Empirical Price Modeling for Sponsored Search

Kuzman Ganchev, Alex Kulesza, Jinsong Tan, Ryan Gabbard, Qian Liu, and Michael Kearns

University of Pennsylvania
Philadelphia PA, USA
{kuzman,kulesza,jinsong,gabbard,qianliu,mkearns}@cis.upenn.edu

Abstract. We present a characterization of empirical price data from sponsored search auctions. We show that simple models drawing bid values independently from a fixed distribution can be tuned to match empirical data on average, but still fail to account for deviations observed in individual auctions. Hypothesizing that these deviations are due to strategic bidding, we define measures of "jamming" behavior and show that actual auctions exhibit significantly more jamming than predicted by such models. Correspondingly, removing the jamming bids from observed auction data yields a much closer fit. We demonstrate that this characterization is a revealing tool for analysis, using model parameter values and measures of jamming to summarize the effects of query modifers on a set of keyword auctions.

1 Introduction

Much of the academic literature on sponsored search to date has been theoretical in nature [7, 4–6], characterizing behavior or payoffs under strong assumptions that may fail in practice—especially when the markets in question are young, bidders are inexperienced, and relevant pieces of information (such as the exact rules of the game) are frequently kept secret. To our knowledge there has been no exploratory study of actual bidding data on a large scale to determine how real-world auctions can best be analyzed and understood. This paper provides a simple but needed first look at such questions. We utilize sponsored search data drawn from a wide array of Overture/Yahoo! auctions and examine how bids are distributed, what kinds of models of advertiser value can reasonably be proposed, and the evidence for strategic behavior.

Our analysis serves two immediate purposes. First, a better understanding of empricial bidding behavior improves the quality of data that can be synthetically generated for further study. We show that simple models used in practice fail to account for significant strategic effects, and suggest improvements that meaningfully enhance the "realism" of such models. Second, our characterization of sponsored search auctions includes measurable quantities and model parameters that can be used to summarize important features of an auction for further analysis. To demonstrate the insight provided by such summaries, we show how groups of query modifiers can influence bidding on a wide array of keyword auctions. We find, for example, that adding modifiers like "cheap" or "deal" to automobile brand names tends to increase the amount of bid jamming at the first slot of the corresponing sponsored search auctions.

2 Methodology

Our data was obtained from the Overture bidview tool from approximately November 28 to December 2, 2006[1]. We collected bid data for two sets of queries. The first, smaller set includes the keywords used by Rusmevichientong et al [6] and comprises 859 queries related to travel. It is used for aggregate analysis in Section 2.1. The second set, used in the latter sections of the paper, comprises a wide array of 36,900 queries. For the purposes of further analysis (such as that carried out in Section 5), the second set is structured as a cross product of 450 base keywords—e.g., "lawyer"—intended to reflect basic searches that would generate advertiser interest, and 81 modifiers—e.g., "Philadelphia"—intended to capture the ways in which users might further specify searches. The base keywords and modifiers are further structured by placement in

[1] `http://www.overture.com/`; the bidview tool was discontinued shortly after we collected our data.

groups; there are nine groups of base keywords and six groups of modifiers. Table 1 and Table 2 give some summaries and examples for the groups of keywords in this data set. A complete query in the second set pairs one base keyword with zero or one modifiers (e.g., "lawyer" and "Philadelphia lawyer").

Due to resource constraints, we ran each query only once; our data provides no information on dynamic bidding behavior. The data returned by the bidview tool include up to 40 bids, ranked in order from highest to lowest. Advertiser names and ad text are provided, but not used for our analysis. Furthermore, we throw out the first bid in every auction. This is due to the method by which prices are determined: an advertiser pays a price equal to the bid of the next advertiser in bid order, so that the first bid is not relevant to money changing hands except insofar as it is higher than all other bids. For the remainder of the paper we use the term "price" assuming the convention that the k^{th} price is equal to the $(k+1)^{th}$ bid.

