Multi-session Product Information Search Behavior, Context, and Influencing Factors: A Postprint Study

Authors: Liu Honglian, Zhang Pengyi, Wang Jun

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

Abstract

【Objective】 To investigate the behavioral characteristics of product information search, session time interval features, and shopping website usage features of e-commerce users in multi-session online shopping processes, and to explore the context, reasons, and motivations behind these behaviors. 【Method】 Based on 1,409,160 access logs from 4,285 users of an e-commerce website, sequential analysis and clustering analysis methods were employed to mine user information search behavior characteristics, combined with interviews to study influencing factors such as context and reasons behind the behaviors. 【Results】 Multi-session online shoppers do not have urgent information needs and prefer viewing product details over searching; the average session time interval is 3-4 days; motivations for continuing original shopping tasks include factors such as personal preferences, need status, payment ability, and time; users primarily return to original shopping tasks through search, shopping cart, favorites, browsing same-store or same-item pages, and personalized product recommendations. 【Limitations】 Interview results are limited by sample size and lack universality. 【Conclusion】 This study helps understand users' complex online shopping behaviors and guides shopping websites to improve service quality and user experience.

Full Text

Abstract

Objective: This study examines the behavioral characteristics of product information seeking, session interval patterns, and e-commerce platform usage features among online shoppers engaged in multi-session shopping processes, while investigating the contextual factors, motivations, and driving forces underlying these behaviors.

Methods: Based on 1,409,160 access logs from 4,285 users of an e-commerce website, we employed sequence analysis and clustering methods to uncover patterns in users' information seeking behaviors. Complementing this quantitative analysis, we conducted interviews to explore the contextual factors and motivations influencing these behaviors.

Results: Multi-session shoppers exhibited non-urgent information needs and demonstrated a stronger preference for viewing product details over search-based exploration. The average interval between sessions was 3-4 days. Drivers for resuming previous shopping tasks included personal preferences, need states, payment capacity, and time availability. Users primarily returned to prior shopping tasks through search, shopping carts, bookmarks, browsing within the same store or similar products, and personalized recommendations.

Limitations: Interview findings are constrained by sample size and may not be universally generalizable.

Conclusions: This research contributes to understanding complex online shopping behaviors and provides guidance for e-commerce platforms to improve service quality and user experience.

Keywords: Multi-session online shopping; Information seeking behavior; E-commerce website

Classification Number: G354

Introduction

As of December 2014, China' s online shopping user base reached 361 million, an increase of 59.53 million from the end of 2013, representing a 19.7% growth rate. The proportion of internet users engaging in online shopping rose from 48.9% to 55.7%, with the market size reaching 2.8 trillion yuan [**?**]. The convenience and diversity of choices in online shopping have transformed consumer decision-making processes. A notable characteristic of online shopping is that users may visit e-commerce sites multiple times within a single shopping task, engaging in repeated information seeking and product comparison before finalizing purchase decisions [**?**].

Existing research has modeled online shoppers' information behaviors at the session level using clickstream data to analyze behavioral patterns [**?**], yet few studies have focused on user behavior characteristics across multiple sessions within a single shopping task [**?**, **?**]. Consequently, our understanding remains limited regarding the complex information seeking behaviors of online shoppers who make purchase decisions through repeated visits. This study aims to investigate the behavioral characteristics of product information seeking, session interval patterns, and the contextual factors, motivations, and influencing factors during users' repeated visits to shopping websites. Specifically, we examine: (1) the starting behaviors, ending behaviors, typical behaviors, behavior transition patterns, and session interval characteristics during multi-session product

information seeking; (2) the pathways through which users return to previous shopping tasks; (3) the contexts and need states in which users require multiple visits to make purchase decisions; and (4) the reasons for shopping interruption and the methods for resuming tasks, which collectively enhance our understanding of user behaviors in multi-session online shopping.

