A novel approach for improving Real Time Bidding Strategy

Shikha Gupta, Atul Mishra
Department of Computer Engineering, JCBOSE UST, YMCA, Faridabad, Haryana, India

Abstract. Real Time Bidding is a vital part of online advertising. It helps the publisher to publish the most relevant advertisement to the users in real time. Real Time Bidding is a technique which is helpful for both advertisers and publishers. To place an ad automatically and optimally, it is critical for publishers and advertisers to devise a learning algorithm to cleverly place an ad impression in real-time. The bidding process continues for the entire episode for all the advertisers until the episode ends or the advertiser goes out of budget. Thus, it is important to find an algorithm which will help in optimal bidding of advertisers in real time considering both present and future status.

1. Introduction
Online Marketing is one of the most prominent fields in today’s world. It is important from the perspective of the user, the publishers and the advertisers. The user gets benefit by getting relevant advertisements to their query [15]. The advertisers generate revenue by displaying their advertisement and thus selling their product to the customers [1]. The publishers earn revenue from the advertisers by acting as a bridge between the users and the advertisers [2]. It displays the advertisements to the user based on various factors. All these three are interdependent and affects others [3]. The process that is followed is defined below:

1.1. Deciding the Bid Amount (done by advertiser)
Earlier the price of the queries was fixed. The price was not decided in real time. The algorithms used were: Greedy Algorithms, Balance Algorithm, and Generalized Balance Algorithm for selecting the advertiser. Now, various machine learning algorithms exist. Each advertiser have a bidding agent, which keep complete record of its CTR for that query, budget, lifetime, transition probability, Quality and Relevance of your Keywords, winning probability, estimation of other advertisers bid amount, market price and so on depending on the approach used [11]. All past records are maintained and future decisions are predicted based on previous and current data. The bidding agent decides the amount to bid based on this data.

1.2. Mapping of query with advertiser (done by publisher)
When the user fires a query, it reaches to the publisher, who finds the relevant advertisers from the list of all available advertisers. It also considers several factors for selecting the advertisers [17]. It then forwards the query to the selected advertisers asking them that if they are interested to participate in the query or not. Mapping of Query with advertisers is done. A no. of algorithms exists to find the relevant advertisers like Keyword matching, Broad matching.
1.3. Deciding whether to participate in the bidding or not

A request arrives at the advertiser side asking whether they are interest to participate or not. The advertiser considers a number of factors in making the decision. The advertiser sends back their answer to the publisher.

1.4. Ranking of advertiser (done by publisher)

Now, the publisher ranks all the interested advertisers based on their bid amount and quality score [12]. There are various factors considered for calculation the quality score.

1.5. Deciding the amount to pay (done by publisher)

After doing the ranking, the amount to be paid by the winning advertiser is decided depending on the approach used like first price auction, second price auction, hybrid of first and second price auction [14].

1.6. Declaring the result of bid (done by publisher)

Then, the winner is announced to everyone. The details send to the winner are: amount to pay and then, whether the link is clicked or not (CTR), whether the product was purchased or not (Conversion rate) which are required in further processing and making new decisions.

The above steps defined the various steps followed for the Real time bidding process. This paper contains five sections. Section 2 contains a literature review describing the various approaches used in real time bidding process. Section 3 defines the various problems faced by the existing approaches. In the section 4 of this paper, a new approach for handling these issues and improving the performance has been discussed. Section 5 includes the future scope and conclusion of this paper.

2. Literature Review

Bidding Optimization is one of the major research areas in real time bidding which helps in deciding the right price for the auction without compromising on the performance. A number of approaches exist for this problem [4]. Some of them are discussed in this paper. Linear bidding strategy was proposed by Perlich et al. [5] which was based on impression evaluation. Zhan et al [6] proposed that there is a non-linear relationship between impression evaluation and optimal bid. Both the above approaches consider bidding optimization as a static problem and thus fail to deal with the dynamic nature of bidding.

Zhou et al. [7] represented bidding optimization as an online knapsack problem. In this approach high quality ad requests are selected depending on an exponential function with respect to the period of the budget. However, the limitation with this approach is that it considers unlimited supply of ad requests, which is impractical. Babaioff et al. [8] handles dynamic price bidding using multi-armed bandit approach and decides the optimal price using upper confidence bound. However, this technique does not use any information of prior distribution, which can be of great interest in making the decision [9]. There is a need to maintain the balance between the user, advertiser and publisher.

