Optimal Reserve Prices in Weighted GSP Auctions: Theory and Experimental Methodology*

Yang Sun
City University of Hong Kong
Kowloon, Hong Kong SAR
sunyang@live.hk

Yunhong Zhou
Baidu Research
Beijing, China
zhouyunhong@baidu.com

Xiaotie Deng
City University of Hong Kong
Kowloon, Hong Kong SAR
csdeng@cityu.edu.hk

ABSTRACT

The Generalized Second Price Auction (GSP) is the dominant search-keyword auction mechanism of the past decade, and sponsored search ad auctions are one of the most successful Internet business models [4, 14]. The weighted GSP considers both the quality of each ad and also the corresponding bid in the auction mechanism, and most search engines use the weighted GSP to sell ad positions.

We show how to determine the optimal reserve price in the weighted GSP mechanism and we prove that the auction with this reserve price is a Myerson optimal auction [10]. The optimal price can be extended to support the CPA/CPC/CPM hybrid auction. Simulations show that setting a relatively high reserve price causes bidders to transfer more utility (surplus) to payment, resulting in higher revenue for the search engine. Finally, we describe a practical procedure for computing optimal reserve prices in production systems.

Categories and Subject Descriptors

K.4.4 [**Electronic Commerce**]: Payment schemes; K.6.0 [**Economics**]: General

General Terms

Economics, Electronic Commerce

Keywords

Generalized Second Price Auction, Weighted GSP, Reserve Price, Optimal Auction, Sponsored Search Auction

1. INTRODUCTION

Targeted advertising with search results generates billions of dollars of revenue annually and is a major factor in the successful commercialization of search engines [4, 14]. Targeted ads typically contain a headline, a body of text, and the advertiser's URL. Advertiser revenue increases when Internet users click on the ads and subsequently purchase advertised goods. Search providers conduct auctions to set advertising fees. Early keyword auctions employed the Generalized Frist Price (GFP) mechanism, which was used by Goto.com (later acquired by Yahoo! search). The major stages of keyword auction evolution[12] are:

- **GFP**: The unweighted, pay-per-click, Generalized First Price auction, used by Overture/Yahoo! from 1997–2002.
- **uGSP**: The unweighted, pay-per-click Generalized Second Price auction, used by Yahoo! from 2002–2007.
- **wGSP**: The weighted, pay-per-click, Generalized Second Price auction, used by Google, Yahoo! (since 2007), and Microsoft Bing .

The weighted GSP protocol, a variation of GSP, has been widely adopted across the industry in recent years, and has been studied extensively in academia. In weighted GSP, search engines charge advertisers according to their bids as well as ad quality factors (a variant of the predicted click-through rate).

Our analysis of weighted GSP auctions allows the search engine to set a reserve price for each ad for each keyword, assuming that each bidder knows her private value of the keyword, and only bidders who bid greater than their reserve prices can participate in the auction. When users search for the keyword, the search engine will rank the ads in descending order of advertisers' stated value per impression (or cost per mille, CPM), which can be computed by multiplying the advertiser's (per-click) bid by the advertiser's expected click-through rate (CTR). Instead of rank-by-bid in uGSP, we refer to wGSP as rank-by-revenue.

1.1 Related work

Myerson proposed the general optimal auction framework known as the classical Myerson optimal auction [10], which maximizes auctioneer profits. Myerson also determined the optimal reserve price for the classical single-unit auction, assuming that bidders' values are independent and identically distributed (IID).

Varian and Edelman et al. provide the foundation of our study: they define the keyword auction as a Nash equilibrium problem in sponsored search markets, and the refined equilibria are called *locally envy-free equilibria* [4] or *symmetric Nash equilibria* [14]. These two equilibria are equivalent; each advertiser bids at an envy free point where she is

*Work performed while the authors were at Baidu.

The 7th Ad Auction Workshop in conjunction with ACM EC 2011.

exactly indifferent between remaining in her current position and trading places with a bidder above her.

Edelman and Schwarz [5] generalized Myerson's theory to the multi-unit environment and simulated the case of GSP auctions. Ostrovsky and Schwarz [11] conducted a large scale experiment at Yahoo! Search with a new reserve price, which is around the mid-point between the theoretical optimal reserve price and the original uniform reserve price (10 cents), and reported that the new reserve price raised the average revenue per search almost 13%.

