Computing Optimal Bundles for Sponsored Search

Arpita Ghosh[*] Hamid Nazerzadeh[†] Mukund Sundararajan[‡]

Abstract

A *context* in sponsored search is additional information about a query, such as the user's age, gender or location, that can change an advertisement's relevance or an advertiser's value for that query. Given a set of contexts, advertiser welfare is maximized if the search engine runs a separate auction for each context; however, due to lack of competition within contexts, this can lead to a significant loss in revenue. In general, neither separate auctions nor pure bundling need maximize revenue.

With this motivation, we study the algorithmic question of computing the revenue-maximizing partition of a set of items under a second-price mechanism and additive valuations for bundles. We show that the problem is strongly NP-hard, and present an algorithm that yields a $\frac{1}{2}$-approximation of the revenue from the optimal partition. The algorithm simultaneously yields a $\frac{1}{2}$-approximation of the optimal welfare, thus ensuring that the gain in revenue is not at the cost of welfare. Finally we show that our algorithm can be applied to the sponsored search setting with multiple slots, to obtain a constant factor approximation of the revenue from the optimal partition.

1 Introduction

Sponsored search is a very effective medium for advertising as it allows precise targeting of advertisements to users: a user can be presented with advertisements that are directly related to her search query. However, further targeting is possible by using the *context* of a query and the user associated with the query. A *context* in a sponsored search auction is additional information associated with a particular instance of a query that can change an advertisement's relevance or an advertiser's value for that keyword. For example, zip codes can often be inferred from IP addresses, providing a user location context: for certain queries (say pizza delivery, or dentist) local advertisements might be more relevant to the user than non-local ones. Other examples of contexts are age or gender-related demographic information, or 'search intent' gleaned from other searches by the same user.

The formal study of sponsored search with contexts was recently introduced by Even-Dar et al [10], where the authors showed that splitting a keyword auction into mutiple auctions, one for each context (for example, if the context is location, then having one auction for each location), increases welfare. They also gave examples demonstrating that there is a tradeoff: while welfare increases upon splitting contexts, the search engine's revenue may be larger when the keyword is not split (*i.e.* , all contexts stay combined). (To see why, consider the case when the auction for each context has only one participating advertiser; since the mechanism used is a a variant of second price auctions [16, 8], such advertisers face no competition and will generally pay a small reserve. So the revenue to the search engine is very small compared to the situation when contexts are not separated.) However, the search engine's choice is *not limited to the two extreme partitions* of the set of contexts, namely, keeping each context separate (maximizing efficiency) or combining all contexts together (pure bundling): other partitions of the set of contexts may give better points on the

[*]Yahoo! Research, 2821 Mission College Blvd, Santa Clara, CA 95054. Email: arpita@yahoo-inc.com

[†]Management Science and Engineering Deparment, Stanford University, Stanford, CA 94305. Work performed in part while visiting Yahoo! Research. Email: hamidnz@stanford.edu.

[‡]Department of Computer Science, Stanford University, Stanford, CA 94305. Supported by a Stanford Graduate Fellowship. Email: mukunds@cs.stanford.edu.

revenue-efficiency trade-off curve (in fact, we will show that the revenue from the optimal partition can be arbitrarily larger than the revenue from these two extreme partitions, while losing no more than half the maximum efficiency.)

In this paper, we study the algorithmic problem of optimally partitioning a set of contexts to maximize revenue under a second-price mechanism in the full information setting, *i.e.* , when the matrix of bidder valuations for each context is known (for simplicity, we deal with the case of a single slot, although we later discuss generalizing the results to multiple slots). We show that this problem is strongly NP-hard, and then provide a 1/2-approximation algorithm for it. This approximation algorithm also loses no more than 1/2 the maximum possible efficiency (obtained when all contexts are auctioned separately, possibly with great loss in revenue). We emphasize that since the optimal revenue can be arbitrarily larger than the revenue from either selling all contexts separately or combining them all together, the revenue from this algorithm can also be arbitrarily larger than the natural benchmark revenue; the factor 1/2 is with respect to the optimal revenue over *all* partitions, not the larger revenue of the two extreme partitions. Finally we show that our algorithm can be applied to the sponsored search setting with multiple slots to obtain a factor $\frac{(1-q)}{2}$ of the optimal revenue, when slot clickthrough rates decay geometrically [6] as q^j.

