Mathematical Modeling of Competition in Sponsored Search Market

Jerry J. Liu    Dah Ming Chiu
Department of Information Engineering
The Chinese University of Hong Kong
{ljl008, dmchiu}@ie.cuhk.edu.hk

ABSTRACT
Sponsored search mechanisms have drawn much attention from both academic community and industry in recent years since the seminal papers of [3] and [4]. However, most of the existing literature concentrates on the mechanism design and analysis within the scope of only one search engine in the market. In this paper we propose a mathematical framework for modeling the interaction of publishers, advertisers and end users in a competitive market. We study the competition between two search engines as a three-stage dynamic game and prove the existence of Nash equilibrium prices when allowing advertisers to participate in both advertising systems simultaneously. To compare the expected revenues and social welfare under competition and monopoly, we carry out extensive simulation under common parameter setting of participants. Our results can provide useful insight in regulating the sponsored search market and protecting the interests of advertisers and end users.

1. INTRODUCTION
Internet advertising has become a main source of revenue for primary search engines nowadays. According to the newly-released report by Interactive Advertising Bureau and PricewaterhouseCoopers [1], Internet advertising in the United States reached $22.7 billion in total revenue for the full year of 2009, where sponsored search revenue accounted for 47 percent of the total revenue.

A typical Internet search market consists of three parties: publishers (i.e., search engines), advertisers and end users. In the current age of information explosion, more and more people rely on search engines to pin down their favored products or services. Whenever a query is submitted to the engines by end users, their intents or interests can be potentially captured by the engines through the inputted keywords. These intents of search users can then be sold by search engines to companies who are interested in targeting their products to these users. Nowadays, major search engine operators like Google, Yahoo! and Microsoft display advertisements in the form of sponsored links, which appears alongside the algorithmic links (also known as organic links) in the search results pages. For each keyword, there are usually more than one available advertising slot in the results page. How to effectively allocate these slots and charge the advertisers have been studied and discussed extensively in recent years among people in both academia and industry. Take Google’s AdWords program for example. In this advertising platform, advertisers could choose multiple keywords they are interested in, and for each keyword indicate the maximal willingness to pay for each click and the budget to spend over a period of time. Whenever a user clicks on the sponsored link and is re-directed to the advertisers’ site, certain payment is charged by the program until the advertisers’ budget is used up.

Most of the existing works focus on the interaction of the three parties within the scope of only one search engine’s advertising system. However, considering there is usually more than one company providing search service in the market, one natural question would be how the market would evolve when there exists competition among multiple search engines. In particular, will all users and advertisers gradually shift to one leading engine or will the “inferior” engines still earn enough profits to survive when competing with the leading one? What would be the consequences if one search engine monopolizes the market? These concerns arise from the current situation of high levels of concentration in search engine market: Google is widely considered to possess the leading technology and obtain the largest market shares in most countries and regions, followed by Yahoo! and Microsoft Bing.

This paper aims to formulate a reasonable model to study the competition between two search engine operators and help to address some of the intriguing questions mentioned above. We will consider a three-stage dynamic game model. In stage I, the two operators’ services determine how the market of end users is split. In stage II, the two operators simultaneously determine their prices to advertisers. In stage III, the advertisers choose the operator in which they can obtain highest utility based on the announced prices. Based on the monopoly formulation, we then analyze the strategic behaviors of end users and advertisers under duopoly and...
find the Nash Equilibrium prices of both search engines in Section 4. We provide some simulation results and discussion in Section 5 and conclude in Section 6. Due to space limitations, most of the proofs are omitted and we refer interesting readers to our online technical report [19].

