The Effectiveness of Online Sponsored Search from Search Users' Perspective

Yiqun Liu State Key Lab of Intelligent Technology & Systems Tsinghua University Beijing, China P.R.

liuyiqun@tsinghua.edu.cn

Bo Zhou, Shaoping Ma State Key Lab of Intelligent Technology & Systems Tsinghua University Beijing, China P.R.

zhoubo2000@gmail.com

Liyun Ru, Lei Chen State Key Lab of Intelligent Technology & Systems Tsinghua University Beijing, China P.R.

msp@tsinghua.edu.cn

ABSTRACT

With the explosive growth of information available on the Web, more and more users adopt search engines to collection information on the internet. Meanwhile, sponsored search has become one of the most popular forms of Internet advertising because of its effectiveness and feasibility. However, it remains questions to us whether the sponsored search results become obstacles in users' information acquisition process. With analysis into large scale Web user access logs, we obtained several Chinese commercial search engines' sponsored search statistics. We look into users' interaction behavior with sponsored search results. Our conclusion is that as a kind of advertising, sponsored search results can also meet users' information need as ordinary results do. We also found several factors that affect users' clicks on sponsored search results. To the best of our knowledge, this is the first study that uses Web access logs to estimate the effectiveness of sponsored search advertising.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval – Search process; H.3.4 [Information Storage and Retrieval]: Systems and Software – Systems and Software; H.3.5 [Information Storage and Retrieval]: Online Information Services – Commercial services, Web-based services.

General Terms

Experimentation

Keywords

Search Engine; User Behavior Analysis; Sponsored search

1. INTRODUCTION

With the explosive growth of information available on the Web, more and more users adopt search engines to collection information on the internet. Meanwhile, sponsored search has become one of the most popular forms of Internet advertising because of its effectiveness and feasibility. In "The 2007 special report on Chinese search engine market", it points out that in terms of the global market the size of search engine market continues to grow rapidly.... With the growth rate of 17.3%, it has achieved a scale of 2.85 billion U.S. dollars in 2007." Search engines have become an indispensable information access method in people's daily work and life. According to the statistics of China Net Information Center (CNNIC) in Jan. 2008, the prevalence rate of search engines in 210 million Internet users of

China has reached to 72.4%, which has exceeded the prevalence rate of e-mail service. Additionally, 84.5% of search engine users regard search engines as a main way to know a new website.

Therefore, getting higher rank in search results of search engines has become one of the most effective ways to attract users for Internet resources. Sponsored Search Service (provided by Google as the representative) has undoubtedly become the most direct way for advertisers to promote Internet resources. As a result, Sponsored search has also become one of the primary means for a search engine to profit.

Sponsored search refers to the advertising method in which the enterprises pay the search engine for the prominent location of their enterprise information on the search results page. The cost depends on the specific locations and the click-through rate of the information. PPC ads are typically in the form of "sponsored links" labels on the top or the sidebar of the search results page, shown in Figure 1.



Figure 1: example of sponsored search in search result pages (in this example, all results shown either in the sidebar or in the result list are sponsored search results)

Mainly because of its efficient, low risk and flexible features, sponsored search advertising has become the enthusiasm of many enterprises, particularly enterprise of small or middle size. The effectiveness of the sponsored search advertising relies on the quantity of search engine users and the fact that many users regard search engines as their main entrance to the Internet. As a consequence, although the search result pages contain sponsored search, the page views for the search result pages of some hot queries are all most at the same level as the PV of some web portals. And the majority of online users (62%) do not realize the

difference between organic search and sponsored search, so ads with higher ranks attract more attention. [1][2] demonstrate the effectiveness of sponsored search advertising by the establishment of a economic model, and analyze its difference from traditional advertising. Search engines charge for the sponsored search service in accordance with users' click-through rate, which is more reasonable than the charging method used by traditional media. The new advertising method can reduce business costs and risks and guarantee the ad effect at the same time. Enterprise users can choose not only the keywords for sponsored search but also where and when of the sponsored search. Some application process of sponsored search service can be finished all on-line, which leads to the new advertising method more flexible for users. Moreover, users are able to know the effect of their sponsored ads after a very short period of time, so they can conveniently adjust their sponsoring strategy.

