Modeling User Behavior for Vertical Search: Images, Apps and Products

Xiaohui Xie
BNRist, DCST, Tsinghua University
Beijing, China
xiexh_thu@163.com

Jiaxin Mao
BNRist, DCST, Tsinghua University
Beijing, China
maojiaxin@gmail.com

Yiqun Liu
BNRist, DCST, Tsinghua University
Beijing, China
yiqunliu@tsinghua.edu.cn

Maarten de Rijke
University of Amsterdam & Ahold Delhaize
Amsterdam, The Netherlands
m.derijke@uva.nl

ABSTRACT

Search applications such as image search, app search and product search are crucial parts of web search, which we denote as vertical search services. This tutorial will introduce the research and applications of user behavior modeling for vertical search. The bulk of the tutorial is devoted to covering research into behavior patterns, user behavior models and applications of user behavior data to refine evaluation metrics and ranking models for web-based vertical search.

CCS CONCEPTS

• **Information systems → Evaluation of retrieval results**.

KEYWORDS

Image search; Preference judgment; Evaluation metric; User behavior

ACM Reference Format:
Xiaohui Xie, Jiaxin Mao, Yiqun Liu, and Maarten de Rijke. 2020. Modeling User Behavior for Vertical Search: Images, Apps and Products. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3397271.3401423

1 MOTIVATION

Vertical search, e.g., image, app and product search, has helped search engine users to efficiently obtain information in various forms [8, 18, 30]. User behavior data has been successfully adopted to improve general Web searches in result ranking, query suggestion, query auto completion, etc. We therefore believe that understanding user interaction behavior in vertical search scenarios will also provide valuable insight into the optimization of their performances. Vertical search engines differ from general web search engines in the following aspects:

(1) *Different search intent*: In vertical search, the results user searches for are vertical items in heterogeneous forms, e.g., images, rather than web pages. It has been shown that Broder's taxonomy [3], where the search intent of general web search users is categorized into three classes (informational, transactional and navigational), is not applicable for vertical search [9]. Since intent is assumed to be the immediate antecedent of behavior, for any information access service, it is important to understand the underlying intent behind user behavior.
(2) *Different kinds of relevance definitions*: Compared with document results of general web search, vertical results are more heterogeneous. Relevance is a multi-dimensional notion in vertical search in the sense that users will take different aspects into consideration while making relevance judgments. For example, besides topical relevance, other product-related attributes including price and popularity also have a strong effect on user behavior in product search [4].
(3) *Different type of result placement*: Vertical search engines usually place results in a grid-based panel rather than in a one-dimensional ranked list as in web search engines. In that regard, users can not only browse results in a vertical direction but also in a horizontal direction. Differences in result placement lead to different examination behavior of search users, which brings about different distributions of users' attention [24].
(4) *Different interaction mechanisms*: Interaction mechanisms in vertical search are unique. For instance, instead of a query-dependent summary of the landing page, an image snapshot is shown together with some metadata about the image, which is typically only available when a cursor hovers on the result in image search. A similar mechanism is often used in product search, where products with different attributes, e.g., colors, can be previewed while cursor hovering on thumbnails in the result box. Different interaction mechanisms can also lead to different user behavior patterns.

The aforementioned differences bring challenges to modeling user behavior in vertical search scenarios. It is challenging, or even problematic, to apply user behavior models and corresponding ranking models and evaluation metrics that have been proven useful for general Web search to vertical search without adaption.

SIGIR '20, July 25–30, 2020, Virtual Event, China

ACM ISBN 978-1-4503-8016-4/20/07.
https://doi.org/10.1145/3397271.3401423

To address these challenges, there is a growing body of work on modeling user behavior in vertical search scenarios. Besides log-based user behavior, more fine-grained user feedback and behavior data of vertical search has been collected and mined on the basis of user surveys, field studies and user studies [2, 19, 24]. Equipped with this data, taxonomies of search intents of vertical search users have been proposed [17, 19, 23, 30]. For example, Xie et al. [23] propose an image search intent taxonomy consisting of three classes: "Explore/Learn", "Entertain", "Locate/Acquire" and Su et al. [19] categorize search intents of product search users into Target Finding (TF), Decision Making (DM) and Exploration (EP). Moreover, result placement and interaction mechanisms of vertical search have been studied [20, 27].

