University of Liverpool

Department of Computer Science

First Year Progress Report

Revenue Maximization via Signaling

Author:
Argyrios Deligkas

Supervisors:
Dr. Mingyu Guo
Dr. Piotr Krysta

Advisors:
Dr. Giorgos Christodoulou
Dr. Martin Gairing
Dr. Rahul Savani

April 25, 2014

Contents

1 Introduction

My chosen topic of study lies in the field of Mechanism Design with a focus on Revenue Maximization. This PhD aims to carry out research on methods for selling advetrisements on the internet that yield higher revenue than the standard ones. During the initial stages the emphasis has been on study probabilistic single-item second-price auctions where the item is characterized by a set of attributes. The auctioneer knows the actual instantiation of all the attributes, but he may choose to reveal only a subset of these attributes to the bidders.Our model is an abstraction of the following Ad auction scenario. The website (auctioneer) knows the demographic information of its impressions, and this information is in terms of a list of attributes (e.g., age, gender, country of location). The website may hide certain attributes from its advertisers (bidders) in order to create thicker market, which may lead to higher revenue. We study how to hide attributes in a way that the auctioneer gains the maximum expected revenue. We have already shown that it is NP-hard to compute the optimal attribute hiding scheme, we derived a polynomial-time solvable upper bound for the optimal revenue and we have proposed two heuristic-based attribute hiding schemes that experiments show that revenue achieved by these schemes is close to the upper bound.

2 Related research

One advantage of Internet advertising is that it offers advertisers the ability to target customers based on various traits such as demographics. [1] showed that, for sponsored search of a given keyword, instead of running a single auction for the keyword, we can split the whole auction into many separate auctions based on visitors/impressions' *contexts* (e.g., demographics). For example, if we know and only know the visitors' locations, then each location defines a context. In this example scenario, splitting based on context means separate auction for each location. Splitting based on context increases the advertisers' welfare. The explanation is simple: after splitting, advertisers can tailor their bids to the context. As a result, advertisers generally only win (impressions from) visitors that they aim to target. On the other hand, splitting may reduce the revenue received by the auctioneer (publisher, e.g., website) due to the *thin market problem*: there may be few competitors for some contexts. Actually, if for every context, there is only one advertiser interested in it, then the total revenue is 0 under the standard second-price auction.

In [2] Ghosh et al. observed that having a single auction for all contexts and having separate auction for each context are not the only two options. There are other ways to split based on context, and it may lead to much higher revenue. The idea explored in [2] is to *cluster* the contexts into bundles, and run separate auction for each bundle. For example, suppose there are three different contexts: Beijing, Chicago, and London (assuming the only contextual information is the location and visitors are only from these three cities). We can have one auction for the bundle Beijing and Chicago (and a second auction for London only). The interpretation (due to [3]) is that if a visitor is from Beijing or Chicago, then the auctioneer informs the advertisers that the impression is from one of these two cities, *but not exactly which*. When this happens, both advertisers targeting Beijing and advertisers targeting Chicago will compete in the auction. Their bids depend on how much they value impressions from Beijing and Chicago, respectively. Their bids also depend on the conditional probability that the impression is from Beijing (or Chicago) given that the impression is from one of these two cities.

To put it more formally, [2] studied probabilistic single-item second-price auctions (again, interpretation due to [3]). In such an auction, there is only one item for sale under a second-price auction, but the item has different possible *instantiations*. The auctioneer knows the actual instantiation but the bidders do not. The auctioneer may choose to hide certain information from the bidders if this increases the revenue. The probabilistic single-item second-price auction model is an abstraction of the following Ad auction scenario. We have a website that sells one advertisement slot. That is, there is only one item – the only advertisement slot, but the item takes many possible instantiations, due to the fact that visitors/impressions have different demographic profiles. The auctioneer knows every visitor's demographic profile, and he may hide certain information from the advertisers. As mentioned above, [2] considered hiding information by *clustering*: the auctioneer tells the bidders that the actual instantiation is among several instantiations. [3, 4] studied the exact same model and went one step further. These two papers studied hiding information by *signaling*: the auctioneer sends out different signals, and the bidders infer the probability distribution of the actual instantiation, based on the signal received. It is easy to see that signaling is more general than clustering. Interestingly, for full information settings (settings where the auctioneer knows the bidders' exact valuations), [2] showed that it is NP-hard to compute the optimal clustering scheme (optimal in terms of revenue). On the other hand, [3, 4] both independently showed that, under the same full information assumption, it takes only polynomial time to solve for the optimal signaling scheme. This is mostly due to the fact that instantiations are treated as divisible goods under signaling schemes.