3. Problem Statement

The existing process [10], matches the query terms with the keywords specified by the user without giving special attention or privilege to the words which are least common which will later lead to some unmatched queries, i.e., no relevant advertiser with remaining budget is left.

Another problem faced is explore versus exploit, i.e., deciding when to explore and when to exploit [18]. The existing processes keep on exploiting the most common advertisers and do not explore the other ones. If the focus is only on relevancy, our publisher will never achieve the gigantic sum. Instead, it will only exploit the nearest source of rewards, even if this source is small (exploitation). But if our publisher does a little bit of exploration, it can find the big reward. This is
what we call the exploration/exploitation trade off. The most crucial tradeoff faced by the publisher at
every trial is between “exploitation” of the advertiser having the highest expected payoff and
"exploration" to get more information about the expected payoffs of the other advertisers. A solution
must be provided to handle this issue.

4. Proposed Work
The proposed work helps in solving the above two issues identified in this paper. Whenever a query is
fired, a number of steps are followed to get the desired output. Here, the user enters a query (a set of
keywords) to get the result, i.e., the most relevant advertiser. The publisher will show the result
considering the relevancy and the expected future queries also.

Here, we have a set of advertisers \( A_i \) \{\( A_1, A_2, A_3, \ldots, A_n \)\} and each advertiser have a set of pre-
defined keywords \( K_{ij} \) (suppose \( j \) lies between 1 to 3).

4.1. Deciding the bid amount
When a publisher enters the bidding process, the first step is to decide the keywords in which the
advertiser is interested. After selecting them, the bid amount for each keyword is calculated. A number
of factors are considered by the advertiser for deciding the bid amount. They are:

- Selection ratio of each keyword (CTR): How many times the advertiser is selected when
  search term value for this keyword is 1, i.e., when this word exist in the query. By using this,
  the advertiser can focus on the keywords having high selection ratio.

- Weight of keyword: There is a weight associated with each keyword (0 to 1) representing its
  relevance with the advertiser and the sum of all the keywords for an advertiser is 1. The
  advertiser can focus on more relevant keywords.

- Frequency Ratio (FR): This can be done by checking whether the word is common or not. As
  there will be more no. of advertisers bidding for common terms and less for least common. So,
  the advertiser can bid low amount for least common words as the chances of getting selected
  is still high. This helps in estimation of other advertiser’s bid amount. This can be done by
  calculating the ratio of frequency of that term in the past search queries to the total number of
  queries received as the advertiser does not have the list of keywords of other advertisers.

- Budget: The bid amount should be less than remaining budget.

Suppose there are ‘m’ no of keywords given by the advertiser. Here, BD is some constant value as
decided by the advertiser depending on the budget. Then the formula used can be defined as:

\[
\text{Bid Amount} = BD \times (\text{Weight of keyword} \times \text{CTR} \times \text{FR})
\]

The bid amount for all the keywords is calculated and:

\[
\text{Total Bid Amount} \leq \text{Remaining Budget}
\]

\[
\text{Freq ratio} = \frac{\text{Freq of keyword in queries}}{\text{Total number of queries}}
\]

If, freq ratio > 0.75

Then, it is common, FR = 1

Else if 0.25 < freq ratio < 0.75

Then, its occurrence is average, FR = 0.85

Else it is less common, FR = 0.75.

4.2. Mapping of Query with Advertisers
Whenever a query is fired, most relevant advertisers are found for the bidding process. Suppose there
are a total of 100 advertisers then, from them 10 most relevant are found so that bidding and ranking
can be done between them. While finding the most relevant advertisers, a number of things are
considered by the publisher:
Search term: Whether the terms present in the query matches with the keywords specified by the advertiser or not. Value of search term is 1 if found, otherwise zero.

Weight of keyword: There is a weight associated with each keyword (0 to 1) representing its relevance with the advertiser and the sum of all the keywords for an advertiser is 1. It will check the relevancy of the query with the advertiser.

ECTR: Effective Click Through Rate is the conversion rate which is defined as the ratio of the users who completed the goal (purchased the product) to the number of times the user visited the site (the advertiser is clicked by the user). This is also called as conversion rate [16].

Frequency of Keywords: Give preference to advertisers having least common words as their keywords: If an advertiser has specified a least common word as its keyword then preference should be given to that advertiser as there will exist a number of advertisers for common words. By doing this, matching of more queries with the advertisers can be done. This value is calculated by dividing the frequency of that keyword (no. of advertiser having that word as keyword) divided by the total no. of advertisers.