2. RELATED WORK

There are mainly two lines of research work in the sponsored search area. The mainstream of literature focuses on the interaction between advertisers and search engines and aims to understand and devise viable mechanism for the Internet advertising market. There are significant works on the auction mechanism held by major search engines, starting from two seminal works of [3] and [4] which independently investigated the “generalized second-price” (GSP) auction prevailing in major search engines such as Google and Yahoo!. In [11] the authors compared the “direct ranking” method by Overture with the “revenue ranking” method by Google and proposed a truthful mechanism named as “laddered auction”. [16] and [17] relaxed the basic assumption of separable click-through rate in [11] and modeled the externality effect among advertisements in the same search page. There is also an abundance of works on proposing more expressive but still scalable mechanism for sponsored search such as [12–15, 18].

It’s worth pointing out that a few papers considered the practical situation with multiple search engines in the market. In [6] the authors compared the revenues of two search engines with different click-through rates. The study in [7] considered competition between two search engines with different ranking rules. We assert that this assumption is unrealistic since major engines tend to use the same policy which proves to work efficiently and it is unlikely that one engine would switch back to the obsolete rules. The Nash equilibrium solution in the former paper [6] is also not quite practical since it requires advertisers to adopt certain randomized strategy. It would be very difficult for individual advertisers to implement such complex strategies which incur unnecessary maintenance cost.

The other line of work is developed mainly by economists to address the broad issues of search engine competition from social welfare perspective. [8] introduced a quality choice game model where end users choose the search engine with highest quality of search results, and showed that no Nash equilibrium exists in this game. Based on this proposition, the author argues that the search engine market would evolve towards monopoly in the absence of necessary regulatory interventions. [9] proposed a duopoly model which has some similarity to our formulation, however, as many of the technical details of the practical advertising system are ignored, it is doubtful whether this can serve as an accurate model to predict the outcome of search engine market. Similarly, [8] faces the same problem that the vague description of participants’ utility may not be strong enough to support the predictive conclusions in the paper.

These two lines of important work have little intersections so far: the mainstream of work concentrates on the technical progress in designing “better” advertising system, and the other line usually involves less technical details (like the budgets of advertisers in practical advertising system) and targets the macro-effect of competitive market. In view of this, we believe that a comprehensive study of the current search engine ecosystem in a competitive way is vital for addressing many of the unresolved issues in this thriving market. Our work manages to narrow the gap between these two directions of research and makes some initial progress in this direction. This observation helps differentiate our work from most of the existing literature.

3. THE MONOPOLY MARKET MODEL

Suppose there is only one search engine in the market servicing a fixed set of end users and providing advertising opportunity for a set of advertisers denoted by $\mathcal{I}$ ($|\mathcal{I}| = m$). Assume all users are homogeneous and each of them tends to generate the same number of impressions (or clicks) for a particular keyword. Since the search engine owns a fixed number of users, it would be able to supply a fixed number of attentions (in the form of impressions or clicks) for advertisers. For a given interval, let the supply of attentions be $S$. Each advertiser $i \in \mathcal{I}$ has two private parameters: value $v_i$, denoting $i$’s maximal willingness to pay for each attention and budget $B_i$, in the time interval. After advertisers submit their values and budgets to the advertising system, the search engine needs to determine the optimal price per attention to maximize its revenue:

$$R = p \cdot \min(S, D(p)) = \min(p \cdot S, pD(p))$$

(1)

where $D(p)$ is the demand function over price $p$. Reorder the index of advertisers such that $v_j \leq v_{j+1}, j = 1, \ldots, m–1$. Then the aggregate demand can be written as:

$$D(p) = \sum_{i \in \mathcal{I}^+(p)} B_i \frac{p}{p^*}$$

(2)

where $\mathcal{I}^+(p) \triangleq \{ i \in \mathcal{I} : v_i > p \}$ denotes the set of advertisers with positive demand under the current price. Thus $pD(p) = \sum_{i \in \mathcal{I}^+(p)} B_i$ is a non-increasing function over $p$ since $\mathcal{I}^+(p)$ shrinks as price $p$ increases. We first give the following proposition.

