

Gender-Brand Effect of Search Queries on Sponsored Search Performance

Partha Mukherjee

Pennsylvania State University
College of Information Sciences and Technology
University Park, Pennsylvania 16802
pom5109@ist.psu.edu

Bernard J. Jansen

Pennsylvania State University
College of Information Sciences and Technology
University Park, Pennsylvania 16802
jjansen@acm.org

ABSTRACT

In this research, we analyze the relationship among (1) the performance metrics of key phrases used in a sponsored search campaign, (2) the gender orientation of those key phrases, and (3) the occurrence of branded terms in those search queries. The aim of this research is increased personalization of search engine results in order to improve the consumer's online searching experience and potential interest in ads with branding focus. We segregate the key phrases into six different categories combining gender orientation with branding focus of the key phrases from a rich dataset that consists of 7 million daily records of a real time sponsored search campaign. The result of ANOVA shows that the means of sponsored search metrics significantly vary among the categories and exhibits the female's preference over branded terms.

Keywords

search engine marketing, ANOVA, gender orientation, brand effect.

INTRODUCTION

In modern business to consumer e-commerce environments web has become the popular medium for marketing (Constantinides, 2002). In online advertising, the online shopping behavior of the consumers, characterized by information seeking and searching, enables them to satisfy various needs of their everyday life. Keyword advertising (also referred to as sponsored search or search engine marketing (SEM)) has become the most mature form of online advertising (Evans, 2008). Much of this online economy is driven by search queries submitted by the consumers, and the advertisements published in search engine result page (SERP) in response to these queries can be quite profitable for online businesses. Retailers optimize their search engine marketing strategies based on key phrases to understand the search behavior of the consumers.

The online shopping behavior of the consumer is represented by metrics (i.e., number of impressions, the

number of user clicks etc.) that determines the performance of sponsored search. The motivation of this research is to determine whether the combination of gender orientation and brand mention of the key phrases has an effect on these performance metrics per key phrases. This information will help retailers with specific gender-brand targeting for their marketing strategies.

LITERATURE REVIEW

Limited research has been performed measuring the effect of key phrases with branded terms on the performance metrics of sponsored search. Several studies (Ghose and Yang, 2007) report that brand terms have an effect on SEM performance where retailer specific brands achieve high click-through-rates (CTR). Researchers (Jansen, Sobel & Zhang, 2011) have investigated the influence of brand effect on SEM performance and show that combination of branded phrase and branded advertisement generates 15 times more sales revenue than any other combination. In their study, the branded term is the textual representation of the brand name. Brand name is the label that identifies the retailer or retailer's products distinct from those of other retailers.

Regarding the research on the effect of gender on information processing behavior on the web, Sanchez-Franco (2006) identifies that attitude plays a strong role for women's shopping. Jansen, Moore and Carman (2013) carried out research on demographic targeting measure the effect of gender orientation of queries on the sponsored search performance. The researchers probabilistically classify the key phrases of a real time campaign into specific gender orientation using Microsoft adCenter Labs Demographics Prediction Tool. The gender of the searchers is inferred 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.

There are no such prior works that combine the gender specification and brand effect of key phrases on user's behavior. The lack of research leads to the emergence of the motivating question: *What will be the effect of branded and unbranded terms in key phrases with gender differences on the sponsored search performance?*

RESEARCH QUESTION

Our research question is: *Are there significant difference in sponsored search results based on the interaction of the gender oriented and the branded/unbranded terms of search queries?*

To investigate our research question, we created two classifications of key phrases:

Branded key phrases: The key phrases that are associated with the textual representation of brand name.

Unbranded key phrases: The key phrases that do not mention the 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*. Taking the union of the categories formed on brand effect and gender difference, we ultimately have six different 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)*. We are evaluating the effect of these six categories on different performance metrics of sponsored search result that identify the critical user behavior. The metrics investigated in this research are: the number of impressions (i.e., search engine response published in SERP for a query), the number of user clicks (i.e. user clicks on ad hyperlink, published in SERP), the revenue generated by online sales of the products, number of orders placed online against a key phrase and cost-per-click (CPC) (i.e., amount billed by search engine to an ad agent for each user click).

The dramatic effect of positive brand image on consumers' purchasing behavior (Jansen, Sobel & Zhang, 2011), and female consumer's preference of less purchasing risk (Sanchez-Franco, 2006), lead us to assume that *F_B* category will generate more interactions than the other gender categories. Based on the research question and the stated assumption, we develop the following hypotheses on the performance metrics:

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

Hypothesis 02: There will be a significant difference in CTR based on specific gender orientation of branded and unbranded key phrases.

Hypothesis 03: There will be a significant difference in the cost per click based on specific gender orientation of branded and unbranded key phrases.

Hypothesis 04: There will be a significant difference in the average of sales based on specific gender orientation of branded and unbranded key phrases.

Hypothesis 05: There will be a significant difference in the average number of orders based on specific gender orientation of branded and unbranded key phrases.