We consider the full information setting for the following reason. A search engine making the decision to split contexts might want to compute a partition of contexts into auctions just once (or infrequently), rather than dynamically [1]. In this case the search engine will use observed historical data to compute these partitions. Specifically, consider keywords where the value-per-click remains almost constant across contexts, with only click-through rates varying across context. Search engines usually have reasonable estimates of click-through rates across contexts, and also of valuations of advertisers who bid frequently on a keyword (while the GSP auction is not truthful, techniques from [16] can be used to obtain estimates of values). Further, advertisers' actual valuations for keywords do not vary significantly over time, as these values are typically based on the estimated profit from future conversions. Thus it is reasonable and possible to use a partitioning algorithm based on full information in this setting [2]. A similar reasoning applies to another situation where bundling may be valuable, which is for related keywords with thin markets (for instance, bundling together misspellings of a valuable keyword like insurance, where each misspelling might have bids from only a few advertisers).

Related work: The study of bundling in the economics literature was started by Palfrey [15], and later extended to various settings [7, 5, 4]. Recently, Jehiel et. al. [11] proposed a novel framework to study mixed bundling auctions, and proved that under certain distributional assumption over valuations, neither bundling all items together, or selling them separately is optimal. Another related paper in this context is [12], which studies high revenue auctions from the class of virtual valuations combinatorial auctions, and gives an auction which is within a logarithmic factor of the revenue maximizing auction for additive valuations. Bundling has also been studied in the setting of monopoly pricing [1, 14, 13, 3]. Our work differs from all of this literature in that we consider the algorithmic problem of computing the optimal, revenue maximizing partition under a second price mechanism in the full information setting; we give a constant factor approximation for this problem, *along with* an efficiency guarantee.

A different solution for revenue maximization in thin markets is to set a reserve price based on estimates of distributions of advertiser valuations [9]. Bundling is a more robust solution when bidders' values (or distributions of bidder values) change with time, but in a positively correlated fashion, such as temporal or seasonal variations (prominent examples are keywords related to travel, or occasions such as Valentine's day (like flowers)). In such cases the same bundling structure can be maintained as opposed to optimal reserve prices which will need to be updated to maintain high revenues.

[1] In fact, it is not clear what it means to dynamically compute bundles of contexts, since contexts may not appear simultaneously.

[2] Note that once bundles have been computed, the usual equilibrium analysis of a keyword auction can be applied to each bundle.

2 Model

There is a set I of items numbered $1 \dots k$, and and a set U of agents, $1, \dots, n$. There is a single copy of each item (we discuss the multiple-slot case in §5). Let v_{ij} be the value that agent i has for item j. We assume that agents valuations for bundles are additive—the value that agent i has for a bundle $B \subseteq I$, v_{iB}, is $\sum_{j \in B} v_{ij}$.

Items, or bundles of items, are sold according to a second-price auction: the winner of a bundle B is the agent with $\max v_{iB}$, and is charged the second highest valuation for that bundle. An allocation partitions I across the bidders. Let S_i denote the set of items allocated to bidder i. The welfare of the allocation is $\sum_{i \in U} v_{iS_i}$, and the revenue is the sum of the prices paid by each winning agent.

Our problem is the following: Given the matrix of valuations v_{ij}, we want to compute the revenue maximizing partition of items into bundles, when each bundle is awarded to the agent with the highest valuation at a price equal to the second highest valuation for the bundle.