Table 1. Summary of base keyword groups.

group	#	examples
cars	41	BMW, Toyota
drugs	62	Xenical, Prozac
electronics	36	laptop, cell phone, camcorder
local-service	55	carpet cleaning, hair dresser
medical	50	anxiety, plastic surgery
non-local-service	27	car insurance, mortgage
software	67	Microsoft Windows, MySQL
subscription	91	cable, magazine
travel	21	cruise, hotel, vacation
total	450	

Table 2. Summary of modifier keyword groups.

group	#	examples
action	6	buy, purchase
info	11	information, review
location	40	New York, Ohio, Philadelphia
post	6	support, parts, repair
price	10	cheap, expensive, free, discount
quality	8	best, luxury, new, used
total	81	

2.1 Initial Look at the Data

Visualizing the bid books, it is apparent that the data are generally quite noisy. Consequently, we begin by examining the data in aggregate using our smaller, travel domain query set. Figure 1 shows how the price paid by the top bidder is correlated with the total number of bidders in the auction. There is an almost linear relationship between the number of bidders and the mean price of the first position. Figure 2 shows the correspondence between bid position and mean prices. Because prices increase as the number of bidders increases, all the auctions used in computing Figure 2 have at least 19 bids. We normalize the prices so that the first price is 1.0. We find that an exponential decay fits the means surprisingly well, and as a result an exponential model seems a natural choice for fitting individual auctions. We will see, however, that individual auctions show significantly different behavior than the aggregate. Figure 3 shows the differences between the first and second price, the second and third price and the third and fourth price for auctions with at least 19 bids. We will see in Section 4 that the peak in the small price difference ranges may be due to "jamming" strategies.

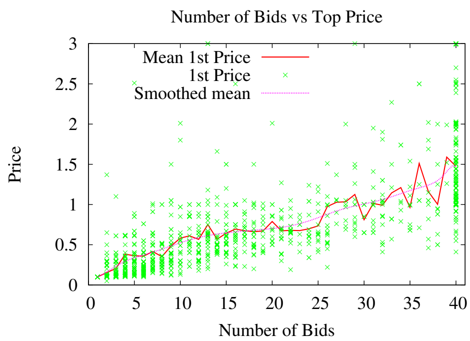

Fig. 1. Correlation between first price and number of bids.

3 Independent Bidding Models

Section 2.1 motivates a simple approach to modeling individual auctions under the assumption of bidder independence. If bid averages follow an exponential decay, independent bidders must be drawing from the unique distribution that yields this curve. In particular, bids can be simulated by sampling prices $p(u)$ where

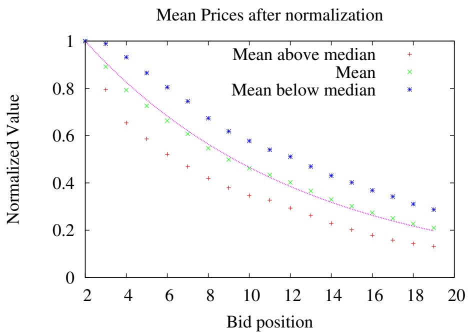

Fig. 2. Correlation between prices and bid position. The curve is an exponential fit to the means.

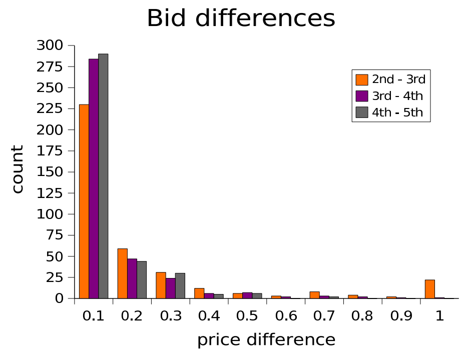

Fig. 3. Bid differences corresponding to differences in price.

Table 3. MSE (*100).

Model	Real	Gauss.	Unif.	Exp.
ab^i	3.8	1.9	3.7	1.7
$ab^i + c$	2.9	1.2	1.3	1.0
linear	14.2	3.2	1.5	7.8
2-linear	1.2	0.6	0.5	0.6

Table 4. Normalized MSE (*100).

Model	Real	Gauss.	Unif.	Exp.
ab^i	9.5	5.7	8.0	7.4
$ab^i + c$	5.9	4.3	3.2	3.0
linear	36.8	15.2	4.2	34.1
2-linear	2.1	1.8	0.9	1.6

Table 5. MAE.

Model	Real	Gauss.	Unif.	Exp.
ab^i	0.10	0.08	0.10	0.07
$ab^i + c$	0.06	0.06	0.06	0.05
linear	0.18	0.09	0.07	0.14
2-linear	0.05	0.04	0.04	0.04

continuous function $p(\frac{k-1}{N-1})$ is the expected price of a bidder who is at position k out of N and u is drawn uniformly at random from $[0, 1]$. For example, $p(u) = ae^{-bu}$ yields an exponential curve like that in Figure 2.

While this method faithfully reproduces aggregate price curves, we show here that it does not realistically generate individual bid books. We take measurements using a variety of simple parametric auction models, comparing the quality of each model's best fit to real data and to synthetic data. If the synthetic data are accurate, then the models should fit both data sets equally well. If there is some consistent difference in the quality of fits, then we can conclude that the generative prodecure above is not realistic.