DATA

The data log used for this research contains daily information on a sponsored search campaign from a large nationwide retail chain that has both a brick and mortar and an online sales presence. The data is a record of the sponsored search advertisement efforts by the company during a 33-month period, spanning 4 calendar years, from 30 September 2005 to 09 June 2008. There are approximately 7 million unique daily records from 36,917 unbranded and 2,655 branded key phrases. Brand classification of key phrases is done based on *client_brand_phrase* attribute of the data set. From gender orientation perspective, we classify 10860 male, 9217 female and 19495 gender neutral key phrases using Microsoft adCenter Labs Demographics Prediction Tool.

METHODOLOGY

Once the six gender-brand categories are constructed, the data is imported into SPSS data view. Our data follows the power law distribution and hence is not multivariate normal. To run ANOVA, we transformed the data via the Box-Cox transformation (Box and Cox, 1964) by using log transformation function $\log(\text{variable} + 0.5)$. The transformation successfully normalizes the data with a bit of skewness on the left as the data is weighed toward lower number of clicks, sales, impressions etc. Despite the existing skewness, ANOVA is proved robust with the data even if it deviates from normality (Box and Andeason, 1955). In SPSS, we run the ANOVA procedure to test the differences between the means of the metrics among the six categories.

RESULTS

The results section contains the evaluation of hypotheses to find the significance among the gender-brand categories.

Hypothesis Testing

To evaluate of our six hypotheses, we use one way ANOVA to compare the means between gender-brand categories. The critical value of ANOVA F-statistic for each metric is 2.214 at 95% confidence interval.

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

The result indicates that there is a significant difference of means of impressions between the gender-brand categories of key phrases ($F(5) = 29.35, p < 0.05$). So hypothesis 01 is fully supported.

From Figure 1 it is found that *F_B* category generates more impressions than the remaining categories.

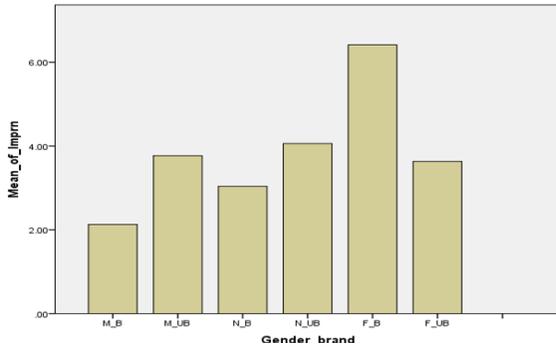


Figure 1. Mean of impressions in each category

Hypothesis 02: There will be a significant difference in CTR based on specific gender orientation of branded and unbranded key phrases.

The result indicates that there is a significant difference of means of clicks between the gender-brand categories of key phrases ($F(5) = 88.13, p < 0.05$). Hypothesis 02 is fully supported.

Similar to Figure 1, it is observed from Figure 2 that F_B category generates more clicks than other categories.

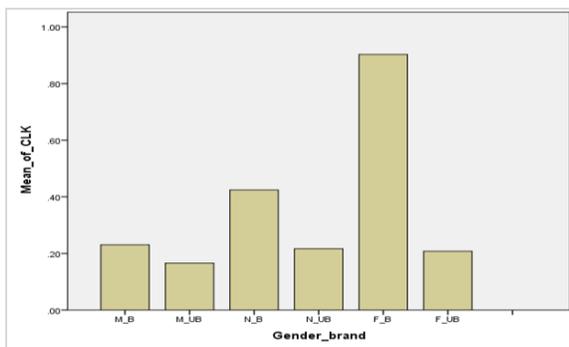


Figure 2. Mean of clicks in each category

Hypothesis 03: There will be a significant difference in the cost per click based on specific gender orientation of branded and unbranded key phrases.

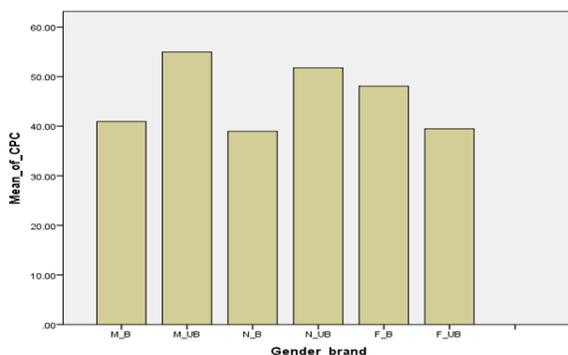


Figure 3. Mean of CPC in each category

The result indicates that there is a significant difference of means of CPC between the gender-brand categories of key phrases ($F(5) = 74.92, p < 0.05$). Hypothesis 03 is fully supported.

From Figure 3 it is found that M_UB category generates more average CPC than other categories. This would indicate that advertisers value the searchers performing web search with *male-unbranded* and *neutral-unbranded* key phrases more as potential customers.

Hypothesis 04: There will be a significant difference in the average of sales based on specific gender orientation of branded and unbranded key phrases.

The result indicates that there is a significant difference of means of sales between the gender-brand categories of key phrases ($F(5) = 80.59, p < 0.05$). Hypothesis 04 is fully supported.

