How Do Sponsored Search Results Affect User Behavior in Web Search?

Zeyang Liu, Yiqun Liu[1], Min Zhang, and Shaoping Ma

State Key Laboratory of Intelligent Technology and Systems,
Tsinghua National Laboratory for Information Science and Technology,
Department of Computer Science and Technology, Tsinghua University, Beijing, 100084,
China P.R.
{zy-liu13,yiqunliu,z-m,msp}@tsinghua.edu.cn

Abstract. Sponsored search has gained considerable attention in Web search-related studies. In this paper, we investigate into the user interaction process on search engine result pages (SERPs) considering the influence of sponsored search results. We integrate different styles of sponsored links into SERPs to systematically study how they affect users' behavioral preferences. Based on an analysis into users' eye movements and click-through behavior, we showed that different presentation styles among sponsored links might lead to different behavior biases, not only for the sponsored search results but also for the organic ones. We also showed that search task type might influence the click-through of sponsored links, whereas this factor has little effect on the visual attention that users devote to top sponsored results. Interestingly, it seems that user satisfaction, which was intuitively measured by participants, was unaffected by the different sponsored search result appearances. These results help us better understand how users interact with search engines and how we can improve the placement of sponsored search results on SERPs.

Keywords: user behavior, sponsored search, eye tracking, click-through

1 Introduction

Web search has revolutionized the way users obtain Web information by providing large quantities of information efficiently and freely. Because of the differences between users' needs and advertisers' desires in search engines, it is difficult to achieve the perfect balance between sponsored and organic search results given the fact that sponsored search is a major source of revenue for commercial search engines. It is therefore important to gain a good understanding of how people interact with search engine result pages (SERPs), especially with sponsored links.

Eye tracking devices, which are designed to record participants' eye gaze information, have been widely adopted to explore search users' examination behavior. Due to the behavior of examination is closely related to the decision whether users

[1] Corresponding author.

would click the results [4, 14], eye-tracking studies have been a major way to improve relevance feedback performance based on click-through behavior logs. In this paper, we perform an eye-tracking analysis to explore the potential impact of sponsored search results on users' examination behavior and study the user's decision process on the SERPs that contain sponsored search results.

There are already a number of existing eye-tracking studies which looked into user's interaction with sponsored search results. For example, Buscher et al. [11] indicated that quality of sponsored search results (i.e. relevance to user queries) affect users' visual attention allocation. Djamasbi et al. [2] showed people behave differently on a same sponsored search result when it is displayed in different devices. Although these studies revealed a number of important findings, the results are based on a specific presentation style. Most of them only concentrated on the content-based influence and rarely considered the presentation biases on sponsored links.

Essentially, there is a large difference in sponsored result styles among different search engines. By considering the users' online experience, most commercial search engines usually adopt their own special layouts for sponsored links, and several of them even provide two different appearances of the same advertisement at different times. Figure 1 shows examples of sponsored results from Google and Baidu (China's most popular search engine). Since existing studies have demonstrated that a presentation bias widely exists in search results [3] and especially in vertical results [4], it is reasonable to presume that the diversity in advertisement presentation styles may also lead to different user behaviors in a web search. Taking it one step further, we seek to understand how user behavior (e.g., examination, click preference) is affected by different presentation of the sponsored search results and why certain existing factors do not suffer from the impact of these differences.

Fig. 1. Different presentation styles of sponsored results on SERPs (Baidu and Google)

Due to the existence of banner blindness [1], sponsored links may only have limited effects on user behaviors. However, according to our findings, the different types of sponsored links, which are displayed on the top position of SERPs, might inevitably capture users' visual attention during the search task. We believe that this phenomena may not only affect the users' behavior on organic results but also have an additional influence on search efficiency. For example, do users still skip banner ads if there is no difference between the sponsored and organic results? How do users change their behavior if we manually vary the sponsored link's style? Taking this case into account, we also analyze the effects of users' examination and click behavior on the organic results to completely analyze the decision process with different presentation forms of sponsored links.