We fit using both exponential and piecewise linear models. Our exponential models have two and three parameters, taking the forms price$[i] = ab^i$ and price$[i] = ab^i + c$, where i is an index of the price positions. Note that the second version allows prices to converge to a nonzero reserve price. We also fit piecewise linear models using dynamic programming. We report only the results for 1 and 2-piece models as MSE drops nearly to zero when 3 or more pieces are used. In all cases, the fits are performed on a per-auction basis to minimize the mean squared error (MSE) of the predicted prices relative to the observed prices. The mean is weighted so that each auction receives equal weight regardless of the number of bids it attracts. We also report normalized MSE (where the highest price in each auction is normzalized to 1.0) and mean absolute error (MAE), computed by averaging the absolute instead of squared differences between predicted and observed prices.

For each auction in our second query set we generated parallel synthetic price data using the sampling technique described above, where $p(\cdot)$ was chosen to be the best exponential fit to the real auction data. Note that the synthetic data is independently sampled from the induced distribution, thus it is not artificially smooth or guaranteed a better fit. We also sampled synthetic bids from Gaussians with mean and variance equal to the empirical prices and from a uniform distribution over the range of empirical prices. These correspond to methods commonly used in practice. Each of the four data sets was fit by each of the four parametric models. The results are presented in Table 3, Table 4, and Table 5.

It is apparent that none of the synthetic methods for generating data displays fit error rates similar to those of real auction data. However, it seems clear that the fits for exponential synthetic data are most similar to those of true auction data, though in magnitude the error measures are much smaller. We conclude, therefore, that the exponential model is the most accurate (as predicted by the aggregate analysis), but that simple, independent-bidder models miss certain key characteristics of sponsored search auctions.

4 Jamming

Figure 4 expands the zero to ten cents region of Figure 3. It is clear that the large number of bid differences below ten cents is due to a peak at differences of one cent or less. We propose that the sharp peak around one cent in the price differences, as well as the non-independence of real auction data, may be due in part to the use of a bidding tactic known as "jamming." Jamming involves bidder A bidding just below bidder B in order to increase B's price while leaving A's price unchanged. This may be an attempt to deplete B's advertising budget or to convince B to drop its bid. This tactic appears to be widely used and is indeed implemented automatically by many bidding packages [3, 2].

Of course, jamming is not the only possible explanation for the unusual number of one cent bid gaps. It is possible that bidders choose to play a strategy in which bids are set to one cent *above* the next lowest bidder in an effort to avoid being jammed, that other bidding strategies create the observed effects indirectly, or that collusion is somehow encouraging clumped bidding. Going forward we will continue to use the term "jamming," but we will define it as a purely statistical measure of bid closeness. Further experiments are necessary to draw valid conclusions about bidder intentions.

Our definition of jamming is as follows: a bid is a jamming bid if it is one cent or less below the next highest bid. We call a series of consecutive bids each of which (except the first) jams the previous one a "jamming region." To reduce the influence of compression artifacts due to the reserve price, we do not consider bids within two cents of the reserve. By comparing measurements of jamming in real and independent-bidder synthetic data, we seek to tease apart *true* jamming (jamming that results from dependencies between bidders) and apparent jamming due merely to the chance clustering of bids, which may be significant in a market with dozens of bids all less than a dollar.

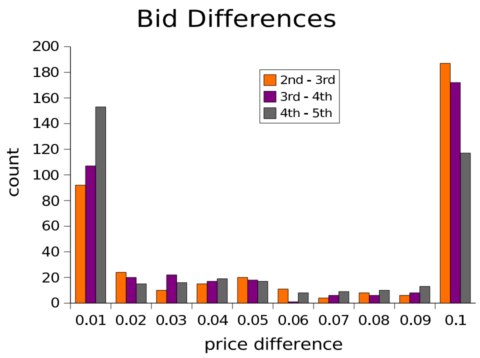

Fig. 4. Bid differences in $0.01 increments. Bars marked "0.1" include all bid differences of $0.10 or greater.

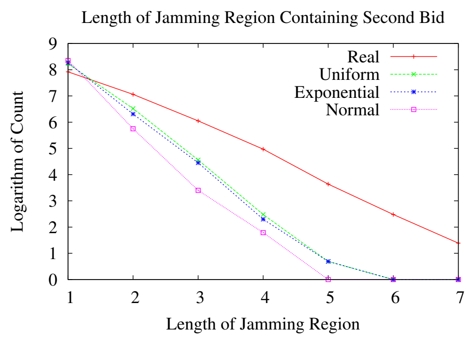

Fig. 5. Counts of the length of the jamming region containing the second bid, on a logarithmic scale.