We continue the study of revenue-maximizing probabilistic single-item second-price auctions. We observe that in practice, *Ad impressions are categorized based on multiple attributes.* Given this, we argue that the most natural way to hide information is by *hiding attributes.* For example, let there be three attributes, each with two possible values:

- Age: Teenager, Adult
- Gender: Male, Female
- Location: US, Non-US

Together there are 2^3 possible instantiations. Under the clustering scheme studied in [2], the website is allowed to hide information by bundling any subset of instantiations. However, not all bundles are natural. For example, consider the bundle {(Teenager, Male, US), (Adult, Female, Non-US)}. By creating this bundle, the website basically may tell the advertisers that a visitor is either a teenage US male or an adult Non-US female. This does not appear natural. The signaling scheme studied in [3, 4] is even more general than clustering, so it may also lead to unnatural bundles.

On the other hand, attribute hiding always leads to natural bundles. For example, the website may hide the location attribute. That is, if the actual instantiation is (Teenager, Male, US), then the website may inform the advertisers that the visitor is a teenage male. By hiding the location attribute, we essentially created a bundle (Teenager, Male, ?), which consists of both (Teenager, Male, US) and (Teenager, Male, Non-US).

Based on the above example, it is easy to see that attribute hiding is clustering with a particular structure. It should be noted that this relationship between attribute hiding and clustering does not mean previous results on clustering apply to our model. For example, one of the two main results from [2] is a constructed clustering scheme that guarantees one half of the optimal revenue (and one half of the optimal social welfare). The construction does not apply to our model since it generally leads to unnatural bundles.

Besides the aforementioned related work in the computer science literature, bundling has also been well-studied in the economics literature. [5] observed that for small numbers of bidders, a revenue-maximizing auctioneer may choose to bundle the items, and this makes bidders universally worse-off. On the other hand, for large numbers of bidders, the auctioneer may choose to unbundle the items, and this hurts the high-demand bidders while benefiting the low-demand bidders. [6] quantitatively analyzed the bundling behavior of the auctioneer. The result is that under a Vickrey auction, for each pair of objects, there is a unique critical number. If there are fewer bidders than this number, the seller chooses to bundle the items, and vice versa. [7] studied more sophisticated bundling policy, including bundling with discounts and probabilistic bundling (the probability of bundling occurring depends on the bids).

3 The model

There is a single item for sale characterized by k attributes (attribute 1 to k). Attribute i has C_i possible values, ranging from 0 to $C_i - 1$. Let m be the total number of possible instantiations. That is, $m = \prod_i C_i$.

An instantiation whose i-th attribute equals a_i is written as

$$(a_1, a_2, a_3, \ldots, a_k)$$

The space of all possible instantiations Ω is

$$\{0, \ldots, C_1 - 1\} \times \{0, \ldots, C_2 - 1\} \times \ldots \times \{0, \ldots, C_k - 1\}$$

We are trying to create bundles only by hiding some of the attributes and we call these bundles as *natural bundles*. To put it more formally, a *natural bundle* b is an element from the following set of all natural bundles:

$$\{0, \ldots, C_1 - 1, ?\} \times \{0, \ldots, C_2 - 1, ?\} \times \ldots \times \{0, \ldots, C_k - 1, ?\}$$

. We say that a natural bundling scheme is valid, if and only if every item belongs to exactly in one natural bundle or it is sold seperately that is without any hidden attribute.

3.1 Assumptions so far

- The probabilities' of different instantiations are based on a publicly known distribution.
- Full information: the auctioneer knows the bidders' true valuations.
- Bidders have additive valuations.