We briefly discuss how the sponsored search setting maps to the above model. Suppose there are n bidders, where bidder i, $1 \le i \le n$, has a value per click v_i. Assume that there is just one slot. Suppose there are k different contexts, and the clickthrough rate (CTR) of bidder i for context j is c_{ij}—this is the probability that the advertisement of bidder i will be clicked on when displayed in context j. Further, let f_j denote the number of impressions corresponding to a specific context. The value that advertiser i has for context j is $v_{ij} = v_i f_j c_{ij}$. We assume that valuations are additive, i.e. the valuation for a set of contexts I' is $v_i \cdot \sum_{j \in I'} f_j c_{ij}$. We note that our model is quite general and can also be applied to bundling different keywords together or the case that different context has different values.

3 Characterizing optimal bundling

An *optimal bundling* is a partitioning of items into bundles that leads to the largest revenue, when items are allocated to the agent with the highest valuation for the bundle at a price equal to the second-highest valuation. In this section, we characterize the structure of bundles in an optimal bundling, and show that bundling to maximize revenue does not lose much efficiency.

Before we discuss optimal bundling, it is natural to ask whether it is sufficient merely to consider two extreme partitions: sell all items separately, or bundle them all together (in fact, much prior work on bundling restricts itself to these two options). However, the larger of the revenues from these two extreme partitions can be arbitrarily worse than the revenue of the optimal bundling, as the following example shows. The same example shows that the efficiency loss can also be arbitrarily large when restricting the choice to these extreme partitions.

Example 1. *Suppose $n = k$, i.e. , there are k items and k agents. The valuation of bidder i is 1 for item i, and 0 for all other items. If all items are sold separately, the revenue is 0 (and welfare is k). Bundling them together gives a revenue of 1 (and welfare 1). However, the revenue of optimal bundling is $\frac{k}{2}$, which is obtained by pairing items, i.e. , partitioning into $\frac{k}{2}$ bundles; this also has welfare $\frac{k}{2}$. Thus choosing between these two options to maximize revenue can lead to revenue and efficiency that are both arbitrarily worse than the optimal revenue and optimal efficiency.*

The following facts follow easily from the above example. (Note that maximum efficiency is always obtained when selling all items separately.)

Fact 1. *An efficiency-maximizing bundling with the highest revenue does not, in the worst case, give a c-approximation of revenue for any constant $c \ge 0$.*

Fact 2. *A revenue-maximizing bundle with the highest efficiency does not, in the worst case, yield better than a $\frac{1}{2}$-approximation of welfare.*

(Note that the revenue maximizing bundle is not unique, and efficiency can vary across optimal partitions: suppose there are 2 items, and 3 bidders with valuations $(10, 0)$, $(5, 5)$, and $(0, 10)$. Both partitions yield the maximum revenue of 10; however one has welfare 20 and the other has welfare 10.)

We will now show, in Theorem 2, that the statement in Fact 2 is tight. Let h_j be the highest valuation for item j, i.e. $h_j = \max_i\{v_{ij}\}$, and let s_j be the second highest valuation for item j. We state the following fact without proof. Consider a bundle B in an optimal bundling. If there is an item j that can be removed from B with no decrease in revenue, the new bundling obtained by selling j separately from B is an optimal bundling with weakly greater efficiency. Note that this lemma implies that in any bundle B (in a revenue-optimal bundling with highest efficiency) with two or more items, we can assume that $h_j > s_j$ for $j \in B$.

