Investigating the Relevance of Sponsored Results for Web Ecommerce Queries

Bernard J. Jansen
College of Information Sciences and Technology
The Pennsylvania State University
University Park, PA, 16801, USA
jjansen@acm.org

ABSTRACT

Are sponsored links, the primary business model for Web search engines, providing Web consumers with relevant results? This research addresses this issue by investigating the relevance of sponsored and non-sponsored links for ecommerce queries from the major search engines. The results show that average relevance ratings for sponsored and non-sponsored links are virtually the same, although the relevance ratings for sponsored links are statistically higher. We used 108 ecommerce queries and 8,256 retrieved links for these queries from three major Web search engines, Google, MSN, and Yahoo!. We present the implications for Web search engines and sponsored search as a long-term business model as well as a mechanism for finding relevant information for searchers.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**] – *Search process*: Measurement, Experimentation, Human Factors

General Terms

Design, Experimentation, Human Factors

Keywords

Web search engines, sponsored search, sponsored results, sponsored links, ecommerce searching, Web searching

1. INTRODUCTION

Web searching engines are the major portals for people as they seek online information, sites, and services. Search engines generally present two types of links in response to user queries -- non-sponsored and sponsored. Non-sponsored links are results returned based on the proprietary indexing and ranking algorithms of the particular search engine. Sponsored links are results returned based on the outcomes of proprietary online auctions where content providers / advertisers bid on query terms. Sponsored links (a.k.a., sponsored search, paid search, sponsored results) are the chief business model for Web search engines, providing profit for the search engine companies and financing the 'free' access to Web content for millions of users worldwide. However, are sponsored links providing relevant results to Web searchers?

Without doubt, sponsored links are one of the most influential innovations in Web search. In 2005, sponsored search was a $12 billion industry for the four largest search engines [6]. Businesses consider sponsored links a reliable marketing and profit avenue,

and search engines certainly consider sponsored search a workable business model. To provide a viable long-term revenue stream, however, sponsored links must be effective at providing relevant information for Web searchers. In this research, we investigate how relevant sponsored links are for Web ecommerce queries.

2. RELATED RESEARCH

Most of the major Web search engines are commercial entities requiring constant revenue to support the free information access provided to millions of searchers every day. There are a number of sponsored search systems on the Web today. As of 2006, Google Adwords and Yahoo! Search Marketing Services are the dominate players. These two systems provide the majority of sponsored links to not only Google and Yahoo! respectively, but also to numerous other search engines via third party agreements. In addition to Google and Yahoo!, Microsoft has entered the sponsored search market, and there are some other players in the field such as Snap.com, FindWhat, and Kanoodle. The economic effect of sponsored search is huge. In 2005, Web search engines displayed approximately 13 billion sponsored links in a given week, according to Nielsen/NetRatings.

There has been limited research into the relevance of sponsored links for Web queries. Jansen and Resnick [3] state that searchers are biased against sponsored links. Jansen and Molina [2] In this paper, results are presented from the examination of several thousand sponsored and non-sponsored links from the three major search engines in response to 108 ecommerce queries.

3. RESEARCH QUESTIONS

Sponsored search is the dominant business model for Web search engines, generating billions in yearly revenue. However, are sponsored links providing online consumers with relevant choices for products and services? Are sponsored links more relevant than non-sponsored links for Web ecommerce queries?

H: *Sponsored links are more relevant than non-sponsored links for Web ecommerce queries.*

Given that sponsored search is primarily a method for commercial entities, it would seem reasonable that sponsored links should be more relevant than non-sponsored links for ecommerce queries. In fact, for sponsored search to be a workable long-term business model, it is critical that the relevance be at least competitive with non-sponsored links.

4. RESEARCH METHODOLOGY

Using popular ecommerce terms from WordTracker, we extracted 108 queries from an AltaVista search transaction log. Examples

include *Where can I buy Diesel shoes online?*, *wholesale car prices*, *buying digital cameras*, and *discount tickets Broadway*. We submitted the 108 ecommerce queries to the three major search engines (i.e., Google, MSN Search, and Yahoo!) and retrieved the results. We first captured two search engine results pages (SERPs), which generally accounts for 80% of Web searchers' results page viewing [4, 7] for each query on each search engine. The total number of search links returned from all three Web search engines was 8,256.

We removed the Search Engine, Sponsored, Location, and Rank fields. We also removed all duplicate links for each query, as we did not want evaluators to review the same URL more than once for a particular query to ensure consistence among evaluations. With duplicates removed, we had 6,162 unique records containing the fields of Query, Title, Summary, and URL. We then designed an application that displayed each result individually with the corresponding query. We had three evaluators independently rank each result for relevance based on whether, for the given query, the result is relevant (1), somewhat relevant (2), or non-relevant (3). When the three users had evaluated all of the search results, we combined and averaged the ratings for each of the search results. Cronbach's alpha [1] among the evaluators was 0.61.

