The Gender-Brand Effect of Key Phrases on User Clicks in Sponsored Search

Partha Mukherjee
Pennsylvania State University.
College of Information Science and Technology.
University Park, PA-16802, USA
pom5109@ist.psu.edu

Bernard J. Jansen
Pennsylvania State University.
College of Information Science and Technology.
University Park, PA-16802, USA
jjansen@acm.org

Abstract
In this research, we analyze the relationship among (1) the numbers of clicks generated by the key-phrases users provide as queries in sponsored search, (2) the interaction between the gender orientation of those key-phrases, and (3) the occurrence of branded terms in those queries. The aim of this research is increased personalization of search engine results for branded queries targeting a specific gender. This will improve the consumer's online searching experience and potential interest in ads with branding focus. Increased personalization of search engine results and enhancement of users' interest in branded ads will also increase the revenue and profit of the advertisers. Our data consists of 7 million daily records from a keyword advertising campaign from a major US retailer. We segregate the key-phrases into six different categories. The gender orientation of terms has male, female and neutral categories each with two groups: - branded and unbranded phrases. Using one way Analysis of Variance, we analyze the effect of gender orientation of both branded and unbranded key-phrases on number of users' clicks on the queries submitted to web search engine. The result shows that the means of users' clicks significantly vary among the categories. Moreover the result shows that the group formed by combination of female gender with branded terms generates maximum number of clicks among all six categories. Females may be more attracted to the use of branded terms due to the positive attitude and customer loyalty generated by brand image.

CHI'13, April 27 – May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2.

Author Keywords
Search Engine Marketing; ANOVA; Gender Orientation; Brand Effect

Figure 1: Snapshot of Google SERP

ACM Classification Keywords
J.4[Social and Behavioral Sciences]: Economics

General Terms
Human Factors, Economics, Experimentation

Introduction

The online economy is driven by search queries, and the advertisements served by the search engines in response to these queries can be quite profitable for online business. To leverage the web as the medium of commerce, more retailers with brick and mortar stores have gone online and the numbers of online-only websites has grown significantly. To thrive in their business, these businesses need to optimize their search engine advertising strategies based on key-phrases. Two major issues marketers face to optimize their marketing strategies are understanding the search behavior of the consumers [6] and the attitude of the consumers towards the brand. Prior research on web demographics and the role of branding in sponsored search [5, 9] illustrates the increased attention in addressing these issues.

In this research, we use a real time search engine marketing (SEM) data of a major US retailer to analyze the relationship between online consumer behavior and the gender specific searches of terms in key-phrases with a branding and non-branding focus. The searching behavior of the consumer is represented by the number of clicks as the modern keyword advertising relies on Pay-Per-Click (PPC) model. When the user clicks on a hyperlink shown in a search engine result page (SERP), the corresponding advertiser pays the web search engine for each click. Figure 1 shows a snapshot of SERP of the Google search engine.

The research motivation is to determine whether the combination of gender orientation of the key-phrases mentioning brand names has an effect on the number of clicks from the users per key-phrases. This information will help retailers in specific gender-brand targeting for their marketing strategies.

In the next section, we provide a brief literature review that studies the research regarding the effect of branding and gender orientation on the online searching behavior of the users. The following sections present research questions, data, results, discussions with implications respectively.

Literature Review

Prior research has studied the brand effects as the antecedents of online trust relating to the company, its website and the product [11] . However, limited research has been performed measuring the effect of key-phrases with branded terms on the performance metrics of sponsored search. Ghose and Young [3, 4] reports that brand terms have an effect on SEM performance where retailer specific brands achieve high click-through-rates (CTR). Jansen, Sobel and Zhang [7] investigate the influence of brand effect on SEM performance and infer that the brand term in key-phrases has a dramatic effect in performance of sponsored search campaign. The branded term is the textual representation of the brand.