The rest of the paper is organized as follows: we first present an overview of the related work. Then, we introduce the experimental design for collecting user behavior information through the use of eye tracking. After analyzing the user behavior data based on the eye-tracking experiment, we provide several conclusions in Section 4. A summary of the experiment results is presented in Section 5.

2 Related Work

2.1 Web Search Behavior

Because eye-tracking technology was first applied in reading research [5, 6], it has become increasingly popular in many domains associated with user behavior research, such as human-computer interaction interface design and autism diagnoses. In 2004, Granka et al. [7] proposed an eye-tracking approach to analyze user behavior on SERPs and to infer the relevancy of documents in the Web search. Comparing the implicit feedback versus the explicit feedback, Joachims et al. [8, 9] indicated that the users' decision processes are influenced not only by the relevance of the results but also by the *trust* and *quality* of the result set. Huang et al. [10] showed that cursor movements are particularly correlated to the viewing position based on the eye-tracking analysis. Bar-llan et al. [3] studied presentation biases and indicated that users tend to choose one of the top results on the SERPs. They also proposed a site reputation bias in their research. Wang et al. [4] found that the existence of presentation bias may also lead to different behavioral biases on federated searches. Most of these studies were concerned about the user behavior from the 10 organic results.

2.2 Studies on Sponsored Search Results

Because sponsored search results are becoming essential and important components of the SERPs in current commercial search engines, considerable work has focused on this regard. Buscher et al. [11] reported that the users' visual attention devoted to SERPs depends on both the task type and ad quality, and the order in which ads of different quality were presented also influence the user behavior. Jansen et al. [12] integrated sponsored and organic links into a single listing on SERPs. They showed that the combination of these results does not appear to increase the clicks on sponsored listings.

Even though they made no difference between sponsored and organic links on SERPs, the ranking and mixture of these results, which depends on the built-in algorithm of meta-search engines, might not be widely representative of the web search. Essentially, the position of the sponsored and organic results is relatively fixed and are separated rather than mixed in the major search engines. Ovens et al. [1] proposed text advertising blindness, which may cause a user to skip ads on a web page. Further, Djamasbi et al. [2] focused on the user's behavior on mobile devices. They conducted a study in which users had to interact with the SERPs that contained advertisements on mobile phones. Interestingly, they found that Generation Y did not exhibit banner blindness on mobile screens when most of them viewed both the advertisements and organic results.

Overall, most prior studies on sponsored links adopted a controlled experimental approach to characterize the visual attention or mouse movements on SERPs. Nevertheless, to the best of our knowledge, previous work rarely considered the presentation effects of sponsored results and strictly analyzed the influence, which might cause different users' examination and clicks on SERPs, especially on the organic results. Considering the existence of presentation biases on result lists (shown in Section 2.1), we are curious about whether the presentation forms of sponsored links could also lead to different user behavior. In this paper, we analyze the influence of presentation forms on the users' examination as well as the click preference when different styles of sponsored links are presented in random order. We also attempt to determine how sponsored result styles affect user satisfaction to provide a reasonable strategy on the presentation of sponsored results.

3 Collecting User Behavior Information

3.1 Participants

Thirty-three participants, all of whom were undergraduate students from a major university in China, were recruited using cash compensation. The students range in age from 17 to 20 with a mean age of 18.5 and have various majors in this university. The gender distribution is split between 14 males and 19 females. Because of the precision limitation of the eye tracker, not all of their eye movement data were available. As a result, 30 of the students that went through the calibration of correctness were finally taken into account. Particularly, all but two participants self-reported that the search engine we provided in this experimentation was their primary search instrument in their daily life, and they were very familiar with its operation.