Key definitions relevant to this study are as follows: A session represents an uninterrupted sequence of request-response interactions between a user and server, with each session constituting one online shopping process [**?**]. In this study, if a user performs no actions for 45 consecutive minutes [**?**], the current session is considered ended, with the user' s next action marking the beginning of a new session. A shopping task refers to the collection of activities including product information seeking and purchase decision-making [**?**]. Shopping tasks may conclude after a single session ("single-session" shopping tasks) or after multiple sessions ("multi-session" shopping tasks).

Literature Review

Consumer decision-making processes consist of several stages: problem recognition, information seeking, evaluation and selection, purchase, and post-purchase feedback. The classic Engel-Kollat-Blackwell model in consumer behavior research comprehensively describes this process: when consumers' actual state fails to match their desired state, need recognition occurs, prompting information seeking to satisfy the need. Consumers then evaluate and compare alternatives based on internal and external information, develop preferences, make purchase decisions, and post-purchase evaluations subsequently influence future decisions [**?**]. Unlike traditional shopping, online consumers face a virtual, interactive environment with more convenient information access and broader selection ranges, making comprehensive evaluation of all alternatives difficult. The core online purchase process comprises three stages: information search, evaluation and selection, and purchase [**?**]. Some research divides online consumer decision-making into two main phases: broad browsing of numerous products to identify preferred options, followed by in-depth comparison of selected items to make final purchase decisions [**?**]. This study draws upon the macro-level process from consumer behavior models to examine users' information query behaviors throughout the shopping process, though whether online shopping exhibits similar characteristics to traditional retail requires further verification. Unlike traditional shopping, online information query costs are substantially reduced, visit frequency increases, and the number of compared products grows, yet research on consumers' multi-session information queries remains scarce.

In the information behavior domain, numerous studies have analyzed and modeled users completing searches across multiple stages for literature or web search tasks. For instance, Lin et al.' s Multiple Information Seeking Episodes (MISE) model [**?**] proposes eight patterns to systematically explain how information seekers resolve the same problem through repeated searches. Vakkari' s task-based information retrieval process theory models the information retrieval pro-

cess in academic writing, identifying key factors in continuous searching [**?**]. Spink' s multiple search sessions model incorporates both user and system factors affecting search outcomes in continuous retrieval processes [**?**]. Komlodi et al.' s information seeking history framework describes information retrieval history as choices in human-system interaction, outlining external and internal factors influencing the use of retrieval history [**?**]. Information behavior research has primarily focused on library and information science domains, with research objects mainly targeting literature or web retrieval. These studies provide methodological and conceptual references for understanding multi-session product information seeking behaviors. However, product information queries differ from general information queries as the obtained information serves consumption decisions. Therefore, this study integrates insights from both consumer behavior research and information behavior research.

Methodology

3.1 Log Analysis

Our raw dataset comprised browser access logs from May 2013, collected by an internet market research firm through browser plugins. Log fields are detailed in Table 1 . Supplementary data included user registration information and product category data (approximately 14,000 categories obtained through Taobao' s open API).

Previous research defined and labeled user sessions based on raw logs and annotated page types and user behavior types [**?**], yielding 1,409,160 usable log records across 81,759 sessions from 4,285 users. Building upon prior work identifying multi-session online shopping tasks [**?**], this study investigates information seeking behavior characteristics during multi-session shopping processes.

We employed statistical methods to analyze session interval patterns in multi-session shopping. Using Microsoft SQL Server Data Tools' sequence analysis and clustering algorithms, we mined users' browsing and search behavior sequences to analyze behavior transition patterns and starting/ending behaviors across sessions. The SQL Server sequence clustering algorithm is a hybrid approach combining clustering methods and Markov chain analysis for sequence data. Sequence data represents transitions between states—in this study, users' sequences of product searches or page clicks. The algorithm examines all transition probabilities, measures distances between all possible sequences in the dataset to determine cluster inputs, and then uses this sequence information as input for the Expectation Maximization (EM) clustering method. In EM clustering, the algorithm iteratively refines the initial classification model, calculating the probability of data belonging to specific clusters, and terminates when the probabilistic model fits the data appropriately.