Lemma 1. *Consider a bundle B in an optimal bundling with highest efficiency. If bidder i has the highest valuation for item j in bundle B, then i has either the highest valuation or the second highest valuation for bundle B.*

Proof. Let i_1 and i_2 be the bidders with the highest and second highest valuations for bundle B, and consider any item $j \in B$. If j is removed from the bundle B, the revenue from $B \setminus \{j\}$ is at least $\min(v_{i_1B} - v_{i_1j}, v_{i_2B} - v_{i_2j}) \geq v_{i_2B} - \max(v_{i_1j}, v_{i_2j})$. Since it is strictly beneficial to bundle j with $B \setminus \{j\}$, we must have

$$v_{i_2B} > v_{i_2B} - \max(v_{i_1j}, v_{i_2j}) + s_j \;\Rightarrow s_j < \max(v_{i_1j}, v_{i_2j}),$$

which proves the claim. □

Theorem 2. *An optimal bundling with the highest efficiency also gives a $\frac{1}{2}$-approximation of welfare.*

Proof. Consider a bundle B in such an optimal bundling, and let i_1 and i_2 be the bidders with the highest and the second highest valuation for the bundle. Since B is allocated to i_1,

$$v_{i_1B} \geq \frac{1}{2}\left(\sum_{j\in B} v_{i_1j} + \sum_{j\in B} v_{i_2j}\right)$$

From Lemma 1, we have

$$\sum_{j\in B} v_{i_1j} + \sum_{j\in B} v_{i_2j} \geq \sum_{j\in B} h_j.$$

Therefore, summing over all bundles in the optimal bundling, the welfare of allocation is at least $\frac{1}{2}\sum_{j=1}^{k} h_j$. The proof follows since the maximum efficiency is $\sum_{j=1}^{k} h_j$. □

4 Computing the Optimal Bundling

We now turn to the question of computing a revenue-maximizing bundle. We will show that this problem (even with no constraints on efficiency of the solution) is strongly NP-hard. We then present an algorithm which approximates the optimal revenue by a factor 1/2; in addition, the efficiency of the bundling is no smaller than 1/2 of the maximum efficiency.

Theorem 3. *The problem of finding the optimal bundling is strongly NP-Hard.*

Proof. The proof is by reduction from 3-partition, which is strongly NP-hard: Given a multiset S of $3n$ positive integers, can $S = \{x_1, x_2, \ldots, x_{3n}\}$ be partitioned into n subsets $S_1, S_2, \ldots, S_n$ such that the sum of the numbers in each subset is equal.

Let $w = \sum_{i=1}^{3n} x_i$. We reduce the problem by constructing an instance of the bundling problem with $n+1$ bidders and $4n$ items. The instance is given in the table below. Each row corresponds to a bidder and each column represents an item. All empty values are 0.

$\frac{w}{n}$							
	$\frac{w}{n}$						
		$\ddots$					
			$\frac{w}{n}$				
				x_1	x_2	$\cdots$	x_{3n}

It is easy to see that the revenue of the optimal bundling for the instance above is w if and only if there exits a 3-partition. $\square$

4.1 Approximation algorithm

Recall that h_j and s_j are defined as the highest and second highest valuations for item j. Let A_i be the set of items for which agent i has the highest valuation, i.e. $A_i = \{j \mid v_{ij} = h_j\}$, and let $w_i = \sum_{j \in A_i} h_j$. Number agents so that $w_1 \geq w_2 \geq \ldots \geq w_n$. Let $A_{n+1} = \emptyset$, and $w_{n+1} = 0$.

Algorithm $\mathcal{B}$:

$r_1 \leftarrow \sum_{j \in A_1} s_j + \sum_{i=1}^{\lfloor n/2 \rfloor} w_{2i+1}$;
$r_2 \leftarrow \sum_{i=1}^{\lfloor n/2 \rfloor} w_{2i}$;
If ($r_1 \geq r_2$):
    Sell all items in A_1 separately;
    For $i \leftarrow 1$ to $\lfloor n/2 \rfloor$
        Bundle items in A_{2i} and A_{2i+1};
else
    For $i \leftarrow 1$ to $\lfloor n/2 \rfloor$
        Bundle items A_{2i-1} and A_{2i};

Theorem 4. *Algorithm $\mathcal{B}$ obtains at least half the revenue from an optimal bundling.*

Proof. Let OPT be the optimal revenue. We prove the following inequality.

$$OPT \leq \sum_{j \in A_1} s_j + \sum_{i=2}^{n} w_i = r_1 + r_2. \tag{1}$$

The claim then follows since the revenue of $\mathcal{B}$ is at least $\max\{r_1, r_2\}$.