Regarding the research on the effect of gender on information processing behavior on the web, studies [2, 8] find significant behavioral gender differences. To measure the effect of gender orientation of queries on the sponsored search performance Jansen, Moore and Carman [5] probabilistically classified the key-phrases

The brief description of the remaining performance metrics for sponsored search is given below:

1. Impression: Response of search engine shown in SERP against a user query

2. Conversion: Online purchase made by consumer

3. Cost-per-click (CPC): The amount billed by search engine to an ad agent for each user click

4. Sales revenue: Revenue generated by the advertiser by selling the products/items online

5. Orders: The number of orders from the advertisement for that day for a given key phrase

6. Items: Number of items purchased from that advertisement on that day for a given key phrase from all orders. One order could have one or more items.

of a real time campaign into specific gender orientation using Microsoft adCenter Labs Demographics Prediction Tool. The authors infer the gender of the searchers based on the search query. Their result shows that gender specification of key-phrases has significant effect on SEM performance and identifies the importance of personalization of web results.

Though the aforementioned research [5, 7] discuss the effect of gender orientation and brand effect on the users' online searching behavior, they do not investigate the interaction effect of the combination of gender and branding focus of terms in key-phrases on sponsored search campaigns. This is important to know as the results help retailers to identify whether the gender difference with branding focus affect advertising cost and profits by means of increased personalization of the web search results.

Research Question

Our research question is: *Is there any significant difference in the CTR of sponsored search results based on gender differences of branded and unbranded terms in key-phrases?*

To investigate our research question, we created two classifications of key-phrases: *Branded key-phrases* and *Unbranded key-phrases.* The *branded key-phrases* are associated with the textual representation of brand names while the unbranded ones are key-phrases without accompanying brand names. In addition to the brand related classes, we have generated three different categories based on the gender orientation of the key-phrases:- *female key-phrases, male key-phrases* and *neutral key-phrases* re-using the data generated in prior research [5]. Taking the union of the categories formed on brand effect and gender difference, we have six different gender-brand classes of key-phrases. We are evaluating the effect of these six categories on the number of users' clicks. The goal of most SEM campaigns is to get clicks on the hyperlinks of ads shown in SERP. The clicks identify the critical user behavior showing potential interest in search engine results.

As positive brand image has a dramatic effect on consumers' interests in advertisements and purchasing behavior [7], we believe that the gender categories mentioning brand name will result in more clicks than the gender-unbranded counterparts. Moreover from the prior work [10], we observe that female consumers prefer lesser purchasing risk. Therefore, it leads us to assume that female_branded category will generate more clicks than the other gender categories. Based on the research question and the stated assumptions, we develop the following hypothesis:

Hypothesis 01: There will be a significant difference in the number of clicks based on specific gender orientation of branded and unbranded key-phrases.

There are other metrics (e.g., impression, conversion, sales revenue, cost-per-click, orders, items etc.) that measure the sponsored search campaign performance. Due to space constraint we limit our study to evaluate the effect of the combined gender-brand difference on the critical metric of users' clicks.

Data

We have used a rich dataset that contains records of the sponsored search advertisement efforts by a major US retailer during a 33-month period, spanning 4 calendar years, from 30 September 2005 to 09 June

Key-phrases	Count
Branded	2655
Unbranded	36917

Table 1. Actual count of branded and unbranded key-phrases

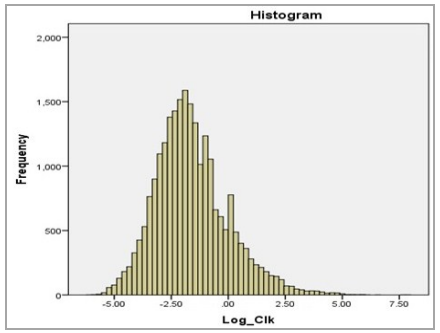

Figure 2: Histogram plot of the log transformed data

Gender of user	Count
Male	10860
Neutral	19495
Female	9217

Table 2. Actual count of different gender category of users

2008. There are approximately 7 million daily records from nearly 40,000 key phrases. There is a unique record for each key phrase for a given day. Table 1 shows the actual count of branded and unbranded key-phrases while Table 2 displays the gender count.

The legend of the data log used in this research is provided in Table 3. From Table 3 we constitute six classes of key-phrases: female_branded (F_B), female_unbranded (F_UB), male_branded (M_B), male_unbranded (M_UB), neutral_branded (N_B) and neutral_unbranded (N_UB).