However, with increasing number of sponsored search ads, people are puzzled by the fact: more and more top results of organic search are replaced by sponsored search, and the search result of some queries are all sponsored ads, such as "ticket", "join", "financial management" and so on. Whether or not it will affect users' experience? Whether or not it will lower the loyalty of users? Whether or not it will finally dilute the effect of ads? In this paper, we address these issues by investigating the following research questions:

- Does the appearance (e.g. the position on search result pages) of sponsored search result affect user clicks?
- 2. How many percentage of search users click on sponsored search results while using commercial search engines?
- 3. Which kind of keywords bring in more sponsored search clicks, hot keywords, cold keywords or the others?
- 4. Does the ranking of sponsored search results affect user clicks?
- 5. Can sponsored search result be regarded as "a medium of information" instead of obstacle in users' information acquisition processes?

Therefore this paper focuses on understanding users' response to the sponsored search ads with the analysis of users' search log, which leads to further understanding of the effect of sponsored ads. This paper is organized as follow: Sec.3 introduces log-collection methods and data analysis methods for the logs of search engines. In Sec.4 the analysis of sponsored ads effect is proposed according to the data of user behavior. Sec.5 concludes with some suggestions for the enterprise users and search engines.

2. RELATED WORK

With the explosive growth of information on the Web, search engine become one of the most important information portal in people's daily lives. Commercial search engines provide free information access to billions of Web search users every day and their primary business model is sponsored search and sponsored links. Sponsored links from sponsored search systems such as Google Adsense appears both on search engine result pages (Figure 2(a)) and on ordinary Web sites (Figure 2(b)). Collecting click-through information from large number of ordinary Web sites would be a rather difficult task and most previous work such as [1] and [3] are based on algorithmic analysis and very small scale experimental corpuses. Therefore, in this paper we focus on

sponsored search results on search engine result pages and try to find out whether sponsored links disturb searchers' information acquisition.



Figure 2 (a) Sponsored search results on search results pages



Figure 2 (b) Sponsored search results on ordinary Web pages

Because of its immense economic impact in commercial search engines, there are lots of researches on the effectiveness of sponsored search in search engine advertising.

Most of these efforts are based on user studies which involve small number of assessors. Due to different experiment settings, many researches got different or even contradictory results to a same question. [4] found that Internet users are more likely to click on an organic (non-sponsored) search link on Google according to an online survey of over 1000 individuals. [5][6][7] also found that sponsored links are less relevant than organic links and searchers have a strong preference for non-sponsored links according to small scale user studies of college students as evaluators. [9] adopted a similar method to evaluate the relevance of both sponsored and non-sponsored results given an E-commence query. However, their conclusion was different from previous works: average relevance ratings for sponsored and non-sponsored links are practically the same.

There are also a few researches which are adopted based on real-world data. [8] collected click-through information from a retailer that advertises on Google. However, they focused on the evidences which affect Google's paid search result rankings. [9] utilized click-through logs of a meta search engine to analyze user

preference on sponsored search results. They found that combining sponsored and non-sponsored links in a same result list does not increase clicks on sponsored listings.

Different from their works, our analysis into the performance of sponsored search is based on browser toolbar logs. With the development of search engine, Web browser toolbars become more and more popular recently. Lots of search engine develop toolbar software to attract more user visits (e.g. Google, Yahoo...). Web users usually adopt toolbars to get instant access to search engine services and to get browser enhancement such as pop-up window blocking and download acceleration. In order to provide value-add services to users, most toolbar services also collect anonymous click-through information from users' navigation behavior. Previous works such as [10] adopts this kind of click-through information to improve ranking strategy. The major advantages of adopting Web access log data in sponsored search research include the following aspects.

Large scale of user data can be collected without interrupting users' Web search behavior at very low cost. Therefore, the research results can be objective and reliable.

Performance of sponsored search results from different search engines can be collected and compared. In our research, most of the frequently used Chinese search engines are involved and their different kinds of sponsored search policies can be compared.



Figure 3. Sponsored search results appear at different positions on a search result page (snapshot of www.baidu.com, which is one of the most popular Web search engine in China)

Sponsored search results appear on search result pages in different forms. For example, they can appear at position B in Figure 3 and position A in Figure 3(a), which indicates that they are different from organic results. They can also appear at position A in Figure 2 and they look almost the same as a non-sponsored result. Analysis into Web access logs makes it possible to find out which form of sponsored search appearance is more effective.