1.2 Our contribution

The previous theoretically optimal reserve prices were computed under the assumption that all bidders have the same quality scores and are ranked solely on their bids [5, 11]. We show how to determine provably optimal reserve prices in weighted GSP auctions where bidders have *different* quality scores, which is the situation in nearly all practical keyword auctions today. We also show that this optimal reserve price can also support CPA/CPC/CPM hybrid auctions.

The rest of this paper is organized as follows: Section 2 describes a typical weighted GSP mechanism. Section 3 describes how to set an optimal reserve price, proves that a Myerson optimal auction results, and extends this optimal reserve price to CPA/CPC/CPM hybrid auctions. Section 4 presents simulation results on search engine revenue, bidders' payment, and utility. Section 5 describes a practical technique to compute the optimal reserve price for each keyword and advertiser.

2. MODEL

There are k positions to be allocated among n bidders, with $k \leq n$. We denote the ad positions as $t \in \{1, \cdots, k\}$, and the bidders as $i \in \{1, \cdots, n\}$. We also denote v_i and b_i as the private value and bid of bidder i. We define a position factor $x_t \in (0, 1]$ and an ad quality factor $e_i \in (0, 1]$, where x_t denotes the probability that an ad in position t will be noticed by users, assuming that the position factor is decreasing by t, i.e., $x_1 > x_2 > \cdots > x_t > 0$; e_i denotes the probability of ad i will be clicked on if noticed, also called the ad's normalized (for position) CTR, which is common knowledge for all bidders.

We assume that the expected CTR of ad i in position t is $e_i \cdot x_t$, which can be separated into the ad factor e_i and position factor x_t. [1, 9]

We analyze a typical ranking rule where ads are ranked in descending order of score $e_i \cdot b_i$ ($e_1 b_1 > e_2 b_2 > \cdots > e_n b_n$). The price (per-click) of bidder i depends on the next bidder's score and his e_i:

$$\text{Price}_i = \frac{e_{i+1} \cdot b_{i+1}}{e_i} < \frac{e_i \cdot b_i}{e_i} = b_i \tag{1}$$

[14] proved that the symmetric Nash equilibrium (SNE) exists in this mechanism and in SNE the bidders' bids satisfy the follow expressions:

$$e_1 v_1 \geq \frac{p_1 x_1 - p_2 x_2}{x_1 - x_2} \geq e_2 v_2 \geq \cdots \geq e_i v_i \geq p_i$$

where $p_i = e_{i+1} \cdot b_{i+1}$.

Each bidder bids with the same strategy $\beta(\cdot)$ in SNE, i.e., $e_i \cdot b_i = \beta(e_i \cdot v_i)$. Bidder i can estimate his position (x_t) when he bids b_i:

$$x_t = \tilde{q}(e_i \cdot b_i) \tag{2}$$

Combining with the pricing function, the bidder's expected payment is

$$\begin{aligned} \text{Payment}_\text{i} &= \text{Price}_i \cdot e_i \cdot x_t = e_{i+1} \cdot b_{i+1} \cdot x_t \\ &= \beta(e_{i+1} \cdot v_{i+1}) \cdot \tilde{q}(e_i \cdot b_i) = \tilde{m}(e_i \cdot b_i) \end{aligned} \tag{3}$$

3. THE OPTIMAL RESERVE PRICE

An increase in reserve price can immediately lead to changes in those bids of advertisers that were previously bidding below the new reserve price, and makes further efforts to the above bidders, finally leading to a new equilibrium. We focus on the *perfect* SNE (the minimum revenue SNE corresponding to the VCG allocation) and find the optimal reserve price which maximizes the revenue of the perfect SNE.

We assume rational (individual-utility-maximizing) bidders who bid with the same strategy $\beta(\cdot)$ in equilibrium; a typical bidder i with value v_i will bid b_i. In this condition, his predicted CTR is $\text{pCTR}_i = e_i \cdot x_t = e_i \cdot \tilde{q}(e_i \cdot b_i)$. Thus, the bidder's utility formula:

$$U_i(b_i) = v_i \cdot pCTR_i - \text{Payment}_i = v_i \cdot e_i \cdot \tilde{q}(e_i \cdot b_i) - \tilde{m}(e_i \cdot b_i)$$

The derivative of U_i with variable b_i is:

$$\frac{\partial U_i(b_i)}{\partial b_i} = v_i \cdot e_i^2 \cdot q'(e_i \cdot b_i) - e_i \cdot m'(e_i \cdot b_i) \tag{4}$$