4 Results obtained so far

1. In our model under the previous assumptions, it is NP-hard (in the number of the items) to compute a valid natural bundling scheme in order to maximize auctioneer's revenue. In this general case we have assumed moreover that the bidders have arbitrary values for the items.
2. We found an upper bound using a linear program which has polynomial size in the number of the items.
3. We propose two heusristic algorithims
   - The first algorithm uses matching and can be applied only for items with binary attributes. This algorithm produces good results in experiments but there are instances that tha revenue from its output can be arbitrarily worse than the optimal one. On the positive side is that it finds the bundling scheme with the optimal revenue, if we add as a restriction that the auctioneer is allowed to hide at most one attribute.
   - The second is more complicated and creates a "tree-structured" bundling scheme. This algorithm begins with the bundle with all the attributes hidden and recursevely splits the bundles into halves by revealing one attribute at a time. It should be mentioned that these revealings are not myopic, but there is some preprocessing before. This algorithm can be

applied for items which have attributes with more than two values and experiments show that it produces bundling schemes which revenues are really close to the upper bound (and many times are the optimal).

4. If each bidder is interested in exactly one item with value 1 (his value for all other instantiation is zero) and the attributes are binary, then we can compute the optimal natural bundling scheme using matching.

5. If each bidder is interested in exactly one item with value 1 (his value for all other instantiation is zero) and each attribute has 8 different values, then it is NP-hard to compute the optimal natural bundling scheme.

It should be mentioned that if the bidders' valuations are commonly known, then the mechanism is not truthful. The first three results are presented in the paper [8] which is accepted for the IJCAI'13 conference.

5 Future goals

1. A first goal is to study theoreticaly the "tree-structured" algorithm in order to see wheter it guarantees a constant fraction of the optimal revenue.

2. A second interesting question is to find the computational complexity for the optimal attribute hiding scheme, if each bidder is interested exactly in one bundle and has zero value for every instatiation outside from this bundle. This is a natural question, because advertisers usually target a specific fraction of the market.

3. Another way is to change our assumptions:

    - What happens if the bidders do not know anything about the items distribution? Can the auctioneer create a two stage signaling mechanism, that in the first stage she decides what information to reveal about the distribution and in the second to send the attribute hiding scheme, such that she gains more revenue?
    - What happens if the auctioneer does not know the actual bidders' valuations?

- If the valuation functions are not additive, how can we attack the problem?

4. It is possible that the values of the attributes have some structure (hierarchy, tree structure) so all bidders have valuations according to these structure. Can we use that structure in order to set reserve prices for every item or for every attribute?

5. Another goal is to compute a different natural bundling scheme for each bidder.

References

[1] Eyal Even-Dar, Michael Kearns, and Jennifer Wortman (2007). Sponsored search with contexts. *In Workshop on Internet and Network Economics (WINE), San Diego, CA, USA, 2007.*

[2] Arpita Ghosh, Hamid Nazerzadeh, and Mukund Sundararajan (2007). Computing optimal bundles for sponsored search. *In Workshop on Internet and Network Economics (WINE), San Diego, CA, USA, 2007.*

[3] Yuval Emek, Michal Feldman, Iftah Gamzu, Renato Paes Leme, and Moshe Tennenholtz (2012). Signaling schemes for revenue maximization. *In Proceedings of the ACM Conference on Electronic Commerce (EC), Valencia, Spain, 2012.*

[4] Peter Bro Miltersen and Or Sheffet (2012). Send mixed signals: earn more, work less. *In Proceedings of the ACM Conference on Electronic Commerce (EC), Valencia, Spain, 2012.*

[5] Thomas R Palfrey (1983). Bundling decisions bya multiproduct monopolist with incomplete information. *Econometrica, 51(2):46383, March 1983.*

[6] Indranil Chakraborty (1999). Bundling decisions for selling multiple objects. *Economic Theory, 13(3):723733, 1999.*

[7] Christopher Avery and Terrence Hendershott (2000). Bundling and optimal auctions of multiple products. *Review of Economic Studies, 67:483497, 2000.*

[8] Mingyu Guo and Argyrios Deligkas (2013). Revenue Maximization via Hiding Item Attributes. *To appear in Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI-13), Beijing, China.*