3. THE COLLECTION OF USER BEHAVIOR DATA

3.1 Data Source

With the development of search engine, Web browser toolbars become more and more popular recently. Lots of search engine develop toolbar software to attract more user visits (e.g. Google, Yahoo, Baidu...). Web users usually adopt toolbars to get instant access to search engine services and to get browser enhancement such as pop-up window blocking and download acceleration. In order to provide value-add services to users, most toolbar services also collect anonymous click-through information from users' navigation behavior. Previous works such as [10] adopts this kind of click-through information to improve ranking strategy. In this paper, we also adopt Web access logs collected by search toolbar because this kind of data contains user behavior information from multiple search engines instead of a single engine. Information included in the click-through logs is shown in Table 1.

One kind of information different from traditional click-through information is the 'ID' of a search session. This information is used to record the query and click sequence given a certain session so that the number of clicks for a certain query and the behavior of changing queries can be recorded. This anonymous information is also used to separate different users while recording as little private information as possible.

Table 1. Click-through Information Used in the Automatic Evaluation Process

Name	Record Content		
Query	The user query submitted		
URL	URL of the result clicked by the user		
Engine	Name of the search engine which the user was using		
ID	Automatically assigned user's identification code according to his query session		
Time	Date and time of the clicking or querying event		
Machine	Software information		

The information shown in Table 1 can be easily recorded and are already recorded by commercial search engine systems. Therefore it is practical and feasible to obtain these types of information and to apply them in the automatic evaluation process. With the help of a popular commercial Chinese search engine, click-through logs of six most frequently-used Chinese search engines were collected from 22nd April 2008 to 13th May 2008. These click-through logs (recording altogether 900 million querying and clicking events) were adopted in our work.

The information shown in Table 1 is generally recorded by toolbars of majority of search engines, so the method proposed in this paper is of a significant reference to other toolbars. Meanwhile, by using ID assigned automatically by systems, there exists no private information, which guarantees the reasonableness of log analysis.

3.2 The analysis of URL format

After investigations into 6 famous Chinese search engines (Google, Yahoo, Baidu, Sogou, SouSou, Youdao), we discovered the rank ordered by click rate is: Google, Baidu, SouSou, Yahoo, Sogou, Youdao. We ignore Youdao, because Youdao only has sponsored-to-top ads but not sponsored-to-side ads. Although SouSou does not have sponsored-to-side ads and it uses the sponsored search ads of Google, we still keep SouSou in

consideration because its click rate is so high that it follows right after Google and Baidu.

Table 2. URL format of different search engines

URL format of search process	Default encoding	URL format
process	encouning	of ads
http://www.baidu.com/s?wd=	GBK	http://www.b
\$KEYWORD		aidu.com/bai
http://www.baidu.com/baidu? word=\$KEYWORD		du.php?
http://www.baidu.com.cn/s?w d=\$KEYWORD		
http://www.google.cn/search? q=\$KEYWORDie=\$CODE	UTF-8	http://www.g
http://www.google.com/search? q=\$KEYWORDie=\$CODE		oogle.cn/aclk ?
http://search.cn.yahoo.com/sear	GBK	http://click.p
ch?p=\$KEYWORDei=\$CO		4p.cn.yahoo.
http://www.yahoo.cn/s?p=\$K		com/ci_im?
EYWORD		
http://www.sogou.com/web?	GBK	http://click.c
query=\$KEYWORD		pc.sogou.co
http://www.sogou.com/sohu? query=\$KEYWORD		m/bill_searc
query=ψικΕ τ w σκΕ		h?
		http://click.c
		pc.sogou.co
		m/bill_biz?
http://www.soso.com/qw=\$	GBK	http://www.g
KEYWORD		oogle.cn/aclk
		?
<u> </u>		

4. THE EFFECT ANALYSIS OF SPONSORED SEARCH

4.1 Organic Search vs. Sponsored Search in Terms of Clicks

Table 3 shows the average clicks per day on search result and sponsored search of the 5 different search engines. The 2nd column is the division of clicks of sponsored ads with clicks of search results (called Ad-CTR in this paper); the 3rd column is the standard deviation of Ad-CTR and it's as an order of magnitude smaller as the average, which indicates there is little variation of Ad-CTR.

According to Table 3, we can discover that although search engines receive different amount of user clicks, for all of the 5 search engines, the daily ads clicks only account for 0.12%-0.27%

of the daily clicks of search results, which shows the promotion space of the sponsored search market.