3.2 Tasks

Participants had to complete the same set of 24 search tasks by using the same search engine, which we assigned. Each of these participants was required to solve these tasks in random order, which could eliminate the effects of trust bias [8, 9]. To make the initial result page comparable, we assigned an initial query for each task. All of these queries were selected from real-world commercial search logs so that they could reflect the users' search intention realistically. We extracted 810 medium frequency queries

automatically from the raw search logs[2], which were supplied by a major commercial search engine in China. Then, the original SERPs related to these queries were examined one by one, and the SERPs that contained more than three sponsored links could be selected into the initial SERP set. Finally, there were seven navigational queries (i.e., to reach a particular web site) and 17 informational/transactional queries (i.e., to acquire specific information) in our experimentation. Moreover, we also provided a short description for each query to avoid possible ambiguities. Several examples of the initial queries are shown in Table 1.

We transferred the search sever that stored the static set of SERPs captured from the original search engine. This allowed us to have a consistent initial SERP for each query and to strictly control the experimental factor, which may affect the user's behavior. After the initial SERP was presented, there were no restrictions on users' behavior on the search result pages. Thus, participants could be free to behave on the SERP as usual in ways, such as clicking the related search links, jumping to the next page of results, and reconstructing the query. The aim of this design is to achieve the real user appearances on search results in a laboratory environment.

Table 1. Examples of search queries adopted in the eye-tracking experimentation

Query	Translation	Intent Type
中国移动网上营业厅	China Mobile Online Service	Navigational
悦达起亚k2报价	Price of KIA K2	Informational
多动症的表现	The symptoms of ADHD	Informational

Unlike previous work that only concentrated on user behavior, we also consider the user's online experience. In this study, user satisfaction on SERPs was collected in our system. Participants were required to evaluate his/her satisfaction intuitively on the initial result page after completing a task session. They could score their evaluation ranging from 1 to 5: a score of 1 represents user disappointment on this page, while a score of 5 stands for the highest level of user satisfaction. To some extent, these scores could reflect the relative level of users' online experience across these tasks.

3.3 SERP Generation

As we described in section 1, the appearance of sponsored links may be largely different among different search engines. Even though this may lead to a wide variety of sponsored styles, there is still commonality among them. After observing the layout of major search engines such as Google, Bing, Baidu and Yahoo!, we classify the presentation styles of sponsored search results into two categories:

- With background color (WBC): this style of sponsored links ordinarily contains a particular color background, which makes the sponsored ads more special and outstanding than the other elements on the page. As a traditional style, it is the most widely used in contemporary search engines, such as Yahoo! and Bing.

[2] http://www.sogou.com/labs/dl/q.html

- No background color (NBC): Compared with the WBC, sponsored links with this style seem to be more similar to the organic results. Intuitively, there is no difference between the sponsored links and organic results except for several unobvious sponsored tags or brands. A good example of this is Google sponsored results.

Depending on the analysis above, we focus on the influence of these two types of sponsored link styles in this paper, which could cover the majority of the application area for the existing sponsored search results. Furthermore, because the negative effects might be caused by the different layouts of SERPs, it is important to provide a uniform interface for each initial result page. For this reason, we finally generate our result lists based on the layout of Baidu SERPs, which could display two types of sponsored links (WBC and NBC style) as a realistic component of the SERPs.

In the study, each of the initially fixed SERPs mainly consists of three top sponsored links and ten organic results. All of the result lists, including the sponsored links and organic results, were captured from the same original search engines so that the initial SERPs for all components could look as natural as possible. Particularly, while the sponsored search results were shown in different styles, both the content of the sponsored results (e.g., link title, snippet etc.) and the order of sponsored results remained consistent for the same query. Figure 2 shows an example of the interface for the SERPs, which has a different sponsored link style.

Fig. 2. Examples of sponsored results with and without background color for the same query

In addition, based on the conclusion of Wang et al. [4], organic results need to be tackled by removing the vertical results (e.g., images, videos, etc.) to eliminate the influence of other elements on the SERPs that we do not expect.

Following the pre-procedure steps above, we implemented 48 initially fixed SERPs in total. Half of them contain WBC sponsored links, and the rest contain NBC sponsored links. The 24 tasks were equally split between these two types of styles. For each of the participants, these two types of sponsored link styles were selected randomly and were shown alternatively for a specific query.