Based on these observations, previous work investigates how user behavior varies with different search intents and how result placement and interaction mechanisms affect user behavior [19, 20, 23, 25, 30]. Among them, Wu et al. [20] jointly model the different stages of the shopping journey, i.e., comparing and clicking a product on a search result page and deciding whether to purchase a product on the product description page, in product search. Zhuo et al. [30] show that for queries with "Fuzzy Search" intent of app search, some user habits known from web search are brought to app search on mobile scenarios. Understanding user behavior provides an opportunity for using user behavior data to improve ranking models and evaluation metrics in vertical search. For ranking models, user behavior data enables personalized preference and context information to be considered into the model construction [4, 6, 16, 25, 26]. For instance, Xie et al. [25] utilize user behavior, i.e., click and cursor hovering, in a search session to capture users' short-term visual preference in image search. For evaluation metrics, different types of result placement and unique user behavior have been considered [20, 27, 31]. Besides topical relevance, aspects such as image quality (image search), revenue (product search) and download (app search) need to be optimized. Also, existing list-based evaluation metrics may not be applicable in grid-based result panels of vertical search.

We believe it is the right time to organize and present this material to a broad audience of interested information retrieval researchers, whether junior or senior, whether academic or industrial.

2 OBJECTIVES

User behavior modeling for vertical search has been investigated by different communities, in information retrieval, machine learning and cognitive science. A key aim of this tutorial is to bring these together and offer a unified perspective. We will do so by guiding participants through the user behavior modeling workflow described by Xie [22]; see Figure 1.

More specifically, our objectives for this tutorial are organized around six challenges that we will use as organizing principles:

Challenge 1 How do search intents, result placement and interaction mechanisms of vertical search differ from those of general web search?
- Motivate the study of user behavior in vertical search.

Challenge 2 How to collect and study user behavior data in vertical search scenarios such as image, app and product search scenarios?

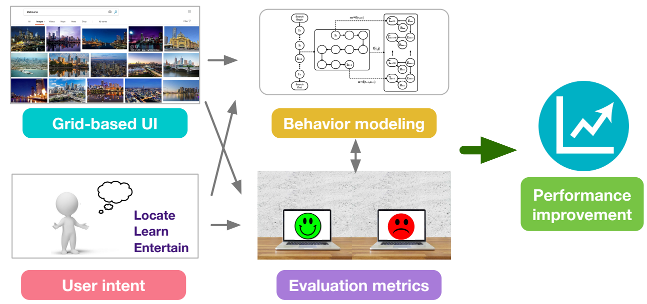

Figure 1: A flowchart of research topics introduced in our tutorial. An arrow (A → B) means that A might have an impact on B. (Figure taken from [22])

- Introduce how to collect and mine user behavior data through lab-based user study, field study and search log given a search scenario.

Challenge 3 How do different search intents, result placement and interaction mechanisms affect user behavior in vertical search scenarios?
- Show how user behavior (e.g., perception patterns and examination behavior) varies with these aspects. Provide both analytical results and insights gained from observations.
- Describe user behavior assumptions behind popular user models while these assumptions are proposed based on previous user behavior observations.

Challenge 4 How can we use behavior data to construct better ranking models to improve the performance of vertical search engines?
- Explain how user behavior can be used to improve the relevance ranking of vertical search engines and guide the construction of evaluation metrics which can better reflect users' satisfaction.
- Review work on relevance ranking that utilizes users' feedback to rank more relevant results to higher ranks.

Challenge 5 How can we use behavior data to design evaluation metrics to better reflect user satisfaction in the context of vertical search?
- Provide an understanding of the diversity and richness of user feedback in vertical search, which can be implicit (e.g, click signal, cursor hovering and mouse movement) or explicit (e.g, interactive re-ranking methods).
- Introduce an evaluation framework for vertical search; since search results in vertical search are more heterogeneous than in web search, we will first introduce how to perform relevance judgments for vertical results.
- Illustrate evaluation metrics and the underlying user behavior models.

Challenge 6 What are future directions of research in user behavior modeling for vertical search?
- Discuss interesting future directions for user behavior modeling for vertical search.

By discussing these research challenges, we aim to provide a clear picture of the state of the art in user behavior modeling for vertical search where multimedia content is provided. Although we mainly focus on behavior-based models, how to combine user behavior and content information, e.g., query and result descriptions, will also be discussed. We believe it is beneficial to bridge the gap between user behavior and search content. This tutorial will inspire researches both in behavior modeling and content modeling.

3 FORMAT AND DETAILED SCHEDULE

The tutorial will be organized in a half-day (3 hours plus break). Collection and analysis of user behavior data as well as model construction and experimental outcomes will be introduced.

[*30 minutes*] Background and objectives.

- Introduction to the context, i.e. vertical search scenarios, where user behavior modeling is investigated.
- Basic concepts and methodologies in search user behavior analysis

[*30 minutes*] User behavior data collection.

- How to collect and mine user behavior data in vertical search settings (e.g., field study [21], eye-tracking user study [24], user survey [2] and log data [11])?
- How to analyze collected behavior data and gain valuable insight from it [12, 29]?