Since the strategy $\beta()$ is an equilibrium, thus $e_i \cdot b_i = \beta(e_i \cdot v_i)$ maximizes his utility, i.e.,

$$\begin{aligned} &\frac{\partial U_i(b_i)}{\partial b_i}|_{e_i b_i = \beta(e_i v_i)} \\ = \;& v_i \cdot e_i^2 \cdot q'(\beta(e_i \cdot v_i)) - e_i \cdot m'(\beta(e_i \cdot v_i)) = 0 \end{aligned}$$

Let *score* $s_i \equiv e_i \cdot v_i$, $\tilde{q}(\beta()) = q()$, $\tilde{m}(\beta()) = m()$, then

$$s_i \cdot q'(s_i) - m'(s_i) = 0$$

If the minimum score of all ads is $\underline{s}$, then its utility is zero, and it satisfies the restraint boundary condition:

$$U(\underline{s}) = \underline{s} \cdot q(\underline{s}) - m(\underline{s}) = 0 \tag{5}$$

Solving equations 5 and 5 yields

$$m(s) = s \cdot q(s) - \int_{\underline{s}}^{s} q(r) dr, \tag{6}$$

where s is a random variable drawn from a known distribution $F(s)$, with the probability density function $f(s) = F'(s)$ in $(0, \infty)$ [2]. The payment of a bidder is determined by his score s only: when $s < \underline{s}$, his utility is 0; and when $s \geq \underline{s}$, his utility is computed by the equation 6. The expected payment of bidder i is:

$$\begin{aligned} E[m(s)] &= \int_0^{\underline{s}} 0 \cdot f(s) ds + \int_{\underline{s}}^{\infty} m(s) \cdot f(s) ds \\ &= \int_{\underline{s}}^{\infty} s \cdot q(s) f(s) ds - \int_{\underline{s}}^{\infty} \left[\int_{\underline{s}}^{s} q(r) dr \right] f(s) ds \\ &= \int_{\underline{s}}^{\infty} s \cdot q(s) \cdot f(s) ds \\ &\quad - \left[\int_{\underline{s}}^{s} q(r) dr \cdot F(s) \right]_{\underline{s}}^{\infty} + \int_{\underline{s}}^{\infty} F(s) \cdot q(s) ds \\ &= \int_{\underline{s}}^{\infty} [s \cdot f(s) - 1 + F(s)] \cdot q(s) ds \\ &= \int_{\underline{s}}^{\infty} \left[s - \frac{1 - F(s)}{f(s)} \right] \cdot q(s) \cdot f(s) ds \end{aligned}$$

There are n bidders, each bidder's expected payment is given by the above expression, and the sum of their expected payments is the expected revenue of the search engine:

$$n \cdot E[m(s)] = n \int_{\underline{s}}^{\infty} \left[s - \frac{1-F(s)}{f(s)} \right] \cdot q(s) \cdot f(s) ds \quad . \quad (7)$$

There is also a probability $F^n(\underline{s})$ that the auction is a dud for this keyword, in this case the search engine gains a potential revenue t_0 of better user experience. Therefore, the expected revenue per search of the search engine is:

$$T(\underline{s}) = F^n(\underline{s}) \cdot t_0 + n \int_{\underline{s}}^{\infty} \left[s - \frac{1-F(s)}{f(s)} \right] \cdot q(s) \cdot f(s) ds$$

If the minimum score $\underline{s}$ makes the expected revenue $T(\underline{s})$ maximum, it satisfies:

$$\frac{\partial T(\underline{s})}{\partial \underline{s}} = n \cdot F^{n-1}(\underline{s}) f(\underline{s}) t_0 - n \left[\underline{s} - \frac{1-F(\underline{s})}{f(\underline{s})} \right] q(\underline{s}) \cdot f(\underline{s}) = 0$$

Then the optimal value of the minimum score $\underline{s}^*$ satisfies

$$\underline{s}^* - \frac{1-F(\underline{s}^*)}{f(\underline{s}^*)} = \frac{F^{n-1}(\underline{s}^*)}{q(\underline{s}^*)} \cdot t_0 \quad (8)$$

where $q(s)$ is the expected position factor when a bidder's score is s. When n bidders compete for k ad positions, a typical bidder i should win a position t between 1 and $\min(n, k)$ with probability:

$$q(s) = \sum_{t=1}^{\min(n,k)} \binom{n-1}{t-1} F(s)^{t-1} [1-F(s)]^{n-t} \cdot x_t \quad (9)$$

