new session [6]. A shopping task refers to a collection of behaviors performed by a user to purchase a specific product. For research purposes, this study focuses on identifying shopping tasks within individual sessions. A multi-task session contains multiple shopping tasks.

2. Related Work

2.1 Task Identification in Search Systems

The primary approaches to search task identification involve comparing the relevance between queries. Glance [7] proposed comparing URLs returned by search engine results, while Raghavan et al. [8] suggested comparing retrieved documents. Based on the similarity of these metrics between two queries, researchers determine whether they belong to the same search task.

Methods for comparing query similarity fall into two main categories. The first compares character-level similarity by extracting term sets from queries, using metrics such as Järvelin et al.' s [9] Jaccard index (the ratio of intersection to union between two sets) and Jones et al.' s [10] Levenshtein distance. The second compares semantic relatedness using vector space models, leveraging external knowledge resources like Wiktionary and Wikipedia to compute similarity between query terms and semantic resources. For each query term t with W resources, where C represents the relevance between term t and the i-th resource, a vector $C(t) = (C , C , \cdots, Cw)$ is generated. Each query' s vector is the sum of vectors for its constituent terms, and the angle between two query vectors serves as the similarity metric.

Lucchese et al. [12] employed the following methodology for task identification: First, preprocess search logs by removing empty and meaningless records, eliminating stop words, using algorithms to reduce morphological variations, and excluding long sessions with excessive queries (likely machine-generated and irrelevant for user studies). Next, calculate character-level and semantic similarity between queries using the methods described above. Finally, combine these two evaluation metrics using either simple weighted averaging or a threshold-based approach: when term similarity exceeds a certain threshold, use it as the final metric; otherwise, use the maximum of term similarity and semantic similarity.

Overall, current mainstream task identification methods analyze character-level or semantic similarity between queries to establish evaluation metrics for identifying multiple queries belonging to the same task.

2.2 Multi-task Search Session Research

Spink et al. [13] identified two primary reasons for multi-task search sessions: users initially having multiple topics to search, or users starting with one topic and deriving additional topics during the search process.

Regarding multi-task session characteristics, Ozmutlu et al. [14] found that multi-task sessions involve longer queries and more time than single-task sessions, a conclusion validated in their subsequent study [15]. Lin et al. [16] discovered that multi-task sessions contain more queries than single-task sessions.

Lucchese et al. [12], Spink et al. [13], and Wang et al. [17] all employed log analysis to study multi-task sessions. Key findings are summarized in .

As shown in , web search sessions average approximately 2-5 queries per session, with multi-task sessions accounting for 70%-80% of all sessions. Current research on multi-task sessions primarily focuses on numerical characteristics (e.g., time, query length, average tasks per session, average queries per task). Further investigations into how task types affect these characteristics and analyses of relationships among tasks within sessions remain limited.

This study processes product search logs using a novel approach that annotates queries with product types to identify tasks. We analyze basic patterns in multi-task product search sessions, examine relationships between product categories and search behaviors, and investigate connections between task relationships and task primacy in multi-task sessions, filling a gap in current research.

3. Methodology

3.1 Data Description and Preprocessing

The log data were collected by a third-party market research firm from user clients accessing Taobao.com in May 2013. The raw log table contains 1,409,160 records across 81,759 sessions from 4,285 users.

Preprocessing steps included: (1) Filtering and removing merchant data. Since this study focuses on consumers, we excluded users with more than 100 sessions (likely merchants). (2) Removing records from other behavior types such as login, browsing, pagination, and filtering, as this study emphasizes search behavior.

The preprocessed dataset includes 53,091 records containing original search queries from 2,754 users (64.27% of the total user base) and 19,704 sessions (24.10% of total sessions).

3.2 Task Identification Method and Evaluation

The task identification process for product search records is illustrated in [Figure 1: see original paper].

First, we obtained all product category data through the Taobao API—over 14,000 categories organized across 4-5 hierarchical levels. Based on this classification system, we established an improved product category hierarchy through supplementation, modification, and deletion of categories.