It is noticeable that SouSou's Ad-CTR 1 times more than Google although SouSou uses Google's ads. The main reason is that there's no ad in sidebar of SouSou and the ads appearing in Google's sidebar appears on the top of the SouSou's search result pages. Although the number of ads is less than 3, ads on the top have more attraction than the ads on the sidebar.

Table 3. Clicks on search results and ads for different search engines (average of 21 days)

Search engine	Ad-CTR	Standard deviation	
Baidu	0.19%	2.26E-04	
Google	0.12%	1.14E-04	
Yahoo	0.18%	3.09E-04	
Sogou	0.16%	2.88E-04	
SouSou	0.27%	3.71E-04	

4.2 The Distribution of Sponsored Search Clicks

Table 4 shows the number of unique terms, the total clicks, the maximum of clicks, the average clicks (total clicks/term number)

Table 4. Statistics on query terms (21 days)

	# Unique terms	Total clicks	Average clicks	Maximum of clicks
Search results	1,304,211	81,185,985	6.102277317	384,798
Ads	69,938	145,186	2.07592439	812
Percentage	0.53%	0.18%	-	-

First, according to Table 4, terms with ad clicks account for a tiny fraction. The average clicks of ads for the terms containing ad clicks are 2 times, which is less than the average clicks of all terms-6 times. The maximum of ad-clicks of an single term is much less than the clicks of overall search results. This is mainly because the number of sponsored ads is much less than the number of search results.

Figure 4 (a) shows the distribution of #Terms in accordance with the frequency of query terms. The horizontal axis represents the query term frequency; the vertical axis represents the proportion of the query term number. From Fig. 4(a) we discovered that the query terms with less than 8 clicks account for 90% of the overall query terms. Figure 4 (b) shows the distribution of term clicks in accordance with the frequency of query terms. The main difference between Fig. 4 (a) and Fig. 4 (b) is that in Fig. 4 (b) the vertical axis represents the proportion of clicks. From Fig. 4 (b) we can see more that query terms with less than 4 ad clicks not only account for the majority of query terms with ad clicks but

also attract most of users' clicks (more than 50%), which also shows the ad clicks of a query term is only a little.

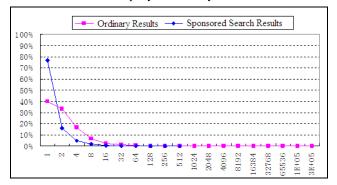


Figure 4 (a) The distribution of #Terms in accordance with the frequency of query terms. (Category axis: keyword frequency, value axis: the proportion of the query term number)

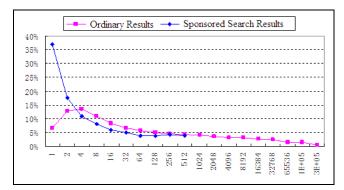


Figure 4 (b) the distribution of term clicks in accordance with the frequency of query terms. (Category axis: keyword frequency, value axis: the proportion of the query clicks)

4.3 The Relation Between Query Clicks and HotDegree

Definition 1. The *HotDegree* of a certain Web search query Q is defined as:

$$HotDegree(query\ Q) = \frac{\#(user\ click\ of\ Q)}{\#(average\ user\ click)}$$

In our work, we use the definition *HotDegree* to describe whether a keyword is a hot keyword. Query terms with *HotDegree* more than 10 are regarded as hot queries. According to our statistics, 493 query terms with more than 20 ad clicks are all hot queries, whose average clicks are 3774 times, highest clicks are 65456 times, and lowest clicks are 64 times. They all satisfy the definition of hot queries and with their clicks 10 times more than the average.

However, there exists a large part of hot queries with no ad click. 2090 hot queries out of 98099 in all (28.08%) have ad clicks. This is mainly because a large part of hot queries consist of navigational queries, such as "Baidu", "Taobao", "Xiaonei", "Sina" etc. There are about 50 navigational queries in the top 100

hot queries. We can find the related websites for these queries in the search results with high rankings, so there's no need for sponsored search advertising. Additionally, since search engines are regarded as the entrance to the Internet for many web users and the descriptions for websites are often with little variations (e.g. when a web user want to access to www.sina.com, his/her queries are often "新浪" or "sina", with no other variations.), the navigational queries obtain the highest clicks.