When the optimal minimum score $\underline{s}^*$ is resolved via equation 8, for bidder i with quality factor e_i, the optimal reserve price p_i^* for this keyword is

$$p_i^* = \frac{\underline{s}^*}{e_i} \quad (10)$$

3.1 Optimal Mechanism

We follow the wGSP mechanism described in section 2, and the bidders' values per impression ($e_i \cdot v_i$) are IID, drawn from a distribution $F(s)$ satisfying that $\frac{1-F(s)}{f(s)}$ is decreasing of s, denote the minimum score $\underline{s}^*$ as the solution of equation 8, and reserve price $p_i^* = \underline{s}^*/e_i$.

PROPOSTION 1. *A wGSP auction with a reserve price $p_i^* = \underline{s}^*/e_i$ for bidder i is an optimal mechanism.*

PROOF. The proof is based on Varian's SNE analysis [14], combined with the framework from Ulku that generalized Myerson optimal auction to multi-unit auction [13]. Ulku's result shows that an optimal multi-unit auction mechanism should have the following properties:

- It allocates units only to bidders whose virtual valuations are nonnegative.
- The sum of virtual valuations of all bidders is maximized.

The virtual valuation V of bidder i with value s_i per impression in position t is

$$V_{i,t} = x_t \cdot \left[s_i - \frac{1-F(s_i)}{f(s_i)} \right]$$

Since we allow only bidders with $b_i \geq p_i^*$ to participate in the auction,

$$s_i = b_i \cdot e_i \geq p_i^* \cdot e_i = \underline{s}^* \quad (11)$$

Since $\frac{1-F(s)}{f(s)}$ is decreasing of s, $s - \frac{1-F(s)}{f(s)}$ is an increasing function of s. According to equation 8,

$$s_i - \frac{1-F(s_i)}{f(s_i)} > \underline{s}^* - \frac{1-F(\underline{s}^*)}{f(\underline{s}^*)} = \frac{F^{n-1}(\underline{s}^*)}{q(\underline{s}^*)} \cdot t_0 \geq 0. \quad (12)$$

With the model description $x_t \geq 0$, it guarantees that each bidder's virtual valuation V is nonnegative. The sum of the winners' virtual valuations is:

$$\sum_{i,t} V_{i,t} = \sum_{i,t} x_t \left[s_i - \frac{1-F(s_i)}{f(s_i)} \right]$$

With the concept of SNE, bidder i's value s_i per impression is positively correlated with the position factor x_t, and winners' $s - \frac{1-F(s)}{f(s)}$ is an increasing function by s, that results the $s_i - \frac{1-F(s_i)}{f(s_i)}$ is positively correlated with x_t, obviously, this assures the sum of $x_t \cdot V_i$ is maximized. □

3.2 Extension to Hybrid Auctions

Besides CPC, two other charge schemes are widely used in keyword auctions [6]:

1. **CPM**, or cost per (thousand) impressions: The search engine charges the advertiser for every instance of an ad shown to a user.
2. **CPA**, or cost per action: The search engine charges the advertiser when a predefined action happens.

The three models are equivalent when precise estimates of the CTR and the action conversion rate per impression are known [6].

Some search engines allow CPC/CPM hybrid bidding, in which each advertiser may choose whether to bid for per-click (CPC) or per impression (CPM). Bids of both types compete for the same ad positions.

The search engine can convert a CPC bid b_i to estimated CPM bid m_i by multiplying his bid b_i with the quality factor e_i, and compete with those CPM bidders j who bid m_j per impression. Bidders will be ranked in descending order of their m_i, and charged depending on their bidding types: $\text{Price}_i = p_i/e_i$ per-click, or $\text{Price}_i = p_i$ per impression, where p_i denote the next bidder's stated value per impression ($p_i = b_{i+1} \cdot e_{i+1}$ or m_{i+1}).

Zhu and Wilbur [16] studied this case, and showed that any GSP equilibrium can be supported in the hybrid auction format. We assume the estimated eCPM is unbias and accurate enough, and with the proposition, the bidders in this hybrid auction satisfy the symmetric Nash equilibrium:

$$s_1 \geq \frac{p_1 x_1 - p_2 x_2}{x_1 - x_2} \geq s_2 \geq \frac{p_2 x_2 - p_3 x_3}{x_2 - x_3} \geq \cdots \geq s_i \geq \underline{s}$$

The optimal reserve price depends only on the distribution of the valuation per impression. With those conditions, the optimal minimum score $\underline{s}^*$ satisfies:

$$\underline{s}^* - \frac{1-F(\underline{s}^*)}{f(\underline{s}^*)} = \frac{F^{n-1}(\underline{s}^*)}{q(\underline{s}^*)} \cdot t_0$$

where $\underline{s}^*$ is the optimal reserve price for the CPM bidders, and $p_i^* = \underline{s}^*/e_i$ is the quality-factor normalized optimal reserve price for the CPC bidders in this hybrid bidding auction.