To prove (1), let B be a bundle in an optimal bundling, and let i and i' be the two agents with highest valuations for B. At least one of these two agents is not agent 1; let i be this agent. Because the mechanism charges the second highest price for each bundle, the revenue of the optimal bundling from B is at most:

$$\sum_{j \in B} v_{ij} = \sum_{j \in B \cap A_1} v_{ij} + \sum_{j \in B - A_1} v_{ij} \leq \sum_{j \in B \cap A_1} s_j + \sum_{j \in B - A_1} h_j$$

Summing over all bundles in the optimal bundling yields (1). $\square$

Proposition 5. *The efficiency of algorithm $\mathcal{B}$ is at least half the maximum efficiency.*

Proof. The maximum efficiency is $\sum_j h_j = \sum_{i=1}^{n} w_i$. The efficiency of $\mathcal{B}$ is at least

$$w_1 + \min(\sum_{i=1}^{\lfloor n/2 \rfloor} w_{2i}, \sum_{i=1}^{\lfloor n/2 \rfloor} w_{2i+1}) \geq \frac{1}{2} \sum_{i=1}^{n} w_i,$$

since the algorithm always sells items in A_1 to bidder 1; this gives us the result. $\square$

5 Multiple Slots

We finally discuss the case of multiple slots, and show that our algorithm gives a constant factor of the optimal revenue when slot clickthrough rates decrease geometrically, which is realistic for sponsored search auctions [9, 6].

Suppose there are m slots numbered $1 \ldots m$. Following [2], assume that the click-through-rate of ad i for context j in slot k is separable into $c_{ij} \cdot \Theta_k$ (*i.e.* , the clickthrough rate can be factored into a term specific to the advertiser-keyword pair and another term specific to the slot). We will show that Algorithm $\mathcal{B}$ continues to give us a constant factor approximation of revenue when the slot-dependent CTR decreases geometrically, i.e. $\Theta_{k+1} = q \cdot \Theta_k$, for some q with $0 < q < 1$.

Fix an instance of the full information, sponsored search problem, *i.e.* , a value per click v_i for each advertiser i, the parameters c_{ij} for each advertiser-context pair, and the slot specific CTRs Θ_j for the slots. Let f_j be the number of impressions from context j. Given any bundling of the keywords, we now define revenue of the generalized second price auction (GSP) [8]. Unlike the second price auction for one slot, this auction is not truthful. We assume that the equilibria of [16, 8] are attained. The prices at such an equilibrium is precisely the prices that VCG would charge in each bundle [8, 2].

Denote revenue from the revenue maximizing bundling as R^*, and let $\mathcal{P}$ denote the partition of keywords in this bundling. Consider a bundle $B \in \mathcal{P}$. Let $v_{iB} = v_i \cdot \sum_{j \in B} f_j \cdot c_{ij}$. Number bidders in non-increasing sequence of v_{iB}s. The equilibrium from [8] predicts that the first m bidders appear in sequence from slot 1 to slot m, and the revenue from player k is $(\sum_{k=i}^{m} (\Theta_k - \Theta_{k+1}) v_{k+1,B})$. Thus the total revenue from all m slots is

$$R^* = \sum_{B \in \mathcal{P}} \left(\sum_{i=1}^{m} \sum_{k=i}^{m} (\Theta_k - \Theta_{k+1}) v_{k+1,B} \right). \tag{2}$$

The contribution of the first slot to R^*, denoted R^*_f, is

$$R^*_f = \sum_{B \in \mathcal{P}} \left(\sum_{k=1}^{m} (\Theta_k - \Theta_{k+1}) \cdot v_{k+1,B} \right). \tag{3}$$

Because the number of clicks to the top slot is a $(1-q)$-fraction of all clicks, we can show (proof omitted):

Lemma 6. $R^*_f \geq (1-q) \cdot R^*$.