3.4 Measurements

The main goal of this paper is to study how users distribute their visual attention and clicks on different components of the result lists when sponsored links are presented on SERPs. For this reason, both the users' eye movements and clicks are collected in our experimentation.

We use the Tobii X2-30 eye tracker, which is able to record the main indicators of ocular behavior (e.g., fixations, saccades, pupil dilation etc.) and to capture the participants' eye movements. Because fixations are the most relevant metric for evaluating information processing in an online search [8], we detect fixations by using built-in algorithms from Tobii Studio. In these algorithms, a fixation is defined as a gaze duration above 100ms around a specific location. Additionally, to analyze the differences among the various parts of the SERP, we labeled each result document boundary manually and recorded the fixation duration per result document.

Because users' click and scroll behavior is another important metric in our study, we also capture and analyze users' mouse movements during the search tasks, including the click time, scroll time, and clicked URL. For example, we count the number of clicks on sponsored links and organic results, respectively. Moreover, user satisfaction regarding the sponsored SERP is taken into account as well, which has been discussed in the previous section.

4 Analysis of User Behavior

Our study focuses on the potential influence of sponsored search results. We look into the users' actual examination and click behavior on sponsored and organic results and attempt to explore several potential impacts of sponsored links on SERPs. The following sections present this work in detail.

4.1 How do users examine/click on sponsored search results with different styles?

First, we study the influence of different styles on sponsored search results. Figure 3 compares the average fixation duration on sponsored links that are displayed in WBC and NBC style.

It is clear that the fixation duration on the NBC sponsored results ($\mu = 2197.667$ $\sigma = 131.783$) was consistently higher than that on the WBC sponsored results ($\mu = 1095.667$ $\sigma = 59.219$), which means that it took substantially more time for participants to view the sponsored results when the NBC style was presented. From this bar chart, we can see the visual attention devoted to the sponsored links that were displayed in the NBC style is approximately twice as much as that displayed in the WBC style. To some extent, this result reflects that participants might tend to spend more time distinguishing the real organic results from the whole result lists because of the extremely similar appearances between the sponsored and organic results.

Similarly, this behavior bias also causes significant differences in the average click numbers, which are shown in Figure 4. We can find that the average click number on the NBC style ($\mu = 0.165$ $\sigma = 0.027$) is almost three times greater than that of the WBC counterparts ($\mu = 0.048$ $\sigma = 0.013$). Interestingly, for each specific style, while the third-ranked result receives the lowest click-through among the three sponsored links, it accounts for the highest proportion from the total amount of visual attention. It can be

observed that participants might more easily recognize the search advertisements and avoid clicking them when they spend more time examining these links.

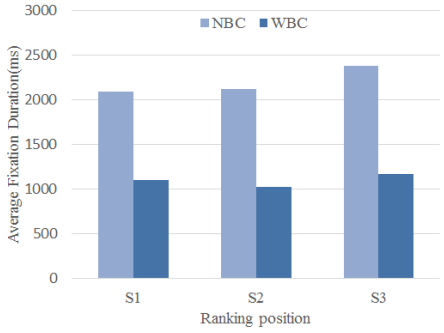

Fig. 3. Comparison of the average fixation duration on sponsored links that were displayed in WBC and NBC style

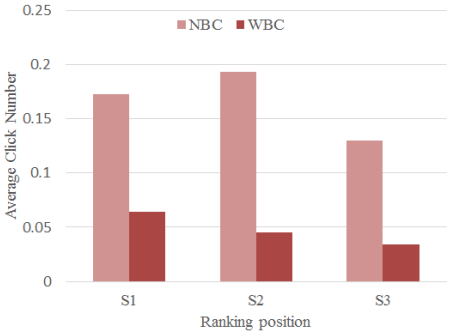

Fig. 4. Average number of clicks on the sponsored links that were displayed in WBC and NBC style

In conclusion, the presentation forms of the sponsored links had a noticeable influence on users' performance, including user examination and click behavior. When the appearances of sponsored links were similar to the organic links, it seemed that these sponsored results might be more likely to attract additional attention and click-through. In addition, we found that additional examination on sponsored results did not absolutely increase the corresponding click-through or even reverse the click-through on them.