Google also allows advertisers to bid for conversions in AdWords. The CPA bidders can participate in this hybrid auction and compete with other bidders with different bidding types. As the same argument above, by introducing an estimated conversion rate α_i per impression, each CPA bid can be converted into an estimated CPM bid, and the optimal reserve price $p_i^* = \underline{s}^*/\alpha_i$ can be set to the CPA bidder i. In summary, the optimal reserve price result can be extended to this CPA/CPC/CPM hybrid auction.

4. SIMULATIONS

Search engines can increase per search revenue significantly by adjusting reserve prices to the optimal one. We simulate this auction process via Monte-Carlo method with different numbers of bidders, and calculate the search engine's revenue, each bidder's payment, and compare bidders' payment and utility. Considering the wGSP mechanism described in section 2, for a specific keyword, all bidders' valuations per impression (or $e_i \cdot v_i$) are drawn from the same *lognormal distribution*, which is derived from a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. The theoretical optimal reserve price (normalized) is 135 cents, which is shown as a dashed line in each figure.

We vary the reserve price from 0 to 1000 cents in one-cent increments, compute the average of the expected revenue per impression for each reserve price with 100,000 iterations, a randomizer generating bidders' values per impression according to the *pdf* function of the lognormal distribution:

$$f(s) = \frac{1}{\sqrt{2\pi}s\sigma} e^{\frac{-(\ln s - \mu)^2}{2\sigma^2}} \tag{13}$$

We assume that each bidder bids at the envy-free point, and plays a perfect SNE. We also assume the position factor ratio $x_{i+1}/x_i = 0.7$, which is a common knowledge for all bidders (the same as in [14]).

4.1 Revenue Effect

Figure 1 shows the per-search revenue variation on one keyword auction with different numbers of bidders. Each curve in the figure reflects the per-search revenue as a function of reserve price, from 1 bidder to 5 bidders participating in the auction, corresponding to bottom-to-top curves.

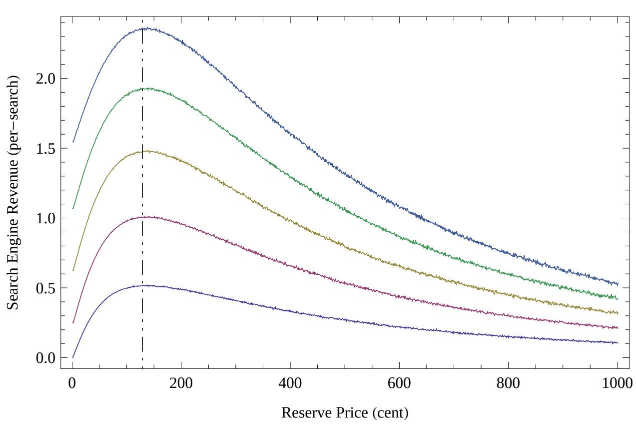

Figure 1: Search engine's revenue (per-search)

These curves represent the same optimal reserve price (around 135 cents, normalized), which indicates that the optimal reserve price for one keyword is independent of the number of bidders, and depends only on the distribution of bidders' valuations per impression, which is consistent with our theoretical result.

Table 1: Revenue increse

Number of bidders	Reserve price (normalized) 30 cents	135 cents (optimal)	Increase (%)
1	0.26577	0.51411	93.81%
2	0.61477	1.00834	63.92%
3	1.01315	1.47661	45.89%
4	1.43771	1.92383	33.95%
5	1.87221	2.35522	25.77%

The figure also indicates that the per-search revenue increase obtained by increasing reserve price from lower ones to the optimal one is substantial when few bidders bidding on this keyword; the increase is less pronounced when more bidders participate. Table 1 shows that along with the dense of bidders, the revenue increase rate actually declines.