[*30 minutes*] Analytical results of user behavior data.

- Search goals of vertical search users including both the "what" dimension and the "why" dimension [7, 17, 19, 23, 30].
- How search intents and interaction mechanisms affect user behavior in vertical search [20, 24]?

[*40 minutes*] User-centric ranking models.

- Introduce previous work on modeling user behavior to construct ranking models for vertical search, including models and corresponding user behavior assumptions [1, 5, 13, 15, 25, 26].

[*40 minutes*] User-centric evaluation metrics.

- Introduce previous work on evaluation metrics for vertical search, including offline metrics and online metrics [20, 27, 29].

[*10 minutes*] Conclusion and future directions.

- Wrap up and introduce promising future directions for this topic.

4 SUPPORTING MATERIALS

We will provide the materials:

(1) Slides.
(2) Protocol (processes and settings) to conduct field study, user study to collect and mine user behavior data.
(3) Code and data samples to follow experimental segments of the tutorial.
(4) Extensive annotated bibliography.

Those materials will be made available at https://github.com/THUxiexiaohui/SIGIR-2020-Tutorial.

5 RELATED MATERIALS

No tutorial is an island. This tutorial aims at offering a comprehensive picture of user behavior modeling for vertical search, a topic that is increasingly attracting attention but that has not been treated systematically in any of the the tutorials listed. Other recent tutorials offer further in-depth treatments of some of the topics that we cover in this tutorial.

For example, Lalmas and Hong [10] present a tutorial on metrics of user engagement at WSDM 2018. They introduce how to leverage collected knowledge about the daily online behavior of millions of users to understand what engage them short-term and long-term. A similar tutorial has also been presented at WWW 2019. Moreover, Omidvar-Tehrani and Amer-Yahia [14] review research on User Group Analytics (UGA) and discuss different approaches and open challenges for group discovery, exploration, and visualization in a CIKM 2018 tutorial. At SIGIR 2019, Zarrinkalam et al. [28] introduce methods for extracting, mining and predicting users' interests from social networks. Information source, mining techniques and evaluation methodologies are covered in this tutorial.

We encourage our participants to follow up and deepen their understanding of modeling user behavior for vertical search by also studying these materials.

ACKNOWLEDGMENTS

This work is supported by the National Key Research and Development Program of China (2018YFC0831700), Natural Science Foundation of China (Grant No. 61622208, 61732008, 61532011, 61902209), Beijing Academy of Artificial Intelligence (BAAI) and the Innovation Center for Artificial Intelligence (ICAI). All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

REFERENCES

[1] Qingyao Ai, Yongfeng Zhang, Keping Bi, Xu Chen, and W Bruce Croft. 2017. Learning a Hierarchical Embedding Model for Personalized Product Search. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 645–654.
[2] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W Bruce Croft. 2018. Target Apps Selection: Towards a Unified Search Framework for Mobile Devices. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 215–224.
[3] Andrei Broder. 2002. A Taxonomy of Web Search. In *ACM Sigir forum*, Vol. 36. ACM New York, NY, USA, 3–10.
[4] David Carmel, Elad Haramaty, Arnon Lazerson, Liane Lewin-Eytan, and Yoelle Maarek. 2020. Why Do People Buy Seemingly Irrelevant Items in Voice Product Search? On the Relation between Product Relevance and Customer Satisfaction in eCommerce. In *Proceedings of the 13th International Conference on Web Search and Data Mining*. 79–87.
[5] Yingtong Dou, Weijian Li, Zhirong Liu, Zhenhua Dong, Jiebo Luo, and Philip S Yu. 2019. Uncovering Download Fraud Activities in Mobile App Markets. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 671–678.
[6] Yangyang Guo, Zhiyong Cheng, Liqiang Nie, Xin-Shun Xu, and Mohan Kankanhalli. 2018. Multi-modal Preference Modeling for Product Search. In *Proceedings of the 26th ACM international conference on Multimedia*. 1865–1873.
[7] Bernard J Jansen. 2008. Searching for Digital Images on the Web. *Journal of Documentation* (2008).
[8] Shubhra Kanti Karmaker Santu, Parikshit Sondhi, and ChengXiang Zhai. 2017. On Application of Learning to Rank for E-commerce Search. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 475–484.
[9] Diane Kelly, Jaime Arguello, Ashlee Edwards, and Wan-ching Wu. 2015. Development and Evaluation of Search Tasks for IIR Experiments Using a Cognitive Complexity Framework. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval*. 101–110.
[10] Mounia Lalmas and Liangjie Hong. 2018. Tutorial on Metrics of User Engagement: Applications to News, Search and E-commerce. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 781–782.

