A Study on Factors Influencing the Click Through Rate, CTR, of Internet ads by the Charging Mode of CPM (Cost Per Mile) Based on Information Transfer

Xingjun Shi

School of Economy & Trade, Zhejiang Industry & Trade Vocational College, Wenzhou, Zhejiang, China

Abstract — CTR (Click Through Rate) is generally considered as a core indicator in measuring the effect and the ROI (Return On Investment) of Internet ads by the charging mode of CPM (Cost Per Mile). By utilizing relevant theories and research methodologies of modern informatics, the study aims to determine the characteristics of promotion that information is having and the impact on CTR by means of mathematical inference. The functional relationship between them is further explored by quantitative analysis through practical efficiency testing, data statistics and regression analysis on the Internet. The conclusion of the research could be a reference or an instruction for advertisers in designing their Internet ads, so as to improve the transfer efficiency of promotion information and raise their CTR and eventually the ROI.

Keywords - Information Transfer, Cost Per Mile (CPM), Internet ads , Influencing factors

I. INTRODUCTION

Internet ads, yet without an authority definition, is usually referred to as advanced advertising operation models that advertising on Internet in forms of specific ad bars, hyperlink and multimedia representations. There are 4 essential elements of internet advertising: operators, information, media and charging mode. Operators include advertisers (referred to corporations or some economic organizations that design, produce or publish ads for the purpose of promotion), media agents (the bailees of the advertisers), and the audiences who receive ad information willingly or unwillingly. The information refers to what the advertisers intend to express in ads and transfer to the audiences. Media is the medium through which ads are presented and delivered. Charging mode fixes the charge standard in the contract between the advertisers and media agents.

There are various charging modes of Internet ads at present, including CPM(Cost Per Mile), CPC(Cost Per Click), CPA(Cost Per Action), CPS(Cost Per Sales), CPR(Cost Per Response), etc. They differ in product features, emphasis, and applicability, of which CPM is one of the most popular charging way of network media buying currently.

CPM is the average cost of per 1000 audiences seeing or hearing the ads during the advertising. Using this charging mode advertisers usually face higher risks than the website media because the costs are consequently required once the ads are showed. Have the visitors noticed it? Are they interested in it? Have they paid their attentions? These are the points advertisers actually care but nothing to do with their payments. Therefore, using this charging mode, the key to raise the ROI (Return on Investment) lies in the CTR (Click through Rate) of their ads.

According to incomplete statistics, the CTR of Internet ads is usually kept at a range from 0.3 to 0.8% at present, and the actual transaction rate is less than 5% of the CTR. That is to say, per customer transaction, the advertising cost may probably reach a sizable percentage according to the charge standard of the mainstream website media.

Therefore, how to raise CTR? There are many visible and invisible influencing factors. Since the scale, the position, and the performance of the ads can hardly be improved without increasing their investment, one of the most effective ways is to focus on the inherent information the ads intend to express, such as to improve the creativity of the ads constantly, making them more and more attractive and effective.

II. RELATED RESEARCH

There have been already a few relevant works about Internet ads on Influencing factors of their CTR. Here we take some typical ones for instance. Wegert found that banner ads like TV ads play an impactful role when they are well-designed. In other words, excellent banner ads can easily draw much more attentions and bring higher CTR than corny ones. The impact factors were divided by Prof. Huang into 2 categories, the ones from Internet ads and the ones from web visitors. Shay-shay testified to the impact on CTR made by 4 core characteristics: the interaction, the vitality, the coerciveness and the repetitiveness. In Rettie, Robinson and Jenner’s opinion, Internet ads is increasing by leaps and bounds, leading to fiercer competition and negative reactions from audiences. From Guo Yan’s research, the content, the scale, the position, the interaction and the performance of Internet ads were found out as 5 key influencing factors of CTR.

The works above are mainly about the direct performances of ads. While judging from the response of the audiences comes to different conclusions. Mehta found that people with positive attitudes to the Internet ads are much more easily to be influenced by them. Mullaneysh, a
famous advertising professor, pointed out that most of web visitors who have noticed Internet ads, were those who trying to find something stimulated, such as trial products or free samples. Rodgers and Thorson believed that the attitudes visitors towards the Internet ads could be a vital impact factor of their effects. More than 800 experimenters were gathered by Dreze and Hussherr, 2 network market researchers, to test their impressions towards the 10 ads from 10 different brands they had watched one by one. And finally they certified the brand recognition normally plays an important part in the effect. In another test, Todgers focused on attentions the audients towards the ads by testing Internet users with different motivations. Then he testified the importance of the motivations. For example, one searches the information he or she want purposively will pay more attentions on ads than other Internet surfers. After an empirical research, Ren and Song found that there are 2 significant impact factors of Internet ads, the richness of the network resources and the web habits of visitors. Simultaneously, they confirmed that the CTR is in proportion to the ideas the Internet ads embody, as well as their transmission ways.