4.2 How do users examine/click on organic results on a SERP with sponsored results?

Following the analysis in section 4.2, we are also curious about the effects of the organic results on the SERP containing sponsored results. Figure 5 illustrates an average fixation duration for organic links that are broken down separately for the sponsored links, which were presented in the WBC and NBC styles. Compared with the influence on the sponsored results, there is a smaller difference in the examination of organic results in general. Interestingly, for the NBC style, the visual attention that participants distributed to the third ranked sponsored result (mean time = 2383.958 ms σ = 904.634) was very close to that of the first organic result (mean time = 2623.667 ms σ = 827.776). To a certain extent, this reflects the fact that participants were more likely to be confused by the hidden appearance of sponsored links and devote more time to the examination to these links to eliminate misleading clicks. Furthermore, unlike the performance on the NBC style, we can see that participants tended to overlook the WBC sponsored links and focused more on the first organic results. Based on this analysis, we can infer that the presentation forms might be an important factor that can cause users' banner blindness differently.

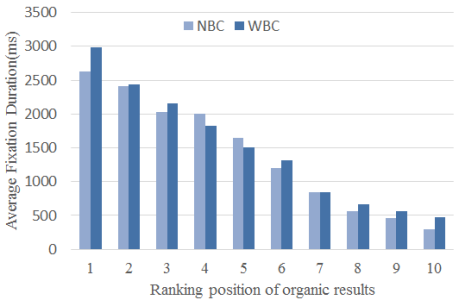

Fig. 5. Average fixation duration on organic results split between the NBC and WBC styles of the sponsored links

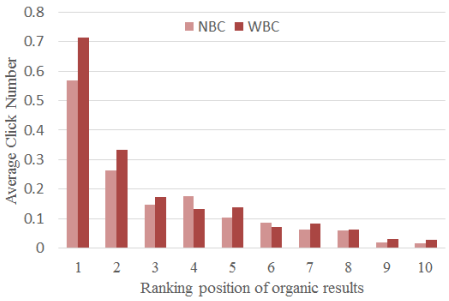

Fig. 6. Average number of clicks on the organic results split between the NBC and WBC styles of the sponsored links

Conversely, we found that the sponsored links' presentation affected the distribution of clicks on organic results. Because the NBC style attracts more visual attention and clicks, the average click number for the organic results in the NBC style ($\mu = 0.150$) is less than that in the WBC style ($\mu = 0.177$). Because of the considerable distraction from the sponsored links, especially those whose appearances are similar to the organic results, it might be easier for users to lose patience when examining and clicking the links on the organic result lists.

To summarize the organic results, the presentation of sponsored links has significant effects on the distribution of clicks but not on the examination behavior. Additionally, we found that users' banner blindness might be influenced by the appearances of sponsored results to some extent.

4.3 How does the task type affect users' examination/click-through behavior?

It can be observed from Figure 7 that most of the organic results attracted more examination for informational tasks than navigational tasks, whereas the distribution of visual attention was unaffected on the sponsored results. These results are consistent with the prior studies concerning the effects of the task type, such as [11, 13, 14].

To take a further step, we analyze the corresponding click-through behavior regardless of the presentation bias of the sponsored links, which is shown in Figure 8. We can see that the distribution of click-throughs on the organic results (organic 2 to 9) also reflected this bias in which users allocated more clicks on the links in the informational tasks rather than in the navigational tasks. However, the distribution of click-throughs on the sponsored links was not in line with this rule even though the visual attention on these click-through was almost the same. For example, the click-through of the third ranked link accounted for the highest proportion among the top three sponsored results in navigational tasks, while it made up the lowest proportion in the informational tasks. This indicated that the task type bias might also lead to the

click-through of sponsored results differently even though it did not have a noticeable effect on the distribution of the examination on them.