[11] Beibei Li, Anindya Ghose, and Panagiotis G Ipeirotis. 2011. Towards a Theory Model for Product Search. In *Proceedings of the 20th international conference on World wide web*. 327–336.

[12] Wendy W Moe. 2003. Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-store Navigational Clickstream. *Journal of consumer psychology* 13, 1-2 (2003), 29–39.

[13] Neil O'Hare, Paloma De Juan, Rossano Schifanella, Yunlong He, Dawei Yin, and Yi Chang. 2016. Leveraging User Interaction Signals for Web Image Search. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 559–568.

[14] Behrooz Omidvar-Tehrani and Sihem Amer-Yahia. 2018. User Group Analytics: Discovery, Exploration and Visualization. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 2307–2308.

[15] Dae Hoon Park, Yi Fang, Mengwen Liu, and ChengXiang Zhai. 2016. Mobile App Retrieval for Social Media Users via Inference of Implicit Intent in Social Media Text. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. 959–968.

[16] Dae Hoon Park, Mengwen Liu, ChengXiang Zhai, and Haohong Wang. 2015. Leveraging User Reviews to Improve Accuracy for Mobile App Retrieval. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 533–542.

[17] Parikshit Sondhi, Mohit Sharma, Pranam Kolari, and ChengXiang Zhai. 2018. A Taxonomy of Queries for E-commerce Search. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 1245–1248.

[18] Yang Song, Hao Ma, Hongning Wang, and Kuansan Wang. 2013. Exploring and Exploiting User Search Behavior on Mobile and Tablet Devices to Improve Search Relevance. In *Proceedings of the 22nd international conference on World Wide Web*. 1201–1212.

[19] Ning Su, Jiyin He, Yiqun Liu, Min Zhang, and Shaoping Ma. 2018. User Intent, Behaviour, and Perceived Satisfaction in Product Search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 547–555.

[20] Liang Wu, Diane Hu, Liangjie Hong, and Huan Liu. 2018. Turning Clicks into Purchases: Revenue Optimization for Product Search in E-commerce. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 365–374.

[21] Zhijing Wu, Yiqun Liu, Qianfan Zhang, Kailu Wu, Min Zhang, and Shaoping Ma. 2019. The Influence of Image Search Intents on User Behavior and Satisfaction. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. ACM, 645–653.

[22] Xiaohui Xie. 2019. User Behavior Modeling for Web Image Search. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. 826–827.

[23] Xiaohui Xie, Yiqun Liu, Maarten de Rijke, Jiyin He, Min Zhang, and Shaoping Ma. 2018. Why People Search for Images Using Web Search Engines. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM, 655–663.

[24] Xiaohui Xie, Yiqun Liu, Xiaochuan Wang, Meng Wang, Zhijing Wu, Yingying Wu, Min Zhang, and Shaoping Ma. 2017. Investigating Examination Behavior of Image Search Users. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 275–284.

[25] Xiaohui Xie, Jiaxin Mao, Maarten de Rijke, Ruizhe Zhang, Min Zhang, and Shaoping Ma. 2018. Constructing an Interaction Behavior Model for Web Image Search. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 425–434.

[26] Xiaohui Xie, Jiaxin Mao, Yiqun Liu, Maarten de Rijke, Qingyao Ai, Yufei Huang, Min Zhang, and Shaoping Ma. 2019. Improving Web Image Search with Contextual Information. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM, 1683–1692.

[27] Xiaohui Xie, Jiaxin Mao, Yiqun Liu, Maarten de Rijke, Yunqiu Shao, Zixin Ye, Min Zhang, and Shaoping Ma. 2019. Grid-based Evaluation Metrics for Web Image Search. In *The World Wide Web Conference*. ACM, 2103–2114.

[28] Fattane Zarrinkalam, Hossein Fani, and Ebrahim Bagheri. 2019. Extracting, Mining and Predicting Users' Interests from Social Networks. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1407–1408.

[29] Fan Zhang, Ke Zhou, Yunqiu Shao, Cheng Luo, Min Zhang, and Shaoping Ma. 2018. How Well do Offline and Online Evaluation Metrics Measure User Satisfaction in Web Image Search?. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 615–624.

[30] Juchao Zhuo, Zeqian Huang, Yunfeng Liu, Zhanhui Kang, Xun Cao, Mingzhi Li, and Long Jin. 2015. Semantic Matching in APP Search. In *WSDM*. 209–210.

[31] Agustin Zuniga, Huber Flores, Eemil Lagerspetz, Petteri Nurmi, Sasu Tarkoma, Pan Hui, and Jukka Manner. 2019. Tortoise or Hare? Quantifying the Effects of Performance on Mobile App Retention. In *The World Wide Web Conference*. 2517–2528.