4.2 Bidder's Payment

Figure 2 shows each bidder's expected payment as a function of reserve price when five bidders bid on one keyword. It actually decomposes the top (five-bidder) revenue curve in figure 1 into each bidder's payment, and demonstrates how the reserve price impacts each bidder's expected payment in this auction.

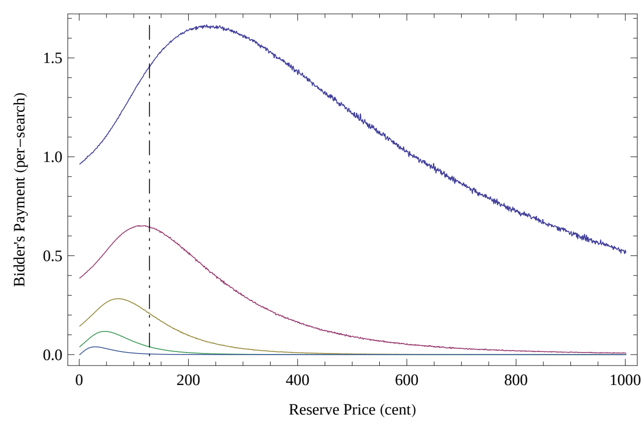

Figure 2: Each bidder's payment (per-search)

From the top to the bottom curve in figure 2, each curve represents one bidder's variation of expected payments, from the top-most ranked to the lowest ranked. The top-most bidder with the highest valuation per impression is ranked at the highest position and pays the most at any reserve price.

Each bidder's expected payment is impacted by the increase of the reserve price. Each represents a similar trend: a bell-shaped curve that reaches a payment peak and then declines to zero. The lowest bidder is most price-sensitive and reaches the peak first, followed by the second lowest, third lowest and the fourth lowest one. The top bidder is least price-sensitive and has not reached the peak when the reserve price exceeds the optimal price (135 cents).

4.3 Compare Revenue, Efficiency and Utility

Figure 3 shows the trend and relationship among the search engine's revenue, efficiency, and bidders' utility. The bell-shaped curve is the revenue curve, which is the same as the top one (5 bidders) at figure 1, the top-most curve is the sum of those 5 bidders' efficiency, and the remaining one is the total of those 5 bidders' utilities.

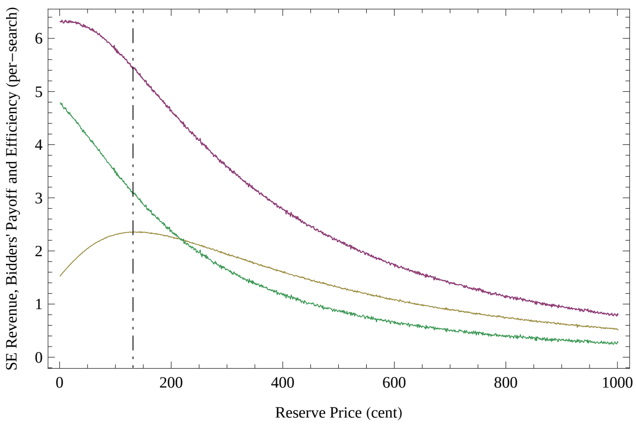

Figure 3: Search engine revenue, efficiency and bidders' utility (per-search) with 5 bidders

The bidders' efficiency [8] equals their payments (the revenue of search engine) plus their utilities. Both the bidders' efficiency and utility decrease with the increase of the reserve price.

Increasing reserve price is a non-zero-sum game between the search engine and the bidders, which decreases bidders' efficiency and forces them to transfer more utility to payment simultaneously, maximizing the search engine's revenue.

5. EXPERIMENT

Practical keyword auctions are complex, e.g., in terms of ad retrieval and matching algorithms. Search engines also consider advertiser budgets and other restrictions to determine rankings. Optimal reserve price theory is a simplified model of the real auction process because it abstracts away such complications. Prior works such as [15] provides a scalable way to compute appropriate reserve prices in complex, real-world settings, while using only input that are readily available in practice.

We collected a large data set from Baidu, the dominant search engine in China. Our data set includes more than 40 million keywords with bids in results page. For each keyword, we collect all bidders' bids and quality factors, and we classify keywords into disjoint industries.

This section describes a practical method for optimal reserve price computation. The process includes 3 steps: keyword semantic clustering, estimating the value distribution of bidders, and setting the reserve price.

5.1 Keywords Semantic Clustering

Before setting the reserve price, since we can never obtain the bidders' neutral private value distribution, we should estimate the value distribution parameters by computing bids for one specific keyword. We observe that the average number of bids for each keyword (also called market depth [5]) is very low, and many keywords have only one bid.