Second, we segmented the 53,091 product search records, yielding 163,617 terms. After frequency analysis, we selected 2,376 terms appearing ten or more times.

We filtered these to identify terms representing product types and matched them with our product category hierarchy. We also incorporated product category names from the classification system into our lexicon. The combined result is a comprehensive lexicon containing most terms directly referencing product categories and frequently used brand names in searches.

For task identification, we matched search records against the lexicon and category mapping. When a record matched a lexicon term, we annotated the corresponding shopping task with that term' s product category. Among 53,091 original search records, 41,486 matched successfully (78.14%). If multiple search records within a user' s session were annotated with the same product subcategory, they were considered part of the same shopping task.

To evaluate accuracy, we randomly sampled 200 matched records for manual task identification. Of these, 164 were accurate or basically accurate, yielding an 82% accuracy rate.

4. Results and Analysis

4.1 Session and Task Distribution

The search logs contained 19,074 sessions. After task identification, 16,050 sessions (84.15%) contained identified shopping tasks, corresponding to 26,182 tasks. The average was 1.63 tasks per session and 1.585 queries per task. To compare average queries per task across major product categories, shows the five categories with highest and lowest ratios of queries to tasks.

Categories with the highest ratios include clothing, bags/leather goods, hardware/tools, and watches, while the lowest ratios include traditional tonics, personal care/healthcare/massage equipment, jewelry/diamonds/jade/gold/silver, and pharmaceuticals/medical devices. Analysis reveals two main factors affecting average queries per task:

(1) **Product facets and quantity**: Product types with multiple facets and large quantities require more queries to narrow search results. For example, user UID: 9242120128831377467, SID: 99456 first searched for "chiffon shirt," then added constraints ("chiffon shirt short sleeve") and further refined ("chiffon shirt short sleeve floral pattern") to reduce selection scope. In contrast, categories like traditional tonics and pharmaceuticals/medical devices contain highly specific products requiring fewer refinement conditions. User UID: 10071878660749185838, SID: 624 purchased "Yuyan Sheng Baoying Dan"(a pharmaceutical) with a single query "Yuyan Sheng+Baoying Dan," yielding a small product range without needing additional queries.

(2) **Description difficulty**: Product types described by infrequently used terms require users to continuously modify queries. In a hardware/tools task, user UID: 17461349388365160511, SID: 48488 used multiple variations ("sliding door lock," "mobile door lock," "sliding lock," "pocket door

lock," "door lock") to find the desired product.

4.2 Dual-Task Session Analysis

(1) Distinguishing Primary and Secondary Tasks
In sessions with two shopping tasks, tasks may have primary-secondary relationships. Users typically focus more on primary tasks, conducting more detailed searches and filtering, thus using more queries. We used query count differences to determine task primacy. When query counts differed significantly, tasks were considered to have primary-secondary distinction (more queries = primary task). When query counts were similar, no such distinction was made.

Specifically, we extracted dual-task sessions with their task query counts and applied these criteria: (a) query count difference of 0 or 1; (b) difference of 2 when neither task had 1 and 3 queries respectively. Sessions meeting either criterion were labeled as having no primary-secondary distinction; others were labeled as having primary and secondary tasks.

Among 3,133 dual-task sessions, 2,591 (82.70%) had no primary-secondary relationship, while 542 (17.30%) had distinguished primary and secondary tasks.

(2) Inter-task Relationships
We classified task relationships into three categories: **strongly related** tasks (e.g., "sofa" and "residential furniture" representing hierarchical categories), **weakly related** tasks (e.g., "trousers" and "jeans" sharing a common parent category "pants"), and **unrelated** tasks (e.g., "jeans" and "facial mask/powder").

Applying these criteria to 3,133 dual-task sessions yielded the results shown in .

Unrelated tasks dominated at 76.30%. Notably, sessions with primary tasks showed significant differences in strong and weak associations. A chi-square independence test using SAS software yielded $^2 = 7.3558$, $p = 0.0067$, rejecting the null hypothesis at $= 0.05$. This indicates that session type and task type are associated: sessions with primary-secondary task distinctions have more closely connected tasks than those without such distinctions.