4.4 The Relation Between Ad Clicks and Rankings

"Promotion" and "Sponsored links" are the two most widely used forms of sponsored search. Google generally adopts the form of "Sponsored links" with most of the "Sponsored links" appear on the side bar and little appear on the top, which is marked with special background colors. The sidebar of Baidu is similar to Google's sidebar, but ads appearing on the top of the search result are called "Promotion" instead, whose format is nearly the same as the organic search. The display of sponsored search in Sogou is generally the same as Baidu, but "Sponsored links" with special background colors appear occasionally on the top of the search result page. In Sogou's search result page, "Promotion" and "Sponsored links" do not appear at the same time. The sidebar of Yahoo is more or less complicated, and its sponsored-to-top ads are less than 3 and are called "Promotion" which is marked with special background colors. SouSou has no sidebar in its search result page, and it display less than 3 "Sponsored links" from Google on the top and sometimes with no special background colors.

To summarize the difference between these two forms of sponsored search: "Promtion" is the form of search which has almost the same displaying format as the organic search with no special background color, and it's labeled only by the word "Promotion".

"Sponsored links" is the form of search which appears on the search result page with apparent differences from the organic search, such as appearing on the sidebar or marked with special background colors. Thus, "Sponsored links" generally have no impact on the number of the organic search.

Due to the fact that "85% of users only look at the top 10 results returned by search engines, i.e. the first page of the whole search results" [12]. We have made an investigation into the clicks of "Promotion" and "Sponsored links" with ranking from 1 to 10. The result is shown in Fig. 5. For "Promotion" and "Sponsored links", the clicks of the 1st result are 7 times and 9 times more than the 2nd result and account for 69% and 75% of the clicks of the top 10 results. Actually, there are often more than 1 sponsored search in the search result. In the other words, most users just click the ad with the highest rank.

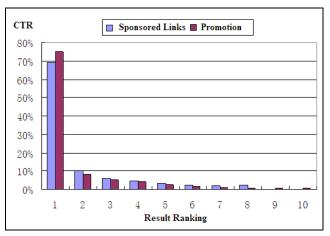


Fig. 5 The distribution of Ad CTR in accordance with ranking

4.5 The Impact of Sponsored Search to User Experience

From the observation of the statistics, we found the phenomenon that for some queries the clicks are not decreasing with the increase of ranking, which indicates that the sponsored search may disturb users. Thus, we crawled the search results of the queries with more than 10 clicks, and check whether there is sponsored ads in them. Through log analysis, we obtained the clicks of the pages containing "Promotion", "Sponsored links" and the clicks of the first result page. The statistics are shown in Fig. 6. The difference between Fig. 6 and Fig. 5 is that Fig. 6 contains the cases in which one page contains ads but has no ad click. In addition, since the different displaying strategies of "Promotion" and "Sponsored links" (discussed in Sec. 4.4), the ranking of "Promotion" is consistent with the ranking of organic search, however the ranking of "Sponsored links" is not consistent with the ranking of organic search. In the other words, in the pages containing "Promotion" the first result should be sponsored ad, but in the pages containing "Sponsored links" the first result is organic search result or "Promotion" instead of "Sponsored links". We can observe the impact of different displaying method to users.

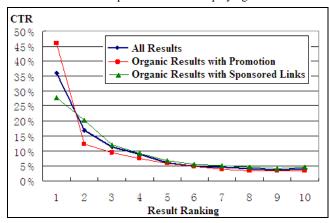


Figure 6 The distribution of clicks on different kinds of pages in accordance with increasing rankings

From Fig. 6, we found that generally clicks are in inverse proportion to ranking although there are some exceptions. In pages containing "Promotion", users are prone to click the ad in the 1st position; in pages containing "Sponsored links", clicks of search result are influenced by the "Sponsored links" and decrease not that sharply.

From the observation, there is no evidence indicating users dislike sponsored ads. On the contrary, users are prone to accept ads for reasons as follow:

- The optimization of the result ranking, which is done by search engines, improves the quality of the sponsored ads.
 E.g. the ranking of sponsored ads in Google's search result is not only related to the price enterprises have paid. If the ad with high ranking has not enough clicks, its ranking will be lowered. [5] also found that sponsored search is more relevant than organic search and the ranking of sponsored search is consistent with its relevance.
- 2. The queries containing sponsored ads are generally related to products, such as lottery, movie, cell phone etc. User's purpose is to search for the product information or do on-line shopping. And the purpose of advertisers is to claim they have various kinds of services, which satisfy the need of the users. Thus, it explain the high clicks.

Apparently, search engine companies ought to note the exceptions in case of degrading user experience.