Among all the statistic and analysis of current research statuses, literature search and questionnaire survey are found as 2 essential methods about CTR, most of which are qualitative and generalized. A few empirical researches are found in them, mainly about analysis of impact factor, and validation test by regression analysis and then the weight of indicators calculated. However, few specific researches are found about the charging mode of CPM, and moreover, empirical researches focusing on information transfer with statistical analysis can be hardly found. That is exactly what this paper aims to do.

III. PROCESS OF RESEARCH

A. Sample

In consideration of the variety, complexity, and diversity of Internet advertising, in order to reduce the obstruction and noise interference as possible as we can, a plain text ad is taken as a sample. It shows “Click Here to Watch HD Hollywood blockbusters for Free” of which most unnecessary information such as the font, the display form, and the color are ignored. Based on the value and the role played in an ad, information included in an ad can be divided into 3 parts, product information, promotion information, and supporting information. In above ad, words and phrases like “Watch” and “Hollywood” that used for the introduction of the product or service, are called product information; those like “HD” and “Free”, used for embellishment, or to strengthening its attractiveness, are classified as promotion information; and those like “here”, “to”, “for”, the hyperlink as well as the bound, merely used for display and dissemination, but of little help to promotion or introduction, are divided as supporting information.

Among this 3 parts, the impact of product information basically lies on the rational demands of visitors, and in the meantime, supporting information is insignificant. Therefore, promotion information plays a decisive role here, and optimizing promotion information is an effective way to bring higher CTR. This is exactly the assumption that we are going to proof.

B. Mathematic Derivation

Noise interference during information transfer is unavoidable. The information theory has been proved by Shannon Weaver and Maletzke that information is transferred in a specific manner along a certain signal path, and will inevitably be affected by noise interference during transfer. That is to say, ad information will not finally be delivered as initially wished because of various interferences. The impact factors include the public image of the media, the forms of expression, the transfer environment, and the cognitive level and mentation state of the visitors. In that case, some words well designed to attract visitors might lead to an opposite result unwillingly, which is called negative transfer.

Suppose the promotion information of the ad is outputted with a probability of x, and delivered to the visitors with a probability of q positively, and q’ negatively. The other part of information is outputted with a probability of 1-x, and delivered to the visitors with a probability of s positively, and s’ negatively, as illustrated in Figure 1.

Also suppose there is functional relationship between C (CTR) and p (Positive transfer), marked as C=f (p), 0<p<1. And to avoid the interference of Noise E, we’d better find the functional relationship between p and x firstly.

Then we mark the information array inputted to the network media as $A=(a_1, a_2, \ldots, a_m)$, and the one outputted from it as $B=(b_1, b_2, \ldots, b_n)$. Due to the Noise E, m and n are not necessary equal. It follows that $P(A)=x$, $P(B)=q$, $p=q+s$, and the total probability of promotion information positively delivering is

$$P(B | A)=P(b_j | a_i) = xq$$

$$i=1,2,\ldots,m; \ j=1,2,\ldots,n$$  (1)
In Equation (1), \( P(B|A) \) is a symbol for the output probability of \( B \) on the premise of input probability of \( A \), called forward probability. Similarly, posterior probability, which means the output probability of \( A \) on the premise of input probability of \( B \), can be deduced.

\[
 P(A | B) = P(A | b_j) = \frac{xp}{xq + (1-x)s} 
\]

\((i = 1, 2, ..., m; j = 1, 2, ..., n)\)  \( (2) \)

And the joint probability of both input and output promotion information is

\[
P(a_j b_j) = P(a_j) P(b_j | a_j) \]  \( (3) \)

Then, it can be deduced according to probability reasoning mode of Bayes’ theorem that

\[
P(a_j | b_j) = \frac{P(a_j b_j)}{P(b_j)} = \frac{P(a_j) P(b_j | a_j)}{\sum_{i=1}^{m} P(a_i) P(b_j | a_i)} \]  \( (4) \)

From the above Equation (1)-(4), the equation can be derived that

\[
P(a_j b_j) = x^2 q = q(\frac{xq}{xq + (1-x)s}) \]  \( (5) \)

According to the known equation \( p = q + s \), the final result is:

\[
p = 2s \times \frac{s}{1+x} \]  \( (6) \)

Since \( s \) maintains a relative stabilization comparing with \( x \), from the above Equation (6) we find that the change of argument \( x \) has positive effect on induced variable \( p \). Thus the conclusion can be drawn that \( p \) is a monotonically increasing function of \( x \). And the exact function relationship between \( C \) and \( x \) will be defined from the empirical research below.

C. The Empirical Research

Still taken as the sample, the ad above was separately embedded onto the homepages of 2 Movie BBS, the PV (Page Viewer) number of which was stabilized at about 4000-5000 per day. Under the premise of maintaining consist information amount, by increasing or reducing the adjunct words frequently for 36 times, the probability \( x \) that promotion information is outputted with will be evaluated at \( 0, 2\%, 4\%, 6\%... 70\% \). To improve the Reliability and validity of the test as far as we can, we choose the same location and the same period per day. The cross test method was applied for the 36 modified ads on the 2 websites, as illustrated in Figure 2.