3.2 User Interviews

Interviews were conducted via telephone from April to May 2015, with each session lasting approximately 20-30 minutes and being audio-recorded. Users were coded chronologically by interview order, and recordings were transcribed and content-analyzed. The interview protocol comprised 13 questions across five dimensions: shopping process, shopping context and motivations, reasons and timing for shopping interruption, motivations for returning to previous tasks, and methods/pathways for resumption.

Participant demographics are shown in Table 2 . Based on the demographic distribution of multi-session shoppers from log data and convenience sampling principles, we selected 15 interviewees, with age and occupational distributions presented in Table 2.

Results

4.1 Information Seeking Behavior Sequence

Using chi-square tests, we found significant differences between multi-session and single-session shoppers in gender distribution ($\chi^2 = 54.949, \alpha = 0.000$) and occupational distribution ($\chi^2 = 41.558, \alpha = 0.000$), but no significant differences in age ($\chi^2 = 8.167, \alpha = 0.698$), education level ($\chi^2 = 3.977, \alpha = 0.782$), or income level ($\chi^2 = 26.789, \alpha = 0.110$). Figure 1 [Figure 1: see original paper] shows the gender and occupational distribution of multi-session shoppers. Occupations were self-reported categories at registration, with "professionals" referring to those with specific technical skills (e.g., programmers, doctors).

Sequence and cluster analysis of search (query-based retrieval) and browse (product detail viewing) behaviors in multi-session shopping tasks revealed distinct behavioral patterns, as shown in Figure 2 [Figure 2: see original paper]. Compared with general online shopping sessions, multi-session shoppers showed greater tendency to "view product details," with higher probabilities of both starting and ending sessions with this behavior, and lower probabilities of "inputting search queries." In terms of behavior transitions, when the previous action was "inputting search queries," multi-session tasks showed higher probability of consecutive searches ("search→search") and lower probability of transitioning to product details ("search→view details") compared to general sessions. When the previous action was "viewing product details," multi-session tasks exhibited higher probability of consecutive detail viewing ("view→view") and lower probability of returning to search ("view→search"). These findings indicate that during multi-session shopping tasks, users tend to repeatedly input search queries, yet have lower probability of viewing product details after searching—suggesting reduced likelihood of discovering potential products. Additionally, the lower probability of returning to search after viewing details suggests that when users find viewed products unsatisfactory, they are less likely to continue searching.

Chi-square tests analyzed whether significant differences existed in frequent behaviors (most frequent behavior types), starting behaviors, ending behaviors, and behavior transition patterns between general and multi-session shopping sessions. Results in Table 3 show significant differences across all dimensions.

Interviews revealed that these behavioral characteristics are closely related to users' shopping contexts, specifically their need states during multi-session shopping. Three primary need states emerged: (1) not urgently needed; (2) needed but not urgent; and (3) needed but uncertain about product specifications, quality, functionality, or usage methods. Ten interviewees mentioned needing but being uncertain about product size, style, quality, functionality, or usage methods; seven mentioned needing but not urgently; six mentioned not particularly needing the product initially. When the need state was "not particularly needed," users were more likely to begin information seeking by browsing, resulting in higher probability of "viewing product details" as the starting behavior. When users "needed but were uncertain about product specifications, quality, functionality, or usage methods," they engaged in more frequent product detail viewing for comparison, leading to higher probability of ending sessions with this behavior.

4.2 Session Time Interval

Session time interval in multi-session shopping tasks refers to the time between adjacent sessions within the same shopping task, calculated as the start time of the later session minus the start time of the earlier session. Basic statistics revealed a maximum interval of 2,595,871 seconds (approximately 30.05 days), a mean of 343,109.55 seconds (approximately 3.97 days), with median and mode closely aligned at 2.00 days and 1.89 days, respectively. The distribution of session intervals is shown in Figure 3 [Figure 3: see original paper]. Linear regression yielded an R-squared of 0.6833, indicating that session intervals within the same shopping task approximately follow a power-law distribution: 极少数的购物任务中会话时间间隔很长, 而大多数购物任务中的会话时间间隔较短.