Figure 5 shows the lengths of the jamming regions starting at the second bid. Intuitively, this corresponds to the number of people jamming the first price point. For real data, the length of the first jamming region is nearly exponentially distributed, and falls off far more slowly than for the simulated data.[2] Figure 6 shows the frequency of jamming by position. There is a general upward trend for all distributions, probably due to the increasing compression of the range of possible bid values at lower positions creating more chance clusters. However, across a range of price points that there is a strong and statistically significant jamming effect in the real data over and above that seen in the independent-bidder models. Figure 7

[2] All differences between the real and simulated data here are significant at 95% confidence.

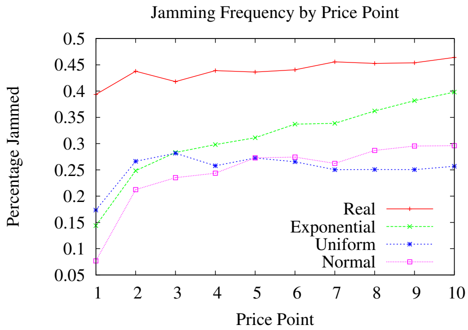

Fig. 6. Frequency of jamming by price position.

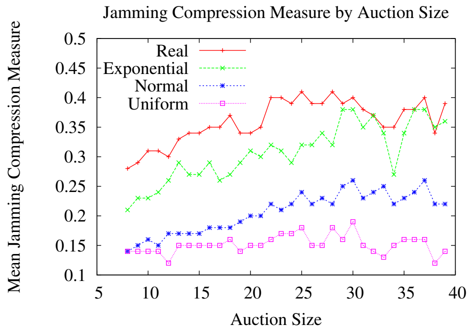

Fig. 7. Jamming ratios by auction length.

shows the jamming ratio, computed as the number of jamming bids over the total number of bids in the auction, versus auction size. As expected, jamming ratios increase gradually for all distributions, due again to chance clustering as more bids are crammed into a small range. Although this data is noisy due to the relatively small number of auctions with certain sizes, the jamming ratio is again significantly higher for real data than exponential data on all auction sizes less than 27 except size 20.

We recomputed our parametric model fits after removing jamming bids from the real auction data. The MSE results are presented in Table 6. Though the gap between synthetic and real data still exists, it has been closed considerably. We hypothesize that further improvements might be obtained by considering other common strategies employed by advertisers. Using similar techniques, it should be possible to formalize and measure the prevalence of such strategies in real-world auction data.

Table 6. MSE results (*100).

Model	Real	Unjam	Exp
ab^i	3.8	2.8	1.7
$ab^i + c$	2.9	1.9	1.0
linear	14.2	13.1	7.8
2-linear	1.2	0.8	0.6

5 Measuring the Effects of Modifiers

Since the queries in our second data set are structured as cross products of keyword/modifier groups, we can visualize interesting trends using the characterization we have developed. We pre-process the data by removing auctions with fewer than 11 bids in order to reduce noise, and display quantites of interest in a series of tables. Each row in a table corresponds to a base group of keywords, each column corresponds to a modifier group or an empty ("null") modifier group. (The groups are described in Section 2.) A cell shows the mean of a particular quantity for queries formed from the cross product of the corresponding groups. If a cell is shaded, the mean is statistically different from the "null" entry in the same row; red/bordered cells indicate an increase, and green/unbordered cells indicate a decrease.

Figure 8 shows the effect of keyword modifiers on the first price parameter for the exponential model (a). The table shows that modifiers from the group "location" significantly increase the highest price paid for the four base groups of "local service", "medical", "subscription," and "cars." This seems natural, as all of these groups involve products that frequently depend on local providers. Conversely, "non-local service" keywords have their prices reduced by the addition of a location, as expected. Figure 9 shows the decay rates for the exponential model (b).