For example,

| Advertiser | Keywords | Weight |
|------------|----------|--------|
| A1         | K11      | 0.6    |
|            | K12      | 0.3    |
|            | K13      | 0.1    |
| A2         | K21      | 0.2    |
|            | K22      | 0.7    |
|            | K23      | 0.1    |
| A3         | K31      | 0.4    |
|            | K32      | 0.3    |
|            | K33      | 0.3    |

Here, suppose a query with terms K21, K32 arrives. Then, the relevancy of Q1 is more with A3 (because K32 > K21), but assume that the word K21 is less common so some preference should be given to A2 while calculating the total score as we will find more queries which will match A2 (as it contains common words as keywords). In this way, more queries can be matched in the future.

This approach will also solve the problem of exploitation as it will explore the advertisers with least common keywords (as the chances of matching of least common word is less in future) and will not keep on exploiting the advertisers having more common terms as their keywords an high bid value (as their chances of matching with the future query term is more often).

The score for all the advertisers is calculated, and from them most relevant advertisers (advertisers with high score value) are selected for the next step. Let’s say there are x terms in the query. Formula to be used is:

\[
\text{Score} = \sum_{j=1}^{m} \sum_{p=1}^{x} \frac{\text{Search term value}_{pj} \times \text{Weight of keyword}_{j}}{\text{Frequency of keyword}_{j}/\text{Total no. of Advertiser}} \times \text{ECTR}
\]

Suppose there are ‘m’ no of keywords given by the advertiser and ‘x’ is the number of terms in the query. The score should be greater than some threshold value to get selected.

The advertisers with high Score value are selected and a message is send to them.

4.3. Deciding whether to participate in the bid or not

After selecting the most relevant advertisers for a query, it was asked by the advertiser that whether they want to participate for that query or not. The decision is made based on the following formula:
Total Value = $\sum_{j=1}^{m} \sum_{p=1}^{x} \text{Search term value}_p \times \text{Weight of keyword}_j \times \text{CTR}_j / \text{FR}_j$

In this the advertiser will give preference to keywords having less frequency ratio. If this total value is greater than some predefined value then it will participate, otherwise not.

4.4. Ranking of Advertisers
Once the decision is made by the advertisers, it is communicated back to the publisher. Now, the ranking of advertisers is done by calculating the Total score for each advertiser by using the formula [11]:

**Total Score = Bid amount, * Quality Score**

**Quality Score= CTR * Score * ECTR * Percentage of balance budget**

Based on the total score value, the ranking of advertisers is done and top rankers are displayed to the user.

4.5. Deciding the amount to pay
Some Hybrid of first and second price auction are used so that there will not be any loss to the publisher and the advertiser too will not pay more than the actual price and will also get profit for maintaining the quality of its advertisement. If the bid price is less than some threshold value then, first price auction is used otherwise, second price auction mechanism is used which is defined as [13]:

\[
\text{Actual cost per click} = \frac{\text{Total Score of Advertiser below you}}{\text{Your Quality score}} + 0.01
\]

Table 1 and Table 2 below shows the database of publisher and advertiser respectively which keeps on updating for making the decisions. The publisher keeps record of various advertisers. It stores the keywords specified by it along with their weight and bid price. It also stores the total budget, Click through rate, Quality Score, Balance budget and Effective click through rate which keeps on changing and updating for future queries. The advertiser keeps record of its keyword along with their weight, bid price, CTR and frequency. It also stores its balance budget and ECTR. Time to time updating is done for making correct decision of future queries.

### Table 1. Database of Publisher

| List of advertisers | List of keywords for each advertiser | Weight of each keyword | Bid Price of each Keyword | Total budget of each advertiser | Click Through Rate | Quality Score | Balance Budget | ECTR |
|---------------------|-------------------------------------|------------------------|---------------------------|--------------------------------|-------------------|--------------|---------------|------|
| A1                  | K_{11}                              | W_{11}                 | P_{11}                    | B_1                            | CTR_1             | QS_1         | BB_1          |      |
|                     | K_{12}                              | W_{12}                 | P_{12}                    |                                |                   |              |               |      |
|                     | K_{13}                              | W_{13}                 | P_{13}                    |                                |                   |              |               |      |
| A2                  | K_{21}                              | W_{21}                 | P_{21}                    | B_2                            | CTR_2             | QS_2         | BB_2          |      |
|                     | K_{22}                              | W_{22}                 | P_{22}                    |                                |                   |              |               |      |
|                     | K_{23}                              | W_{23}                 | P_{23}                    |                                |                   |              |               |      |
Table 2. Database of Advertiser

| List of keywords | Weight of each keyword | Bid Price of each keyword | CTR for each keyword | Frequency of Keyword | Balance budget | ECTR |
|------------------|------------------------|---------------------------|---------------------|---------------------|---------------|------|
| K_1              | W_1                    | P_{11}                     | CTR_1              | FR_1                | BB            | ECTR |
| K_2              | W_2                    | P_{12}                     | CTR_2              | FR_2                |               |      |
| K_3              | W_3                    | P_{13}                     | CTR_3              | FR_3                |               |      |