Interview results showed that reasons for interrupting shopping in multi-session contexts are closely related to shopping contexts, including "needing but uncertain about product specifications, style, quality, functionality, or usage methods," "needing but not urgent," or "not particularly needed." Additional interruption reasons included time constraints, being distracted by other matters, insufficient fragmented time for decision-making, and temporary inability to pay. Motivations for returning to previous shopping tasks included: genuine liking, actual need, intensified need, affordable payment capacity, and accidental re-browsing. Fourteen participants identified "actual need or more urgent demand" as motivation, representing need-driven returns; six mentioned "genuinely liking the item (though not necessarily needing it)" ; three described "casual browsing and accidentally returning" ; one mentioned returning when "payment capacity becomes adequate."

Seven interviewees reported intervals of approximately 2-3 days between adjacent sessions, with three even returning the same day due to "clear shopping needs," stating they "wouldn' t wait long before revisiting the shopping site." Six participants reported intervals ranging from several days to one or two weeks, influenced by need states, decision stages, and time schedules. Two participants reported longer intervals of "one month," "two to three months," or even "half a year," because they "didn' t actually need" the items or used shopping sites infrequently. These findings align with log analysis results showing that 极少数的购物任务中会话时间间隔很长, while most tasks have shorter intervals, with a mode of approximately 2 days.

4.3 Return Pathways to Original Shopping Tasks

Interview results revealed that users primarily returned to previous shopping tasks through: searching again, shopping carts, bookmarks, browsing within the same store or similar products, and personalized recommendations. Some users relied on search pathways during early shopping stages, particularly when they had not yet narrowed down options and were in broad comparison phases. Others habitually used search regardless of shopping stage, or depended more on search to locate previously viewed products under specific contexts—such as when they had particular brand or store preferences, or when seeking a specific product.

Users returned via shopping carts in two scenarios: (1) after broad browsing and comparison, they added selected items to carts for subsequent focused comparison; (2) they added liked or potentially needed items to carts for future reference, essentially using carts as bookmarks, revisiting added items during subsequent visits or when actual need arose. Similarly, users returned via bookmarks in two ways: (1) adding selected items after broad comparison for subsequent focused evaluation; (2) bookmarking liked or potentially needed items for future reference. Analysis of return pathways indicates that for some users, bookmarks and shopping carts serve similar purposes, particularly during broad selection stages. Whether users add narrowed-down items to bookmarks or carts largely depends on personal usage habits.

Conclusion and Discussion

This study employed a mixed-methods approach combining quantitative log analysis and qualitative user interviews to explore product information seeking behavior sequences, session interval patterns, and e-commerce platform usage characteristics in multi-session online shopping, while investigating influencing factors including shopping contexts, interruption reasons, and motivations for returning to previous tasks.

Findings reveal that users in multi-session shopping contexts show higher probabilities of starting and ending sessions with "viewing product details," with this behavior being the most frequent, depending on their shopping context. In

terms of behavior transitions, users tend to repeatedly "input search queries" to seek information, yet exhibit lower probability of viewing product details after searching—indicating reduced likelihood of discovering potential products. Reasons for interrupting shopping and motivations for returning are closely related to users' need states, which also influence session intervals during multi-session shopping. Session intervals are further affected by users' frequency of shopping site usage. Primary pathways for returning to previous tasks include search, shopping carts, bookmarks, browsing within same store or similar products, and personalized recommendations, which relate to users' decision-making stages and shopping site usage habits.