We can also visualize interesting trends using the understanding of jamming developed in Section 4. Figure 10 shows that while adding modifiers in general reduces jamming, adding a location or an action word like "buy" can significantly increase the amount of jamming. Figure 11 shows that adding modifiers such as "cheap" to the name of an automobile manufacturer increases the expected length of the jamming region beginning with the second bid.

	null	act	info	loc	post	price	quality
cars	0.97	1.09	0.88	1.19	1.55	0.9	0.88
drugs	1.31	1.22	1.11	-	-	0.78	-
electr	1.04	0.85	0.39	1.9	0.7	0.78	0.8
local	1.45	0.94	0.99	3.05	1.36	1.18	1.32
med	1.93	-	2.03	3.66	2.19	0.65	-
n-loc	4.76	2.47	1.98	3.58	1.48	2.67	2.65
soft	0.86	0.48	0.42	-	-	-	0.48
subscr	0.87	0.97	0.59	1.47	1.05	0.9	0.83
travel	0.92	0.54	0.53	0.92	-	0.95	1.3

Fig. 8. First price parameter of the exponential model.

	null	act	info	loc	post	price	quality
cars	0.93	0.9	0.87	0.89	0.83	0.88	0.9
drugs	0.86	0.84	0.76	-	-	0.8	-
electr	0.91	0.87	0.88	0.8	0.86	0.88	0.86
local	0.9	0.83	0.82	0.8	0.83	0.86	0.82
med	0.84	-	0.75	0.73	0.81	0.84	-
n-loc	0.9	0.83	0.83	0.89	0.77	0.84	0.86
soft	0.87	0.84	0.88	-	-	-	0.87
subscr	0.88	0.85	0.85	0.8	0.81	0.86	0.86
travel	0.92	0.88	0.9	0.88	-	0.89	0.86

Fig. 9. Decay rate parameter of the exponential model.

	null	act	info	loc	post	price	quality
cars	0.19	0.25	0.2	0.26	0.08	0.23	0.2
drugs	0.16	0.16	0.09	-	-	0.08	-
electr	0.11	0.06	0.01	0.09	0.04	0.06	0.05
local	0.15	0.05	0.06	0.11	0.06	0.09	0.07
med	0.09	-	0.07	0.08	0.08	0.07	-
n-loc	0.27	0.12	0.14	0.23	0.04	0.17	0.19
soft	0.03	0.05	0.02	-	-	-	0.0
subscr	0.08	0.08	0.05	0.12	0.03	0.08	0.09
travel	0.11	0.02	0.05	0.12	-	0.11	0.13

Fig. 10. Fraction of jammed bids over 40 cents.

	null	act	info	loc	post	price	quality
cars	1.98	2.49	2.34	2.47	1.85	2.65	2.11
drugs	2.23	2.1	1.94	-	-	2.04	-
electr	1.47	1.22	1.56	1.78	1.69	1.41	1.38
local	1.69	1.37	1.37	1.6	1.5	1.83	1.53
med	1.29	-	1.26	1.34	1.83	1.2	-
n-loc	1.52	1.39	1.57	1.61	1.0	1.35	1.43
soft	1.33	2.0	1.4	-	-	-	1.0
subscr	1.65	1.3	1.79	1.68	1.17	1.63	1.64
travel	1.68	1.8	2.06	1.57	-	1.5	1.47

Fig. 11. Length of jamming region starting at first price.

6 Conclusion

We collected a large set of empirical sponsored search data and performed an exploratory analysis, attempting to characterize and understand real-world search auction data. We found an aggregate exponential decay of prices across many auctions, but showed that this model does not fully describe bidding behavior on a per-auction basis. We showed that jamming is more prevalent in real data than would be predicted by a model of independent bidders, and that removing jamming from empirical data (or, convesely, adding jamming to synthetic data) improves the similarity significantly. Future work will include studying effects other than jamming that contribute to this disparity. Finally, we demonstrated that our model parameters and measures of jamming provide useful summaries of important auction features, revealing trends in the ways modifiers influence the bids for search keywords.

References

1. A. Animesh, V. Ramachandran, and S. Viswanathan. Quality Uncertainty and the Performance of Online Sponsored Search Markets: An Empirical Investigation. *SSRN eLibrary*, 2006.
2. ApexPacific. Overture bid management software & overture bid tool. `http://www.apexpacific.com/bidmaximizer/overturebidding.html`.
3. Atlas. Rules-based bidding for pay-per-click management. `http://www.atlasonepoint.com/products/bidmanager/rulesbased`.
4. B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. Working Paper, http://rwj.berkeley.edu/schwarz/, 2005.
5. S. Lahaie. An analysis of alternative slot auction designs for sponsored search. In *EC '06: Proceedings of the 7th ACM conference on Electronic commerce*, pages 218–227, New York, NY, USA, 2006. ACM Press.
6. P. Rusmevichientong and D. P. Williamson. An adaptive algorithm for selecting profitable keywords for search-based advertising services. In *Electronic Commerce '06*, pages 260–269, New York, NY, USA, 2006. ACM Press.
7. H. R. Varian. Position auctions. *International Journal of Industrial Organization (to appear)*, 2006.