For our evaluators, there was agreement among all three evaluators for 2,077 links (25%), agreement among two evaluators on 4,826 links (58%). Given the wide personal variations that can occur with rating documents as relevant or not, we see this agreement as quite high.

We re-integrated the Search Engine, Rank, Type, and Location with the judges' evaluations. For the non-unique results, we reintroduced these records and automatically assigned the corresponding average evaluation. In order to make the means more intuitive, we transposed the evaluation results for relevant and non-relevant links (i.e., a rating of 1 – relevant was recoded to a 3 – relevant and a rating of 3 – not relevant was recoded to a 1 – not relevant). With this recoding, the higher mean score indicates a more relevant set of results. Once we had completed this, we had an evaluation between 1 and 3 for each of the 8,256 results. We exported this tabulation from our database to a spreadsheet and then imported the data into SPSS 12.0 for the statistical analysis.

5. RESULTS

H: *Sponsored links are more relevant than non-sponsored links for Web ecommerce queries.*

In order to evaluate the hypothesis, we performed a statistical evaluation to determine if there is a difference of relevancy means among the two types of Web links (non-sponsored and sponsored) tested. We used a one-way ANOVA statistical analysis to compare means and variance between the groups. The ANOVA analysis tests the null hypothesis whether group means differ. The results indicate that there is a significant difference among the groups ($F(1) = 413.77$, $p<0.01$; the critical value of F = *2.37*).

From Table 1, we see that the average relevance rating of the non-sponsored links (1.69) was statistically significantly lower than the average relevance rating for the sponsored links (1.93). Therefore, we fail to reject that hypothesis that *sponsored links are more relevant than non-sponsored links for Web ecommerce queries.*

Table 1. ANOVA Descriptives for the Research Hypothesis

	N	Mean	Std. Deviation	Std. Error	Min	Max
Non-sponsored	5,639	1.69	.47	.006	1.0	3.0
Sponsored	2,617	1.93	.57	.011	1.0	3.0
Total	8,256	1.77	.52	.006	1.0	3.0

6. DISCUSSION

Our results show statistically that sponsored links are more relevant than non-sponsored links based on user evaluation of SERP snippets composed of Title, Summary, and URL. We found this somewhat surprising given the negative bias that Web searchers appear to have concerning sponsored links [3, 5]. Overall, this is a good sign for the long term success of sponsored search as viable business model. With only about 30% of searchers presently interacting with sponsored links [3], there is a potential for substantial growth given that these sponsored links appear to be providing relevant content. However, our sample of queries was ecommerce-related, and sponsored search is designed with this domain in mind. It would interesting to see if this comparison held in other domains outside of ecommerce, such as health or education.

The research results have implications for Web search engines and implementation of the sponsored search model. It appears that sponsored search is an effective method for providing relevant information to Web searchers, as if not more effective than algorithm methods. Our results show this for the ecommerce area. For future research, we plan to analyzes this data to evaluate top and side sponsored results, correlation of relevance and ranking of results, and factors relating the relevance judgments.

7. REFERENCES

[1] Cronbach, L. J., "Coefficient alpha and the internal structure of tests," *Psychometrika,* vol. 16, pp. 297-334, 1951.

[2] Jansen, B. J. and Molina, P., "The effectiveness of Web search engines for retrieving relevant ecommerce links," *Information Processing & Management,* vol. 42, pp. 1075-1098, 2006.

[3] Jansen, B. J. and Resnick, M., "An examination of searchers' perceptions of non-sponsored and sponsored links during ecommerce Web searching," *Journal of the American Society for Information Science and Technology,* vol. 57, pp. 1949-1961, 2006.

[4] Jansen, B. J., Spink, A., and Saracevic, T., "Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web," *Information Processing & Management,* vol. 36, pp. 207-227, 2000.

[5] Marable, L., "False Oracles: Consumer Reaction to Learning the Truth About How Search Engines Work, Results of an Ethnographic Study," Consumer WebWatch, 2003, pp. 1-66.

[6] McCarthy, T., "Yahoo! Goes to Hollywood," *Time,* vol. 165, pp. 50-53, 21 March 2005.

[7] Silverstein, C., Henzinger, M., Marais, H., and Moricz, M., "Analysis of a Very Large Web Search Engine Query Log," *SIGIR Forum,* vol. 33, pp. 6-12, 1999.