**Proposition 1.** The revenue $R$ is maximized when $S = D(p)$, i.e., when the demand equals the supply.

Proposition 1 can be proved by contradiction. If supply exceeds demand under the current price, the search engine will cut down the price to achieve higher revenue (since $R = pD(p)$ is non-increasing over price $p$); if demand exceeds supply, the engine can raise the price and achieve higher revenue (since $R = pS$ is monotonically increasing over $p$).

By letting demand equal to supply, we have

$$p^*(\mathcal{I}) = \frac{\sum_{i \in \mathcal{I}^+(p)} B_i}{S}$$

(3)

Notice that the term of price appears in both sides of the equation. Thus in general we cannot derive the closed-form solution for optimal price. Since $pD(p)$ is piece-wise constant and (weakly) decreasing over $p$, we can illustrate the search engine revenue through examples in Figure 1. Here we assume there are four advertisers ordered by their values and initially when the price is zero, $\mathcal{I}^+(p) = \mathcal{I} = \{1, 2, 3, 4\}$. As the price exceeds $v_1$, advertiser 1 would quit and $\mathcal{I}^+(p)$ becomes $\{2, 3, 4\}$. The intersection of demand and supply gives the optimal price $p^*$ which is located in $[v_1, v_2]$. According to equation (3),

$$p^*(\mathcal{I}) = (B_2 + B_3 + B_4)/S.$$

In Figure 1(b) we also show the other case when there is one advertiser who is indifferent between participating and quitting the ad campaign since the optimal price is equal to its value. In typical advertising systems like Google AdWords, after advertisers submit

1In practice, the optimal price is usually determined automatically by an auction mechanism. Specifically, this automation process can be regarded as an ascending-bid auction [2] where the auctioneer (i.e., the search engine) iteratively raises the price until there is no excessive demand than supply. Considering strategic issues, [10] proposes an asymptotically revenue-maximizing truthful mechanism. For simplicity of analysis, we ignore the detailed implementation of auctions and assume the search engine can solve the revenue-maximizing problem instantaneously.
their maximal willingness to pay (i.e., their values) and budgets, the ad system would automatically allocate attentions to advertisers as long as the price doesn’t exceed their values and the budgets have not been exhausted yet. Thus here for ease of expression we can assume that the indifferent advertiser would continue participating in the ad campaign under its budget constraint. For example, in Figure 1(b), $p^*(Z) = v_2$ and all the remaining supply of $S - \sum_{i=1}^{2} v_i$ is allocated to advertiser 2.

![Figure 1: Search Engine Revenue Over Prices](image)

We can now show a polynomial step algorithm for search engine to compute the optimal price. By inputting the parameters of advertisers (assuming the indexes of advertisers are re-ordered by their values), Algorithm 1 would return the value of optimal price. The time complexity of the algorithm is $O(n^2)$.

**Algorithm 1 Calculate Optimal Price $p^*(Z)$**

1. $v_0 = 0$;
2. for $i = 1 : m$;
3. \[ \text{sum} = 0; \]
4. for $j = i : m$;
5. \[ \text{sum} = \text{sum} + B_j; \]
6. end for;
7. $p = \text{sum}/S$;
8. if ($p \leq v_i$);
9. \[ \text{return max}(p, v_{i-1}); \]
10. end if;
11. end for;
12. \[ \text{return } v_m; \]

The maximal revenue of search engine is:

$$R = p^* \cdot S \tag{4}$$

The aggregate utility of advertisers is:

$$U_A = \sum_{i \in \mathcal{I}^+(p^*)} (v_i - p^* \cdot \frac{B_i}{p^*}) \tag{5}$$

The social welfare of the advertising system$^3$ is:

$$SW = R + U_A \tag{6}$$

**Lemma 2.** Given the set of participating advertisers, for any supply $S_1, S_2 \in [0, \infty)$, if $S_1 \geq S_2$, we have $p^*(S_1) \leq p^*(S_2)$ and $R(S_1) \geq R(S_2)$.