Statistical methods for distribution estimation do not work if only one sample is available and are not reliable with a small number of samples. We face a new challenge of how to estimate the value distribution when the bid sample is sparse, and that motivates us to use keywords semantic clustering before doing the bids statistics.

For example, ads bidding on the keyword "flower" may be shown at the results page of searching term "flowers express,“ due to broad match techniques based on users intention analysis and ads semantic matching. The value distribution of keyword "flowers express“ should apply to semantically related keywords that bid on "flower,“ because they comptete with not only bidders bidding for the same keyword, but also more bidders bidding on a group of similar keywords. Similarly, for keywords in the same industry, those close in semantics have similar competitors and reflect similar values, which is our fundamental idea.

Our algorithm computes the similarity degree between each candidate keyword pair in the keyword pool, following this general approach: for word-segments, stemming, and thesaurus replacements, each keyword pair's similarity degree is based on the OKAPI BM25 ranking algorithm [7]. If there are few bidders bidding on one keyword, then we find the top n similar keywords from the keywords pool in the same industry, according to the similarity degree, where the number of bidders bidding on these n keywords exceeds a threshold value (e.g., 10), and those bidders' bids will be used to estimate the value distribution of the target keyword.

5.2 Estimating the Value Distributions

We picked millions of keywords from dozens of industries, and for each keyword, we collect the bidders' and related bidders' score $(e_i \cdot v_i)$ via the above approach. We assume that bidders were playing a perfect SNE. Following Varian's method, we compute their $(e_i \cdot v_i)$ from the known $(e_i \cdot b_i)$, then compute two moments of $(e_i \cdot v_i)$: μ as the observed average, σ as the observed standard deviation. The position factors x_t should be calculated from the click and impression logs.

For different keywords, we assume the distributions of bidders' $e_i \cdot v_i$ fall into the same type of continuous distributions. The next step is to fit a typical continuous distribution for these $e_i \cdot v_i$. We select several candidate distributions $\mathrm{dist}_1, \mathrm{dist}_2, \cdots, \mathrm{dist}_d$ whose definition domains are in $[0, \infty)$, then compute their distribution parameters $\mathrm{dist}_j(\mu, \sigma, \cdots)$. As the bidders' scores are collected when a minimum score $\underline{s}$ is already set, this sample just reflects the distribution at $[\underline{s}, \infty)$. We compute the distribution parameters $\mu, \sigma, \cdots$ via *the method of moments* in $[\underline{s}, \infty)$. [1] Then compare the sample and each candidate distribution $\mathrm{dist}_j(\mu, \sigma, \cdots)$ via a distribution fit test method in $[\underline{s}, \infty)$ (e.g. the Pearson χ^2 test, Anderson-Darling test, Cramér-von Mises test), finally select the most likely distribution and its parameters μ, σ as the estimated distribution.

Table 2 shows an example of Cramér-von Mises test[3] (a small p-value suggests that it is unlikely that the data came from $dist_i$). *Empirically we found the lognormal distribu-*

[1]A more computation-intensive method uses maximum likelihood estimation to recover the truncated distribution, and it will be discussed in a followup paper.

tion fits better than other well-known distributions. Figure 4 presents the example valuation histogram and the fitted lognormal distribution.

Table 2: An example of Cramér-von Mises Test

Distribution	Parameters	Statistic	p-value
Lognormal dist.	[1.053,0.882]	0.13945	0.4234
Weibull dist.	[1.179,4.462]	0.39886	0.0728
Gamma dist.	[1.460,2.870]	0.42102	0.0636
Exponential dist.	[0.23864]	0.56207	0.0278

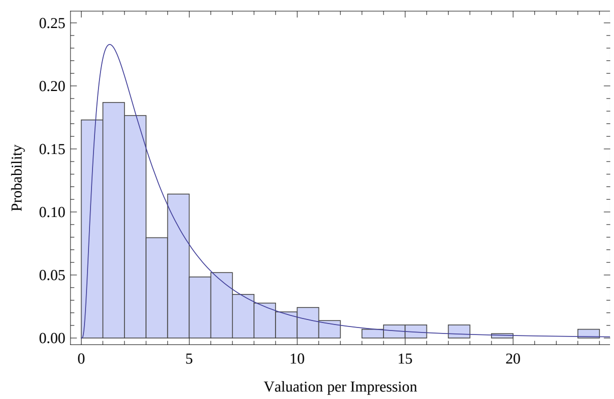

Figure 4: valuation histogram and fitted distribution

5.3 Setting the Reserve Price

The final step is to set the minimum score for each keyword and optimal reserve price for each bidder, according to the ideal distribution parameters. By testing sample keywords in each industry, we fit bidders' $e_i \cdot v_i$ bidding on each keyword with a lognormal distribution.