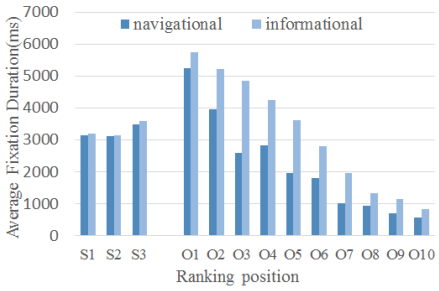

Fig. 7. Average fixation duration on the sponsored and organic results split between the different task types (navigational/informational and transactional)

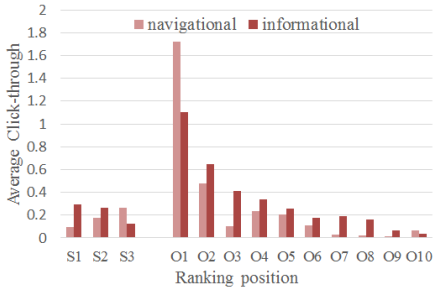

Fig. 8. Average click-through of the sponsored and organic results split by different task type (navigational/informational and transactional)

4.4 How does the sponsored result style affect user satisfaction?

Because the sponsored result style causes a different influence on both the sponsored and organic results, we are also concerned about how users feel about the SERPs when the different sponsored link style was displayed. Figure 9 shows the average score for the initial SERPs of 24 tasks, which contained the NBC or WBC style.

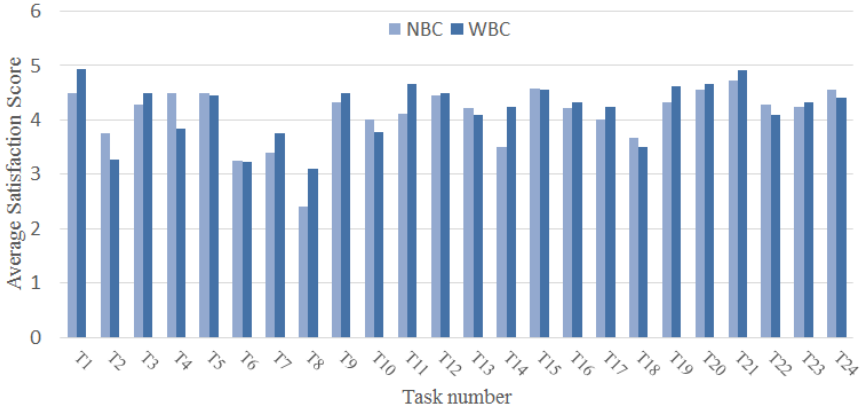

Fig. 9. Average satisfaction score for the initial SERPs of 24 tasks that contained the NBC or WBC style

The average satisfaction score for the NBC style is 4.098 ($\sigma = 0.275$), while the score for the WBC style is 4.188 ($\sigma = 0.259$). Interestingly, user satisfaction seemed to be

unaffected by the varied style of sponsored links in our experimentation. In other words, there was nearly no difference between the users' online experience on SERPs when the sponsored links were presented in the WBC or NBC style ($r = 0.798$ $p < 0.01$). We deduced that part of the reason for this is that participants might not really discover the existence of sponsored links if their presentation styles were similar to the organic results. This means that participants might regard these hidden advertisements as organic results and might be more tolerant toward the misleading clicks on SERPs. Thus, the presentation forms, which are similar to the organic style, may provide better strategies for improving the placement of sponsored search results on SERPs.

5 Conclusions and Future Work

We presented a study using eye-tracking techniques to explore the effects of sponsored search results on user decision strategies for web searches. Different from most previous approaches that focused on a specific presentation form of sponsored links, we systematically investigate how variations in the appearance of sponsored results affect user behavior on SERPs.