**4. THE DUOPOLY MARKET MODEL**

In this section we switch from the monopoly model to the competitive model with more than one search engine. Considering the likely situation where there is one leading search company and one major competitor in the market$^1$, we describe a duopoly model where one search engine has an advantage of technology over the other. We formulate their competition as a three-stage dynamic game as follows.

**4.1 Competition for End Users in Stage I**

In Stage I search engines would choose different strategies for attracting end users with different tastes. The user bases they attract in this Stage would be the decisive factor for determining their supply of user attentions to advertisers in subsequent stages. We assume that there are two horizontally and vertically differentiated search engines $\mathcal{J} = \{1, 2\}$ providing search results to users and selling ad opportunity to advertisers. Here by horizontal difference we mean the different design of their home pages and diversity of extra services such as email, news and other applications. Different users may have different tastes and preferences and hence be attracted by different search engines. Vertical difference denotes the quality of searching results. The higher the quality is, the better users and advertisers would perceive. We assume that search engine 1 possesses the leading technology to match ads to search queries and can provide better service for both users and advertisers than search engine 2. The dynamics of the game can be described as follows: each provider chooses a location in the characteristic space which denotes the specific feature of service it provides to users. And each user is characterized by an address reflecting his individual preference of ideal features search engines should provide. Searching at engine $j \in \mathcal{J}$ involves quadratic transportation cost for a user if engine $j$ is not located in his ideal position.

Assuming that a continuum of users are spread uniformly with unit density on the circumference of a unit circle as shown in Figure 2$^5$. Denote the address of a particular user as $t \in [0, 1)$. Without loss of generality, let search engine 1 locate at $x_1 = 0^\circ$ and engine 2 at $x_2 \in [0, 1)$. Then the utility of user searching in either engine would be as follows:

$$u_1(t) = q - C(t, x_1) = q - \min\{t^2, (1 - t)^2\} \tag{7}$$

$$u_2(t) = \zeta q - C(t, x_2) = \zeta q - (t - x_2)^2 \tag{8}$$

where $\zeta \in [0, 1]$ denotes the comparative “disability” of search engine 2 to provide the best search results for users; $q$ is the positive payoff users perceive when certain information is returned by the search engine for a particular query; $C(t, x_j)$ is the transportation cost incurred when there is certain distance between user’s address $t$ and search engine $j$’s location $x_j$.

Let $n_1(t) = u_2(t)$ we can find the locations of two indifferent users as $\xi_1$ and $\xi_2$ ($\xi_1 < \xi_2$) respectively. Then the market share of search engine 2 is $n_2(x_2) = \xi_2 - \xi_1$ and search engine 1 obtains the remaining market share: $n_1 = 1 - n_2$. By applying the first-order condition on $n_2(x_2)$, we can derive that $x_2 = \frac{1}{2} - 2(1 - \zeta)q$. We see, when two search engines provide the same quality of service $^3$We don’t consider search users’ utility in the expression here.

$^1$For example, in US around 70 percent of search volumes are conducted on Google while the biggest competitor Yahoo! accounts for about 15 percent. Source from “Top 20 Sites & Engines,” Hitwise, May 22, 2010 (http://www.hitwise.com/us/datacenter/main/dashboard-10133.html).

$^4$Since it is a circle, engine 1 can also be regarded as locating at the ending point $x_1 = 1$. 

---

$\text{Algorithm 1 (continued)}$
For simplicity of notation, we have normalized the discount factor \( \rho \) where 
\[ \rho = \frac{\pi_2}{\pi_1}, \]
subject to 
\[ \rho \leq \frac{p_2}{p_1}. \]

Assuming that advertisers are re-ordered according to \( \rho_i \). Then the division of advertisers can be depicted in Figure 3 where \( I_1(p_1, p_2) = \text{sponsored search market.} \)

- \( \{i \in I : \rho_i \leq \frac{p_2}{p_1}\} \) denotes the set of advertisers who prefer search engine 1 and \( I_2(p_1, p_2) = \{i \in I : \rho_i > \frac{p_2}{p_1}\} \) the set of advertisers preferring engine 2.