Figure 4 displays the means of sales revenue generated among six categories. It is found that F_B category generates more revenue than the remaining categories. Combined with impressions and clicks, it is found that females are more inclined to purchase branded products.

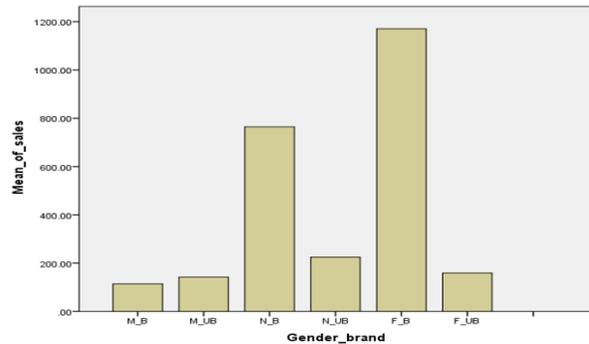


Figure 4. Mean of sales revenue in each category

Hypothesis 05: There will be a significant difference in the average number of orders based on specific gender orientation of branded and unbranded key phrases.

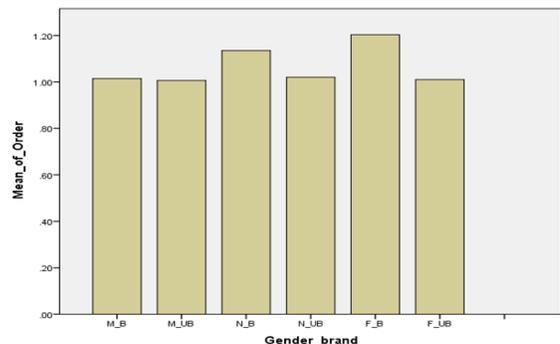


Figure 5. Mean of orders placed in each category

The result indicates that there is a significant difference of means of orders placed between the gender-brand categories of key phrases ($F(5) = 167.13, p < 0.05$). Hypothesis 05 is fully supported.

From Figure 5 it is found that F_B category generates highest number of orders followed by the N_B one. The average orders placed by unbranded key phrases are similar in nature irrespective of gender of searchers.

DISCUSSION AND IMPLICATIONS

Discussion of Results

In this research, we investigate the gender orientation of key phrases with branding focus from the perspective of personalization of web results and human information processing. 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 (Jansen, Moore & Carman, 2013) 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.

From Figure 1, Figure 2, Figure 4 and Figure 5, female searchers engage themselves more as potential customers in online marketing of items with branding focus than males. However this perception is not reflected in CPC. From Figure 3, the F_B category is cheaper than the M_UB and N_UB groups. Advertisers bid more on these unbranded key phrases to draw more potential customers towards them. However, it is not justified, as branded key phrases with female gender generate higher sales revenue.

Theoretical Implications

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 buy the branded products relative to males. The finding is consistent with the theory of social categories (Deaux, 1984) in the sense that brand presence helps females to perceive online shopping environment less precarious.

Practical Implications

Regarding practical significance, results show that branded phrase with female gender orientation performs better across the sponsored search metrics. The result implies that to gender target key-phrases for generating higher sales and profit, 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 regarding evaluation of online advertisement performance, reported in this research indicate that combination of female gender orientation and brand focus for the key phrases generates higher sales and are relatively cheaper than other unbranded combinations. We believe that the outcome observed from the result marks an important step in empirical research for demographic targeting and managing positive brand reputation in sponsored search.

For future work, we like to perform an additional analysis on effect size to ensure that the observed differences between the gender-brand groupings are worth meaningful. We further want to investigate the correlation of personalized web results combined with fluctuation in brand focus to sponsored search performance over time.

ACKNOWLEDGEMENT

Our thanks to The Rimm-Kaufman Group, particularly George Michie, for providing data for this study.

REFERENCES

- Box, G. E., Cox, D. R., 1964. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 211-252.
- Box, G. E. P., Andeesson, S., 1955. Permutation theory in the derivation of robust criteria and the study of departures from assumption. *Journal of the Royal Statistical Society (Series B)*, 17(1), 1-26.
- Constantinides, E., 2002. The 4S web-marketing mix model. *Electronic Commerce Research and Applications*, 1(1), 57-76.
- Deaux, K., 1984. From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist*, 39(2), 105.
- Evans, D. S., 2008. The economics of the online advertising industry. *Review of network economics*, 7(3).
- Ghose, A., Yang, S., 2007. An empirical analysis of paid placement in online advertising. In *Proceedings of the ninth international conference on Electronic commerce (ICEC'07)*.
- Jansen, B. J., Moore, K., Carman, S., 2013. Evaluating the performance of demographic targeting using gender in sponsored search. *Information Processing & Management*, 49(1), 286-302.
- Jansen, B. J., Sobel, K., Zhang, M., 2011. The Brand Effect of Key Phrases and Advertisements in Sponsored Search. *International Journal of Electronic Commerce*, 16(1), 77-106.
- Sanchez-Franco, M. J., 2006. Exploring the influence of gender on the web usage via partial least squares. *Behaviour & Information Technology*, 25(1), 19-36.