5. Conclusion and Future Scope
The In this paper, we propose a novel approach for Real Time Bidding approach. It helps in solving the two issues identified. It will map the queries with the relevant advertisers keeping in mind the expected future queries. Thus providing profit to both publishers and advertisers. It also deals with Explore versus Exploitation issue. It will not consider advertisers based only on their bid amount but will also consider their future scope. Instead of local maxima it will consider global maxima. Hence, the proposed approach is better than previous approaches. A lot of work can still be done to improve its performance; various other factors can be considered for calculating the different values.

References
[1] “The Arrival of Real-Time Bidding and What it Means for Media Buyers”, Google white paper.
[2] Adikari S., Dutta K. (2015) “Real Time Bidding in Online Digital Advertisement”, In: Donnellan B., Helfert M., Kenneally J., VanderMeer D., Rothenberger M., Winter R. (eds) New Horizons in Design Science: Broadening the Research Agenda. DESRIST 2015. Lecture Notes in Computer Science, vol 9073. Springer, Cham.
[3] Kan Ren, Weinan Zhang, Ke Chang, Yifei Rong, Yong Yu, and Jun Wang, “Bidding Machine: Learning to Bid for Directly Optimizing Profits in Display Advertising”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 30, NO. 4, APRIL 2018.
[4] Jun Wang, Weinan Zhang and Shuai Yuan (2017), "Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting", Foundations and Trends® in Information Retrieval: Vol. 11: No. 4-5, pp 297-435. http://dx.doi.org/10.1561/1500000049.
[5] Claudia Perlich,Brian Dalessandro, Rod Hook,Ort Stitelman,Troy Raeder and Foster Provost, 2012,”Bid optimizing and inventory scoring in targeted online advertising” In 18th SIGKDD, ACM, 804-812.
[6] Weinan Zhang, Shuai Yuan, Jun Wang, 2014, “Optimal Real-Time Bidding for Display Advertising”, In 20th SIGKDD, 1077-1086.
[7] Y.Zhou, D.Chakrabarty, R.Lukose, “Budget Constrained Bidding in keyword auctions and online knapsack problems”, Proceeding of the 17th International Conference on World Wide Web, 2008.
[8] M.Babaioff, S.Dughmi, R.Kleinberg and A.Slivkins, “Dynamic Pricing with limited supply”, The 13th ACM Conference on Electronic Commerce, 2012.
[9] Kuang-Chih Lee Ali Jalali Ali Dasdan, “Real Time Bid Optimization with Smooth Budget Delivery in Online Advertising”, arXiv:1305.3011v1 [cs.GT], 2013.
[10] Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman,“Mining massive datasets”.
[11] https://blog.tryadhawk.com/google-adwords/how-google-adwords-works/#how-does-the adwords-auction-work.
[12] Tsan-Ming Choi, Xin Li, and Cheng Ma, “Search-Based Advertising Auctions With Choice Based Budget Constraint”, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, VOL. 45, NO. 8, AUGUST 2015.
[13] Yong Yuan, Feiyue Wang, Juanduian Li1, Rui Qin, “A Survey on Real Time Bidding Advertising”, 978- 1-4799-6058-3/14 ©2014 IEEE.
[14] Shalinda Adikari, Kaushik Dutta, “Real Time Bidding in Online Digital Advertisement”, Springer International Publishing Switzerland 2015, LNCS 9073, pp. 19–38, 2015.

[15] Shuai Yuan, Jun Wang, Xiaoxue Zhao, “Real-time Bidding for Online Advertising: Measurement and Analysis”, In ADKDD 2013, arXiv:1306.6542v1 [cs.GT].

[16] https://www.wordstream.com/conversion-rate

[17] https://support.google.com/adsense/answer/160525?hl=en

[18] Shuai Yuan, Jun Wang, Bowei Chen, Peter Mason, Sam Seljan, “An Empirical Study of Reserve Price Optimisation in Real-Time Bidding”, DOI: 10.1145/2623330.2623357, Conference: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014.