Figure 3: The Division of Advertisers

After initial price \( p_1 \) and \( p_2 \) are announced, the advertisers set is divided into \( I_1 \) and \( I_2 \). Then each search engine can compute its optimal price \( p_1^*(I_1) \) and \( p_2^*(I_2) \) independently as the monopoly case and price ratio \( \frac{p_2^*(I_2)}{p_2^*(I_1)} \) gets updated. If it happens that the new price ratio divides the advertisers set into \( I_1 \) and \( I_2 \), we say this is a Nash equilibrium (NE) price pair as \( (p_1^NE, p_2^NE) \) and neither search engine has incentive to deviate unilaterally. Otherwise, the process will iterate until the prices become stable.

Defining first the set of advertisers who participate the advertising campaign as follows.

\[ I_1^+(p_1, p_2) \triangleq \{i \in I : \rho_i \leq \frac{p_2}{p_1}, v_i \geq p_1\} \tag{12} \]
\[ I_2^+(p_1, p_2) \triangleq \{i \in I : \rho_i > \frac{p_2}{p_1}, \rho_i v_i \geq p_2\} \tag{13} \]

We now give the formal definition of NE price pair.

**Definition 3.** A price pair of \( (p_1, p_2) \) is called a Nash equilibrium price pair if \( p_1 = p^*(I_1^+(p_1, p_2)) \) and \( p_2 = p^*(I_2^+(p_1, p_2)) \) where \( p^*(I) \) is computed via Algorithm 1.

It remains to be shown whether NE price pair would exist in our formulation. The answer turns out to be negative in general. A simple counter-example is when there is only one advertiser in the system. No matter which search engine the advertiser chooses initially, the price in the other engine would be zero since it attracts no advertisers. Hence the advertiser would keep switching between two engines and no stable prices could ever be achieved. An immediate remedy to this “oscillation” problem is allowing advertisers to split their budgets and invest them into two engines simultaneously. With this assumption, the system is expected to achieve a stable state finally. We summarize this result in the following theorem.

**Theorem 4.** Assuming advertisers can purchase service from both search engines simultaneously, Nash equilibrium price pair always exists for any set of advertisers and supplies of search engines.

We now turn to compare the prices under competition and monopoly. Suppose that under competition the supplies of engine 1 and engine 2 are \( S_1 \) and \( S_2 \) \((S_2 \leq S_1)\) respectively, and under monopoly engine 1 would obtain all the supply as \( S_1 + S_2 \). The main results are given in the following theorem.

**Theorem 5.** Denoted by \( (p_1^{NE}, p_2^{NE}) \) the NE price pair and \( p^* \) the optimal price when engine 1 monopolizes the market, it must hold that \( p_2^{NE} \leq p^* \leq p_1^{NE} \).

Theorem 5 implies that the leading engine can raise much more revenues than its competitor due to higher supply \((S_1 \geq S_2)\) as well as higher price \( p_1^{NE} \geq p_2^{NE} \). This advantage in revenue makes the leader capable of investing more on the improvement of search technology which may further intensify its dominance in the
5. SIMULATION RESULTS AND OBSERVATIONS

In this section we present our simulation results and observations. There are four major criteria we would like to explore in the system:

(a.1) Prices: We would show the price differences under competition and monopoly.

(a.2) Revenues: It would be intriguing to study the comparative results of total revenues under competition and monopoly. The gap between revenues under competition and monopoly would serve as a signal of whether the leading company would like to propose a merger or acquisition to its competitor.