![Figure 2: The Cross Test Method Applied](image)

After the test of 36 phases, the data was collected for parallel-forms reliability analysis and chi square test. The value of parallel-forms reliability is 0.917, and the value of chi square is 0.015. Thus we believe that the data from the test is congruous with expectation, as listed in Table 1.

<table>
<thead>
<tr>
<th>Creative</th>
<th>Probability x</th>
<th>PV in Total</th>
<th>Click Number in Total</th>
<th>CTR(C) i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creative1</td>
<td>0</td>
<td>8887</td>
<td>12</td>
<td>0.135%</td>
</tr>
<tr>
<td>Creative2</td>
<td>0.02</td>
<td>8703</td>
<td>17</td>
<td>0.195%</td>
</tr>
<tr>
<td>Creative3</td>
<td>0.04</td>
<td>9279</td>
<td>21</td>
<td>0.226%</td>
</tr>
<tr>
<td>Creative28</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Creative29</td>
<td>0.54</td>
<td>8068</td>
<td>66</td>
<td>0.818</td>
</tr>
<tr>
<td>Creative30</td>
<td>0.56</td>
<td>8965</td>
<td>62</td>
<td>0.820</td>
</tr>
<tr>
<td>Creative31</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Creative32</td>
<td>0.58</td>
<td>8670</td>
<td>73</td>
<td>0.842</td>
</tr>
<tr>
<td>Creative33</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Creative34</td>
<td>0.66</td>
<td>10160</td>
<td>70</td>
<td>0.683</td>
</tr>
<tr>
<td>Creative35</td>
<td>0.68</td>
<td>9324</td>
<td>65</td>
<td>0.697</td>
</tr>
<tr>
<td>Creative36</td>
<td>0.70</td>
<td>9228</td>
<td>63</td>
<td>0.683</td>
</tr>
</tbody>
</table>

From the above set of data we can basically judge that CTR is positively correlated with \( x \), the transfer rate of promotion information. To confirm that the effects of product information and supporting information to CTR is negligible, the following test is to check the effect of increasing or reducing words belonging to the product information and supporting information separately.

After double 36-times tests, the scatter gram comes out and shows that product information plays an important role in the beginning, but the influence turns insignificant later, while there is no distinct correlation between CRT and supporting information (shown in Figure 3). Analyzed by Kano mode, product information is classified as Must-be quality, a few of which is enough but a lot is unnecessary; supporting information is classified as Indifferent quality, no matter more or less, negligible to the result.
XINGJUN SHI: A STUDY ON FACTORS INFLUENCING THE CLICK THROUGH RATE, CTR, OF INTERNET …

Here we come back to the relationship between the dependent variable C and independent variable x, pictorialized in the following scatter gram, Figure 4.

According to the set of data, the Pearson correlation coefficient can be calculated as $r=0.916$. Since that the inference that $C$ is positively linear correlated with $x$ has been confirmed. Suppose the relationship is linear, marked as $C=ax+b+\epsilon$, and then the regression equation can be calculated by means of the least square fit method as

$$C=0.0091x+0.0019$$

By hypothesis testing, $t=6.77$, so the significance level of difference $p$ is far less than 0.05, demonstrating that both the linear relationship and the regression coefficient are significant.

It can be inferred from the scatter gram, obviously, that the linear relationship is much more significant on the definition range of $(0, 0.5)$. However, with the continual increase of $x$, especially over 0.6, $C$ tends to decrease oppositely, and according to Table 1 and Figure 3, $C$ reaches its maximum 0.842 when $x=0.58$. Thus by correlation analysis, we conclude that $C=f(x)$ is a subsection function, and the critical point is $x=0.5$, as Equation (7).

$$C=f(x) = \begin{cases} 
  a'x+b' & x \in [0, 0.5) \\
  0.0841 - 0.252(x - 0.58)^2 & x \in [0.5, 0.7) 
\end{cases}$$

(8)

IV. CONCLUSION

By means of mathematical inference and empirical research, we have proved that CTR is a monotonically increasing function of the transfer rate of promotion information and then found the exact function expression. Therefore, we can come to the conclusion that in order to raise CTR by the charging mode of CPM, advertisers can increase the transfer rate of promotion information by improving the creativity of their ads constantly, so as to find the best one and achieve the top result. However, it is not unlimited. Once the eyeballs of the visitors are fully filled of promotion information, it is not easy for them to find what the ad wants to explain actually, like what the product is and what it is used for. So too much promotion information may lead to an unfavorable result oppositely. Maintaining the promotion information at the rate of about 0.58, we can come to a best result and satisfying ROI.

ACKNOWLEDGMENTS

This work was supported by Zhejiang philosophy & social sciences projects from Zhejiang Federation of Humanities and Social Sciences Circles (Grant No. 16NDJC182YB) and Zhejiang education planning projects from Zhejiang Education Department (Grant No. 2016SCG225).

REFERENCES

[17] Greer, J.D. Advertising on traditional media sites: Can the traditional business model be translated to the Web?[J]. The Social Science Journal, 2004(41):107-113