5. CONCLUSIONS

This paper analyses the sponsored advertising form of five major search engines (Google, Baidu, Yahoo, SouSou, Sogou), and mines users' response toward the ads sponsored by these 5 search engines based on web user log containing 900 million user-clicks. Baidu, as the largest Chinese search engine, has the click rate of more than 10 million per day, while the click rate of sponsored ads accounts for 0.19% of daily click rate for search result. The space for sponsored search market is still great. The queries (more than 10 clicks) containing sponsored search ads only account for 5% of the total queries. Additionally there is 2 clicks in average for every query; most users will only look at the first search result page and click ads on the top. As a consequence, instead of advertising on several hot queries, more than advertising on relatively less popular queries and obtaining a higher ranking, which is possible to reduce the total cost.

According to the present situation, users do not have the tendency of disliking the sponsored ads. Instead, in search result page with sponsored search advertising, sponsored ads on the top increase click rate by 50%. However, the phenomenon that the inversion of click rate in some cases should be noticed by search engines supplier. The consistency between ads ranking and relevance rate still need further consideration. The form of sponsored search advertising is the issue in concern as well. "Sponsored links" is worse than "promotional" in terms of attracting click rate for ads. Although SouSou uses "Sponsored links" of Google, it modified its display strategy and has achieved ads-click rate of 0.27%, which is the most among the 5 search engines mentioned above.

6. REFERENCES

[1] Animesh, A., Ramachandran, V., and Viswanathan, S. 2007. An empirical investigation of the performance of online sponsored search markets. In Proceedings of the Ninth international Conference on Electronic Commerce (Minneapolis, MN, USA, August 19 - 22, 2007). ICEC '07, vol. 258. ACM, New York, NY, 153-160.

- [2] Anindya G, Sha Y. An Empirical Analysis of Sponsored Search Performance in Search Engine Advertising. NET Institute Working Paper, 2007, p7-35
- [3] Li, Y., Jhang-Li, J., and Lee, Y. 2007. Efficiency analysis for display ads and contextual search. In Proceedings of the Ninth international Conference on Electronic Commerce (Minneapolis, MN, USA, August 19 - 22, 2007). ICEC '07, vol. 258. ACM, New York, NY, 361-368.
- [4] Robyn Greenspan. Searching for Balance. The ClickZ Network in Apr 30, 2004. Online at http://www.clickz.com/showPage.html?page=3348071.
- [5] Jansen, B. J. and Resnick, M. 2006. An examination of searcher's perceptions of nonsponsored and sponsored links during ecommerce Web searching. J. Am. Soc. Inf. Sci. Technol. 57, 14 (Dec. 2006), 1949-1961.
- [6] Jansen, B. J. and Molina, P. R. 2006. The effectiveness of web search engines for retrieving relevant ecommerce links. Inf. Process. Manage. 42, 4 (Jul. 2006), 1075-1098.
- [7] Jansen, B. J. 2007. The comparative effectiveness of sponsored and nonsponsored links for Web e-commerce queries. ACM Trans. Web 1, 1 (May. 2007), 3.
- [8] [Ghose 2008] Ghose, A. and Yang, S. 2008. An empirical analysis of sponsored search performance in search engine

- advertising. In Proceedings of the international Conference on Web Search and Web Data Mining (Palo Alto, California, USA, February 11 12, 2008). WSDM '08. ACM, New York, NY, 241-250.
- [9] [Jansen and Spink, 2007] Jansen, B. J., & Spink, A. (2007). The effect on click-through of combining sponsored and non-sponsored search engine results in a single list. Sponsored Links Workshop. WWW 2007:16th International World Wide Web Conference, May 8-12, Banff, Canada.
- [10] [Bilenko 2008] Bilenko, M. and White, R. W. 2008. Mining the search trails of surfing crowds: identifying relevant websites from user activity. In Proceeding of the 17th international Conference on World Wide Web (Beijing, China, April 21 - 25, 2008). WWW '08. ACM, New York, NY, 51-60.
- [11] Bernard J. The Comparative Effectiveness of Sponsored and Nonsponsored Links for Web E-commerce Queries. ACM Transactions on the Web,2007, Vol. 1, Article 3.
- [12] Huijia Yu, Yiqun Liu, Min Zhang, Liyun Ru and Shaoping Ma, Research in Search Engine User Behavior Based on Log Analysis (in Chinese). Journal of Chinese Information Processing. Vol. 21(1): pp. 109-114, 2007.