The potential revenue t_0 to holding any keyword is infinitesimal for search engines, and we set $t_0 = 0$. Combining with the lognormal expression (equation 13), the optimal minimum score $\underline{s}^*$ should satisfy $\underline{s}^* \cdot f(\underline{s}^*) = 1 - F(\underline{s}^*)$, which is simplified to

$$\underline{s}^* = e^{\mu+\sqrt{2}\sigma x^*}, \text{ where } \int_{x^*}^{\infty} e^{-t^2} dt = \frac{e^{-(x^*)^2}}{\sqrt{2}\sigma}. \tag{14}$$

$\underline{s}^*$ can be solved by numerical methods when the distribution parameters μ, σ are known. For each individual bidder i with quality factor e_i, his reserve price p_i^* is set as $\underline{s}^*/e_i$.

6. CONCLUSIONS

Keywords auction is a key part of the business model in huge search-engine companies, generating billions of dollars from these auctions every year. Examining the optimal reserve price for these GSP auctions is an extremely important problem.

Our work generalizes the optimal reserve price to wGSP auctions, and proves that the wGSP auction with this optimal reserve price is a Meyerson optimal auction. We prove that the optimal reserve price in wGSP auctions depends only on the distribution of bidders' valuations per impression, and is independent of the number of bidders. We further extend this theory to the CPA/CPC/CPM hybrid auction, and show that the optimal reserve price can be calculated similarly.

Simulations shed additional light on our theory and reveal a significant revenue increase by adjusting reserve prices to the optimal one. We also observed revenue and payment effects by increasing reserve price in simulation. Besides theory and simulation, we also propose a practical process to compute the optimal reserve price, and resolve some key issues via statistical techniques.

7. REFERENCES

[1] G. Aggarwal, A. Goel, and R. Motwani. Truthful auctions for pricing search keywords. In *EC*, 2006.

[2] J. Bulow and J. Roberts. The simple economics of optimal auctions. *J. Polit. Econ.*, 97(5), October 1989.

[3] D. A. Darling. The Kolmogorov-Smirnov, Cramér-von Mises tests. *Ann. Math. Stat.*, 28:823–838, 1957.

[4] B. Edelman, M. Ostrovsky, and M. Schwarz. Internet Advertising and the Generalized Second-Price Auction. *Amer. Econ. Rev.*, 97(1), 2007.

[5] B. Edelman and M. Schwarz. Optimal auction design and equilibrium selection in sponsored search auctions. *Amer. Econ. Rev.*, 100(2), May 2010.

[6] A. Goel and K. Munagala. Hybrid keyword search auctions. In *WWW*, 2009.

[7] K. S. Jones, S. Walker, and S. E. Robertson. A probabilistic model of information retrieval. In *Information Processing & Mgmt.*, 2000.

[8] S. Lahaie and D. M. Pennock. Revenue analysis of a family of ranking rules for keyword auctions. In *EC*, 2007.

[9] S. Lahaie, D. M. Pennock, A. Saberi, and R. V. Vohra. Sponsored search auctions. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, chapter 28. Cambridge U. Press, Sept. 2007.

[10] R. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1), 1981.

[11] M. Ostrovsky and M. Schwarz. Reserve prices in internet advertising auctions: A field experiment. In *EC*, 2011.

[12] D. R. M. Thompson and K. Leyton-Brown. Computational analysis of perfect-information position auctions. In *EC*, 2009.

[13] L. Ulku. Optimal combinatorial mechanism design. Working Papers 0903, Centro de Investigacion Economica, ITAM, 2009.

[14] H. Varian. Position auctions. *Int'l. J. of Indust. Org.*, 25(6), Dec. 2007.

[15] W. E. Walsh, D. C. Parkes, T. Sandholm, and C. Boutilier. Computing reserve prices and identifying the value distribution in real-world auctions with market disruptions. In *AAAI*, 2008.

[16] Y. Zhu and K. C. Wilbur. Hybrid advertising auctions. *Marketing Science*, 30:249–273, March 2011.