Attribute	Description
Ad	Identified with unique id
Key-phrase	Key-phrase that triggered the advertisement
Average Clicks	Number of user clicks averaged over the days grouped by distinct ad
Classification	Gender orientation of the key-phrases
Client Brand Phrase	=0 for unbranded key-phrases. >0 for branded key-phrases

Table 3. Fields and descriptors of the data log

Methodology and Results

Once six categories were constructed, we imported the data into SPSS. In SPSS, we aimed to test our hypothesis using one way analysis of variance (ANOVA) procedure among six groups to test the differences between the means of clicks among the six categories. However, our data follows the power law distribution and hence is not multivariate normal. We used the log transformation *log(variable+0.5)* to make the data normal before conducting the ANOVA test. The data is successfully normalized by means of log transformation as shown in Figure 2. The approximate straight line representation of normal Q-Q plot shown in Figure 3 also establishes the view of normality of transformed data.

The result of ANOVA test shows that there is a significant interaction between number of clicks and gender-brand classes ($F(5) = 88.1$ with $p\text{-value} \approx 0$) as shown in Table 4. We observe that there is at least one category which is significantly different from other categories in terms of number of clicks. The result in Table 4 supports our research hypothesis.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1115.43	5	223.085	88.13	0.00
Within Groups	59430.03	23478	2.531		
Total	60545.45	23478			

Table 4. The result of ANOVA test

We use the Games-Howell test with the evidence of non-homogeneity of variances for post hoc analysis. Table 5 provides the evidence of non-homogeneity of variances. The significance level of Levene's statistic should be greater than 0.05 for satisfying the condition of homogeneity of variances. So, we follow the test of equality of means where the Welch statistic value is significant as observed in Table 6. Our data follows the equality of means assumption.

The result of post hoc analysis is presented in Table 7. It is seen from the magnitude of reported T-values that female_branded and neutral_branded categories separately have a significant difference in number of clicks with all of the remaining categories.

Figure 5 shows the means of clicks generated in six categories in the perspective of original data. It is

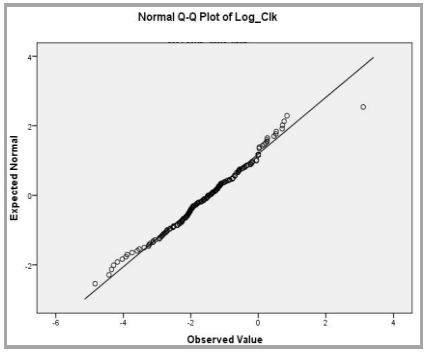

Figure 3: Normal Q-Q plot of transformed data

Levene Statistic	df1	df2	Sig.
54.234	5	23478	0.00

Table 5. Non-Homogeneity of Variances test

Welch Statistic	df1	df2	Sig.
76.187	5	824.96	0.00

Table 6. Equality of means test

worth noticeable that female_branded class will generate most number of clicks. This could indicate that branded key phrases are much more focused and worthwhile targets for attracting potential female online consumers to a Web site.

	M_B	M_UB	N_B	N_UB	F_B	F_UB
M_B	x	2.76	5.08*	0.52	7.6*	0.86
M_UB	2.76	x	16.2*	10.8*	12.3*	8.3*
N_B	5.08*	16.2*	x	11.6*	5.13*	12.1*
N_UB	0.52	10.8*	11.6*	x	10.4*	1.58
F_B	7.6*	12.3*	5.13*	10.4*	X	10.7*
F_UB	0.86	8.3*	12.1*	1.58	10.7*	X

Table 7. Magnitude of T-values between different categories. The significance among the categories are shown by asterisk (*) notation

Discussion and Implications

In this research, we investigate the combined effect of gender orientation and the branding focus of the key-phrases on user's behavior in search engine marketing campaigns. The study with six different gender-brand key-phrase categories highlights that brand focusing of key-phrases produces more clicks with female and neutral gender orientation.

The gender orientation of key-phrases is investigated as an aspect of human information processing and personalization [5]. Intuitively, growth in personalized results leads to increase in marketer's revenue.

This research shows that key-phrases with more personalization (female) offers better result if the key-phrases are focused with brand effect unlike the previous study [5] done without considering branding focus. This in turn also implies that women shoppers might perceive less risk for branded advertisements. Another major finding is that male consumers may not differentiate between the branded and unbranded products unlike female counterparts as there is no significant difference (i.e. T-value = 2.76) observed between M_B and M_UB categories in Table 7.