First, we found that the presentation styles of sponsored links have a significant influence on users' examination and clicks, where participants contributed to both the sponsored and organic results. There is a shape difference in the result for the examination and click-through behavior when the different sponsored styles were presented. In particular, we found that users' examination and clicks were not strictly correlated with each other on the same sponsored link style. In other words, the more visual attention that users devoted to the sponsored results might not result in more click-through to them.

Second, our study showed that the organic result entries capture more visual attention from users in informational tasks than in navigational tasks as opposed to task biases against sponsored links, which is consistent with previous studies [11]. However, the users' examination behaviors on sponsored results are not influenced by this factor, and there is still a real difference in click-through for the different types of tasks.

Finally, based on the analysis into users' satisfaction on different interfaces of the SERPs, it is interesting to find that the organic-like sponsored results, which are similar to the real organic result entries, did not significantly decrease the users' online experience even though these results might more easily lead to users' distraction on the SERPs.

In the near future, we hope to extend this work to incorporate the effects of sponsored results into a search click model and further apply this to large-scale data. Moreover, we plan to analyze the influence of sponsored links on federated searches based on our findings in this paper.

6 Acknowledgements

This work was supported by Natural Science Foundation of China (61073071). Part of the work has been done at the Tsinghua-NUS NExT Search Centre, which is

supported by the Singapore National Research Foundation & Interactive Digital Media R&D Program Office, MDA under research grant (WBS:R-252-300-001-490)

References

1. Owens, J. W., Chaparro, B. S., Palmer, E. M.: Text advertising blindness: the new banner blindness? In: Journal of Usability Studies, Vol. 6, No. 3, pp. 172-197 (2011)
2. Djamasbi, S., Hall-Phillips, A., Yang, R. R.: SERPs and ads on mobile devices: an eye tracking study for generation y. In: Universal Access in Human-Computer Interaction. User and Context Diversity, pp. 259-268. Springer Berlin Heidelberg (2013)
3. Bar-Ilan, J., Keenoy, K., Levene, M., Yaari, E.: Presentation bias is significant in determining user preference for search results—A user study. In: Journal of the American Society for Information Science and Technology, Vol. 60, No. 1, pp. 135-149 (2009)
4. Wang, C., Liu, Y., Zhang, M., Ma, S., Zheng, M., Qian, J., Zhang, K.: Incorporating vertical results into search click models. In: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, pp. 503-512. ACM (2013)
5. Jacob, R. J., Karn, K. S.: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In: Mind, Vol. 2, No. 3, pp. 573-605 (2003)
6. Reichle, E. D., Rayner, K., Pollatsek, A.: Eye movement control in reading: Accounting for initial fixation locations and refixations within the EZ Reader model. In: Vision research, Vol. 39, No. 26, pp. 4403-4411 (1999)
7. Granka, L. A., Joachims, T., Gay, G.: Eye-tracking analysis of user behavior in WWW search. In: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 478-479. ACM (2004)
8. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Gay, G.: Accurately interpreting clickthrough data as implicit feedback. In: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 154-161. ACM (2005)
9. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F., Gay, G.: Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. In: ACM Transactions on Information Systems (TOIS), Vol. 25, No. 2, pp. 7-32 (2007)
10. Huang, J., White, R. W., Dumais, S.: No clicks, no problem: using cursor movements to understand and improve search. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1225-1234. ACM (2011)
11. Buscher, G., Dumais, S. T., Cutrell, E.: The good, the bad, and the random: an eye-tracking study of ad quality in web search. In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval, pp. 42-49. ACM (2010)
12. Jansen, B. J., Spink, A.: Investigating customer click through behaviour with integrated sponsored and nonsponsored results. In: International Journal of Internet Marketing and Advertising, Vol.5, No. 1, pp. 74-94 (2009)
13. Broder, A.: A taxonomy of web search.In: ACM Sigir forum, Vol. 36, No. 2, pp. 3-10. ACM (2002)
14. Cutrell, E., Guan, Z.: What are you looking for?: an eye-tracking study of information usage in web search. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 407-416. ACM (2007)