(a.3) Aggregate Utility of Advertisers: We compare the aggregate utility of advertisers to see whether monopoly would be detrimental to the interest of advertisers, and if so, how severe the loss would be.

(a.4) Social Welfare: Social welfare can be regarded as the realized values of advertisers which denotes the interest of the advertising community as a whole. Under competition, the social welfare is computed as follows:

\[ SW = \sum_{i \in \mathcal{S}_1} v_i q_i + \sum_{i \in \mathcal{S}_2} p_i v_i q_i, \]  

where \( q_i \) is amount of supply allocated to advertiser \( i \).

We consider two search engines equally dividing the market and the total supply is normalized to unity. Thus the supply of either search engine is \( S_1 = S_2 = 0.5 \). Advertisers’ values are uniformly distributed over \((18, 20)\), and their budgets are also drawn from uniform distribution with expectation \( E(B) = 4 \). Discount factors of advertisers are uniformly distributed over \((0.5, 0.9)\). The simulation results are presented in Figure 4. We can make the following observations through Figure 4(a)-(d):

1) As the number of advertisers grows, the prices, revenues and social welfare would all increase. This is because as more advertisers participate, the demand for the limited supply would get boosted, which would finally drive up the unit price per supply and raise the revenue of search engines. When more advertisers compete, only those advertisers with higher values can stay and obtain some of the total supply, therefore the realized values of advertisers would be larger and the social welfare get enhanced.

2) After the number of advertisers reaches about five, the growth of prices and revenues seems saturated: more advertisers would not bring evident enhancement in prices and revenues. This can be derived from our parameter setting: \( E(B)/E(v) = 4/19 \approx 0.2 \) is the approximate amount of demand for each advertiser, and since the total supply is one, in expectation it would be sufficient for five advertisers to consume all the supply.

3) Figure 4(a) corresponds with the conclusion in theorem 5. Denote \((p_1, p_2)\) as the duopoly prices and \( p_M \) as the monopoly price. We notice that \( p_M \) is actually very close to \( p_1 \) but much larger than \( p_2 \). This is because the monopoly engine and engine 1 in competition face advertisers with the same distribution of values. Recall that value is the maximal willingness to pay for advertisers, thus when there are too many advertisers competing...
with each other, the price would approach the maximal possible value, which is 20 according to the distribution range. However, for search engine 2, the actual values of advertisers are the original values discounted by ρ, thus p2 is much smaller.

4) Figure 4(b) shows that revenue of search engine 1 is larger than that of engine 2. This can be easily deducted since the revenue of each engine is \( R_1 = p_1 \cdot S_1 \), \( R_2 = p_2 \cdot S_2 \) and we have \( p_1 > p_2 \), \( S_1 = S_2 \). As we have mentioned, \( p_M \) is approximately equal to \( p_1 \). Therefore we have the monopoly revenue as \( R_M \approx p_1 \cdot (S_1 + S_2) = R_1 + \frac{S_1}{S_2} R_2 > R_1 + R_2 \). This explains the gap between total revenue under competition and monopoly in figure 4(b).

5) Figure 4(c) shows that advertisers as a whole would benefit from competition since engine 2 would provide a much lower price. Figure 4(d) indicates that social welfare under competition is lower than that in monopoly since the realized values in equation (14) get discounted due to the effect of ρ.

6. CONCLUSION

We propose an analytical framework to model the interaction of publishers, advertisers and users. We give the analytical results of price and revenue for monopoly market. For duopoly market, we formulate a three-stage dynamic game model and prove the existence of Nash equilibrium. By carrying out simulations, the comparative results of revenues and social welfare under competition and monopoly are then presented and discussed extensively in the paper. Our analysis shows that although the cooperation between search engines can probably bring more total revenues, advertisers and users may be averse to such plan which eliminates their freedom to choose from diverse services provided by different search companies.

7. ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their valuable comments and suggestions. This work was partially supported by Direct Grant 2050387.

8. REFERENCES