This research contributes to deeper understanding of complex online shopping behaviors, informing navigation design and recommendation strategies for e-commerce platforms. Limitations include: due to privacy constraints in log data, interviewed users were not the same individuals who generated the logs, preventing investigation of the original users' shopping contexts, interruption reasons, motivations, and pathways. We could only select interviewees approximating the demographic distribution of log users. Additionally, interview method limitations mean results are representative within a certain range but may not generalize to other demographic groups. Future research should combine log analysis and interviews with the same user cohort to more deeply excavate behavioral characteristics in multi-session online shopping.

References

[1] China Internet Network Information Center. The 2014 Report of China's Online Shopping Market Research [EB/OL]. https://www.cnnic.net.cn/gywm/xwzx/rdxw/2015/201509/W020150

[2] Senecal S, Kalczynski P J, Nantel J. Consumers' Decision-Making Process and Their Online Shopping Behavior: A Clickstream Analysis[J]. Journal of Business Research, 2005, 58(11): 1599-1608.

[3] Moe W W. Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream[J]. Journal of Consumer Psychology, 2003, 13(1): 29-39.

[4] 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.

[5] Lee J, Podlaseck M, Schonberg E, et al. Visualization and Analysis of Clickstream Data of Online Stores for Understanding Web Merchandising[A]. // Applications of Data Mining to Electronic Commerce[M]. Springer, 2001: 59-84.

[6] Wang Lei. Information Seeking Behavior Research Based on Information Need of Consumers[J]. Information Studies: Theory & Application, 2013, 36(7): 90-93.

[7] 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.

[8] Gillenson M L, Sherrell D L. Enticing Online Consumers: An Extended Technology Acceptance Perspective[J]. Information & Management, 2002, 39(8): 705-719.

[9] Engel J F, Blackwell R D, Miniard P W. Consumer Behavior[J]. Journal of Consumer Policy, 1986, 9(4): 481.

[10] Teo T S, Yeong Y D. Assessing the Consumer Decision Process in the Digital Marketplace[J]. Omega, 2003, 31(5): 349-363.

[11] Häubl G, Trifts V. Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids[J]. Marketing Science, 2000, 19(1): 4-21.

[12] Lin S J, Belkin N J. Modeling Multiple Information Seeking Episodes[C]. In: Proceedings of the 63rd Annual Meeting of ASIS, 2000.

[13] Vakkari P. A Theory of the Task-Based Information Retrieval Process: A Summary and Generalisation of a Longitudinal Study[J]. Journal of Documentation, 2001, 57(1): 44-60.

[14] Spink A. Multiple Search Sessions Model of End-User Behavior: An Exploratory Study[J]. Journal of the American Society for Information Science, 1996, 47(8): 603-609.

[15] Komlodi A, Soergel D. Attorneys Interacting with Legal Information Systems: Tools for Mental Model Building and Task Integration[J]. Proceedings of the American Society for Information Science and Technology, 2002, 39(1): 152-163.

[16] Liu Honglian, Zhang Pengyi, Wang Jun. Product Information Seeking Behavior of Multi-session Online Shopping Tasks[J]. Library and Information Service, 2015, 59(14): 117-125.

Conflict of Interest Statement

All authors declare no conflict of interest.

Supporting Data

Supporting data is self-archived by authors, E-mail: pengyi@pku.edu.cn.

[1] Liu Honglian, Zhang Pengyi, Wang Jun. Interview Outline.doc. Interview outline and question list.

[2] Liu Honglian, Zhang Pengyi, Wang Jun. Interview Users.xls. Demographic information of interview participants.

[3] Liu Honglian, Zhang Pengyi, Wang Jun. Clustering Results.csv. User clustering analysis results.

Author Contributions

Liu Honglian: Data analysis, conducted interviews, drafted manuscript; Zhang Pengyi: Designed research framework, revised manuscript; Wang Jun: Proposed research idea, revised final manuscript.

Received: 2015-11-12 Revised: 2016-01-07

Note: Figure translations are in progress. See original paper for figures.