4.3 Three-Task Session Analysis

(1) Identifying Primary Tasks
In three-task sessions, a task was considered primary if its query count exceeded the sum of the other two tasks' queries, or if it equaled the sum while both other tasks had more than one query. Among 1,230 three-task sessions, 291 (23.66%) contained a primary task.

(2) Inter-task Relationships
For three-task sessions, we analyzed relationships pairwise, categorizing them

as: all three tasks related; two tasks related with the third unrelated; or all three pairwise unrelated.

We randomly sampled 120 sessions (90 without primary tasks, 30 with primary tasks) for relationship classification, with results shown in .

Using SAS software, we calculated attribute association between session type and relationship type. Kendall's Tau-b statistic ranges from -1 to 1, where values near 1 indicate positive association (sessions without primary tasks have higher inter-task relatedness) and values near -1 indicate negative association. The 95% confidence interval for Kendall' s Tau-b was (-0.3404, -0.0056), entirely left of zero, indicating negative association. Thus, sessions without primary tasks have lower inter-task relatedness, while sessions with primary tasks have higher relatedness.

4.4 Multi-task Session Relationships

Combining findings from dual-task and three-task sessions, we conclude that sessions with primary tasks (i.e., tasks using more queries) exhibit closer relationships among constituent tasks than sessions without primary tasks.

Sessions without primary tasks correspond to users' pre-defined multiple shopping tasks of equal importance that may lack direct connections. For example, user UID: 10987349518420796011, SID: 6367 searched for both phone cases and dresses in one session, using approximately equal queries (five each) for both products.

In sessions with primary tasks, users may derive related secondary tasks during primary task searching, creating close connections between tasks. Alternatively, users may initially search for one product but realize their true need during the process, shifting to primary task searching. For instance, user UID: 15691965703667985508, SID: 37988 searched for slippers using three queries, then encountered "jelly shoes" in results and derived a new search for this product. In such cases, query counts differ between tasks, establishing primary-secondary distinctions with close product-type relationships.

5. Conclusions and Discussion

This study identified and analyzed tasks in product search, yielding two main findings:

(1) **Factors affecting queries per task**: First, product facets and quantity—products with multiple facets and large inventories require more queries to narrow results. Second, description difficulty—products requiring specialized terminology necessitate query modifications to obtain accurate results.

(2) **Multi-task session relationships**: Using query volume to define primary and secondary tasks, we found that sessions with primary-secondary

distinctions exhibit closer task relationships than those without such distinctions.

Limitations remain. The task identification method could be improved regarding product classification, lexicon construction, and matching rules. Relying solely on search queries limits comprehensive behavioral analysis, as queries only reflect search behavior without capturing browsing, comparison, or decision-making processes. Other behavioral features like click patterns and page transitions could inform task identification. Future research should incorporate additional data to study multi-task information behavior. Furthermore, examining mobile shopping behavior and comparing it with PC-based behavior represents a valuable research direction.

References

[1] Feng Wei. Empirical Study on Factors Influencing Trust of Consumers in Online Shopping [D]. Hangzhou: Zhejiang University, 2010.

[2] CNNIC. The Research Report of Online Shopping Market in China, 2014 [EB/OL]. [2015-06-30]. http://www.cnnic.cn/gywm/xwzx/rdxw/2015/201509/W020150908609566580083.pdf.

[3] Spink A, Ozmutlu H C, Ozmutlu S. Multitasking Information Seeking and Searching Processes [J]. Journal of the American Society for Information Science and Technology, 2002, 53(8): 639-652.

[4] Spink A, Park M, Jansen B J. Information Task Switching and Multitasking Web Search [J]. Proceedings of the American Society for Information Science and Technology, 2004, 41(1): 22-28.

[5] Zhang Wenjun, Wang Jun, Xu Shanchuan. The Probing of E-commerce User Need States by Page Cluster Analysis—An Empirical Study on Women' s Clothes from Taobao.com [J]. New Technology of Library and Information Service, 2015(3): 67-74.