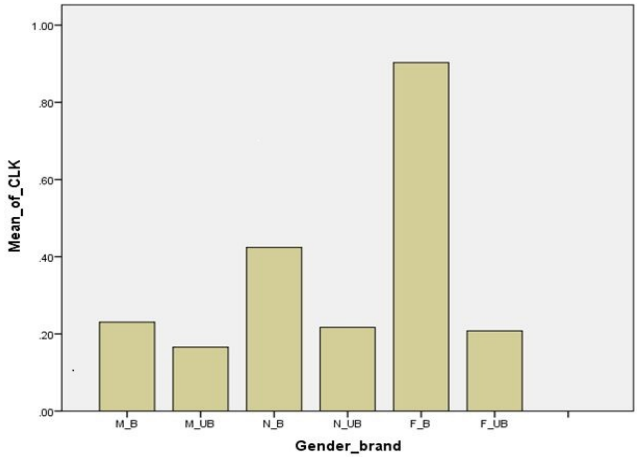

Figure 5: Bar chart representing means of clicks in each category

From the perspective of human information processing, our result highlighting the differences between male and females in terms of the information gathering and processing behavior. Our result shows that in sponsored search, female users are more inclined to click on the branded advertisements relative to males. The theory of social categories [1] indicates that female gender orientated phrases perform better with branding focus (i.e., in less risk perceiving shopping environment). But there is a limitation in cultural bias as the trend depicted by our result is applicable to US searchers only.

Regarding practical significance, results show that branded phrase with female gender orientation

generates higher mean clicks. This implies that to gender target key-phrases, a retailer should include the relatively niched female oriented key phrases common to branded queries along with the generic phrases for unbranded queries in keyword advertising campaigns.

Conclusion

The results of this research conclude that branded key-phrases with female gender orientation generates more clicks. This helps the retailers to personalize the web search results to improve the overall searching experience. It is beneficial for retailers to devote resources, create ad recommendations, and leverage the brand image or reformulate the marketing strategies targeting the female gender oriented key-phrases.

We restrict the combined effect of brand and gender orientation to CTR, but there are other metrics such as impressions, CPC, sales revenue, conversions, orders, items etc. to measure the performance of user's behavior in sponsored search. In our future work, we will investigate the combined gender-brand effect on these remaining metrics.

Acknowledgement

Our thanks to The Rimm-Kaufman Group, particularly George Michie, for providing data for this study. We also acknowledge and thank Alan Rimm-Kaufman, who departed this world on 18 July 2009.

References

[1] Deaux, K. From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist, 39, 2* (1984), 105-116.

[2] Garbarino, E. and Strahilevitz, M. Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation. *Journal of Business Research*, *57, 7* (2004), 768-775.

[3] Ghose, A. and Yang, S. An empirical analysis of paid placement in online advertising. in *Proc. ICEC 2007,* ACM Press (2007).

[4] Ghose, A. and Yang, S. An empirical analysis of sponsored search performance in search engine advertising. in *Proc. WSDM 2008*, ACM Press (2008).

[5] Jansen, B.J., Moore, K., and Carman, S. Evaluating the performance of demographic targeting using gender in sponsored search. *Information Processing & Management, 49, 1* (2012), *286-302*.

[6] Jansen, B.J. and Resnick, M., An examination of searcher's perceptions of nonsponsored and sponsored links during ecommerce Web searching. *Journal of the American Society for Information Science and Technology, 57, 14,* (2006), 1949-1961.

[7] Jansen, B.J., Sobel, K., and Zhang, M. The brand effect of key phrases and advertisements in sponsored search. *International Journal of Electronic Commerce,* 16*, 1,* (2011), 77-106.

[8] Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., and Gay, G. The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management, 42, 4,* (2006), 1123-1131.

[9] Rodgers, S. and Harris, M.A. Gender and e-commerce: an exploratory study. *Journal of Advertising Research, 43, 3,* (2003), 322-329

[10] Sanchez-Franco, M.J. Exploring the influence of gender on the web usage via partial least squares. *Behaviour & Information Technology, 25, 1,* (2006), 19-36.

[11] Schultz, C.D. *Consumer trust in E-Commerce: an analysis of means communicating trustworthiness from a buying transaction life cycle perspective*, Verlag Dr. Kovac, Skövde, 2007.