We are now ready to prove the main theorem of this section.

Theorem 7. *Algorithm $\mathcal{B}$ is $\frac{1-q}{2}$-competitive with the optimal bundling.*

Proof. Suppose, instead, that we run a second price auction with one slot with the same bundle partition $\mathcal{P}$ (if bundles are pre-computed, truthful bidding is a dominant strategy here). Let R^*_2 be the revenue in this case. We show $R^*_2 \geq R^*_f$. Therefore, by Theorem 4 and Lemma 6 algorithm $\mathcal{B}$ obtains a $\frac{1}{2} \cdot (1-q)$ fraction of the revenue. Now we prove the claim:

$$\begin{aligned} R^*_2 &= \sum_{B \in \mathcal{B}} (\sum_{j \in B} f_j) \cdot \Theta_1 \cdot v_{2,B} \\ &\geq \sum_{B \in \mathcal{B}} (\sum_{j \in B} f_j) \sum_{k=1}^{m} (\Theta_k - \Theta_{k+1}) \cdot v_{2,B} \\ &\geq \sum_{B \in \mathcal{B}} (\sum_{j \in B} f_j) \sum_{k=1}^{m} (\Theta_k - \Theta_{k+1}) \cdot v_{k+1,B} \\ &= R^*_f. \end{aligned}$$

□

Acknowledgements: We are grateful to Mohammad Mahdian, David Pennock, Amin Saberi and Michael Schwarz for helpful discussions.

References

[1] W. Adams and J. Yellen. Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics 90*, pages 475–498, 1976.

[2] G. Aggarwal, A. Goel, and R. Motwani. Truthful auctions for pricing search keywords. In *EC '06: Proceedings of the 7th ACM conference on Electronic commerce*, pages 1–7, New York, NY, USA, 2006. ACM Press.

[3] M. Armstrong. Multiproduct nonlinear pricing. *Econometrica 6*, pages 64, 51–75, 1996.

[4] M. Armstrong. Optimal multi-object auctions. *Review of Economic Studies 67*, pages 455–481, 2000.

[5] Y. Bakos and E. Brynjolfsson. Bundling information goods: Pricing, profits, and efficiency. *Management Science*, 1999.

[6] N. Brooks. The Atlas Rank Report: How search engine rank impacts traffic. *Atlas Institute*, July 2004.

[7] I. Chakraborty. Bundling decisions for selling multiple objects. *Economic Theory 13*, pages 723–733, 1999.

[8] B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *Stanford GSB Research Papers: Paper 1917*, 2005.

[9] B. Edelman and M. Schwarz. Optimal auction design in a multi-unit environment: the case of sponsored search auctions. *Manuscript in preparation*, 2007.

[10] E. Even-Dar, M. Kearns, and J. Wortman. Sponsored search with contexts. *Third Workshop on Sponsored Search Auctions*, 2007.

[11] P. Jehiel, M. Meyer-ter-Vehn, and B. Moldovanu. Mixed bundling auctions. *Journal of Economic Theory*, 2007.

[12] A. Likhodedov and T. Sandholm. Approximating revenue-maximixing combinatorial auctions. In *Proceedings of the National Conference on Artificial Intelligence(AAAI)*, 2005.

[13] A. Manelli and D. Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Discussion paper, Arizona State University*, 2001.

[14] P. McAfee, J. McMillan, and M. Whinston. Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics 104*, pages 371–383, 1989.

[15] T. Palfrey. Bundling decisions by a multiproduct monopolist with incomplete information. *Econometrica 51*, pages 463–484, 1983.

[16] H. Varian. Position auctions. *To appear in International Journal of Industrial Organization*, 2006.