[6] Yuan Xingfu, Zhang Pengyi, Liu Honglian, et al. Modeling E-commerce User Session Behaviors Based on Click-through Sequences [J]. Library and Information Service, 2015, 59(1): 119-126.

[7] Glance N S. Community Search Assistant [C]. In: Proceedings of the 6th International Conference on Intelligent User Interfaces. New York: ACM, 2001: 91-96.

[8] Raghavan V V, Sever H. On the Reuse of Past Optimal Queries [C]. In: Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 1995: 344-350.

[9] Järvelin A, Järvelin A, Järvelin K. S-grams: Defining Generalized N-grams for Information Retrieval [J]. Information Processing & Management, 2007, 43(4): 1005-1019.

[10] Jones R, Klinkner K L. Beyond the Session Timeout: Automatic Hierarchical Segmentation of Search Topics in Query Logs [C]. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management. New York: ACM, 2008: 699-708.

[11] Salton G, McGill M J. Introduction to Modern Information Retrieval [M]. McGraw-Hill Book Company, 1983.

[12] Lucchese C, Orlando S, Perego R, et al. Identifying Task-based Sessions in Search Engine Query Logs [C]. In: Proceedings of the 4th ACM International Conference on Web Search and Data Mining. New York: ACM, 2011: 277-286.

[13] Spink A, Park M, Jansen B J, et al. Multitasking During Web Search Sessions [J]. Information Processing & Management, 2006, 42(1): 264-275.

[14] Ozmutlu S, Ozmutlu H C, Spink A. Multitasking Web Searching and Implications for Design [J]. Proceedings of the American Society for Information Science and Technology, 2003, 40(1): 416-421.

[15] Ozmutlu S, Ozmutlu H C, Spink A. A Study of Multitasking Web Search [C]. In: Proceedings of International Conference on Information Technology: Coding and Computing (ITCC 2003). Washington DC: IEEE, 2003: 145-148.

[16] Lin S, Belkin N. Validation of a Model of Information Seeking over Multiple Search Sessions [J]. Journal of the American Society for Information Science and Technology, 2005, 56(4): 393-415.

[17] Wang H, Song Y, Chang M W, et al. Learning to Extract Cross-session Search Tasks [C]. In: Proceedings of the 22nd International Conference on World Wide Web. Geneva: International World Wide Web Conferences Steering Committee, 2013: 1353-1364.

Author Contributions: Zhang Pengyi: research design, paper writing and revision; Zhou Xiang: experiment implementation, data analysis, paper writing and revision; Wang Jun: research design, paper revision.

Conflict of Interest Statement: All authors declare no conflict of interest.

Supporting Data: Supporting data is self-archived by the authors, E-mail: zhouxiang.im@pku.edu.cn.
[1] Zhang Pengyi, Zhou Xiang, Wang Jun. query_by_product_category.csv. Query distribution by product category.
[2] Zhang Pengyi, Zhou Xiang, Wang Jun. task_analysis.csv. Dual-task and multi-task relationship analysis.
[3] Zhang Pengyi, Zhou Xiang, Wang Jun. first_class_product_category.csv. Numerical characteristics of major product categories.

Received: 2015-10-19
Revised: 2015-11-09

Multi-task Session Identification and Analysis in Product Search

Zhang Pengyi, Zhou Xiang, Wang Jun
(Department of Information Management, Peking University, Beijing 100871, China)

Abstract: [Objective] This research aims to identify shopping tasks from product search, and then analyze the characteristic of multi-task sessions. [Methods] Using the product classification of Taobao, and a list of manually selected product terms, we identified online shopping tasks based on query terms from 19,704 search sessions by 2,754 users. [Results] First, factors influence the number of queries per shopping task: product characteristics, the amount of available products, and the difficulty in describing product category with query terms. Second, we found that in sessions with a major task, the relationship among the shopping tasks is closer. [Limitations] The task identification method based on query terms cannot completely describe the complex consumer shopping behaviors. [Conclusions] This study provides an exploratory understanding of the relationships among various shopping tasks, and may be used to improve product recommendation algorithm, as well as predict shopping behaviors.

Keywords: Product search, Shopping task identification, Shopping task analysis, Multi-task session