Pricing Digital Goods via Auctions: A Survey

Mahdi Tayarani Najaran
tayarani@cs.ubc.ca

April 24, 2008

Abstract

The cost of reproducing a digital good is upper bounded by some low cost, providing the seller with an unlimited supply of items for sale. In such cases, or even in cases that the supply is limited but at least as many as the number of bidders, the agents have no motivation to compete with each other. This lack of competition decreases the total revenue of the seller, making traditional auctions to fail to collect optimal revenue under such circumstances. This paper surveys a class of auctions specifically designed for such conditions. The goal of these auctions is to limit the number of sold items in a manner that once again, the best strategy for each is to bid the amount he truly values the good, and not any value lower than that. The auctions have been tuned to maximize the seller's revenue, the money he collects from the agents.

1 Introduction

There are n agents interested in buying a specific item, where each agent i has a private valuation for the item, u_i, which is the maximum price he is willing to pay for it. However, the utility each agent gains from the auction is the difference between the valuation he has and the price he pays for the item. The auctioneer decides the price of the item based on the values b_i, the agent's bids. It can clearly be inferred that agents have no motivation to bid truthfully, i.e. declare their valuation as their bid, unless the rules of the auction are set in way that agents gain the most by bidding truthfully, or at least can't do better by not bidding truthfully. Such auctions are said to be *dominant strategy truthful* or simply *truthful*. The Vickrey-Clarke-Groves (VCG) mechanism is an instance of a truthful mechanism and has been used for selling items in auctions.

For most of the text, the auctions of interest are single round sealed bid multiple unit single demand auctions, unless stated otherwise. The auctioneer (or seller) is willing to sell k identical items to n agents, where each agent declares his bid only to the seller and is interested in possessing at most 1 item. The seller decides which agents are assigned an item and at what prices, with the intention to maximize his own revenue which is the sum of the payments of all the agents. The agents that are not assigned an item pay nothing. It is important to note that everyone is aware of the rules of the auction in advance.

In normal auction settings, the number of items is less than the number of agents and, hence, VCG charges the first k highest bidders a price equal the the *k+1*-st highest bid. However, in the case where the seller is selling digital goods like mp3 downloads or pay per view movies, i.e. goods that have a small marginal cost of reproduction, or in the general case the number of items is more than the agents, this question arises that at what price should the items be sold to achieve an optimal revenue for the seller? Unlike the case where the supply was limited and agents had to compete for the goods, lack of competition for these types of goods prevents the agents from having any incentive to bid high values.

In such cases, VCG would assign an item to every agent at zero cost which is the worst possible outcome for the seller. For long, market analysis has been widely used to accommodate the price of the goods with the demand of the market, mainly referred to as *fixed pricing*. If perfect knowledge of consumer valuations were available, fixed pricing would yield optimum revenue. Unfortunately such information is not always available and the optimality of fixed pricing depends heavily on the accuracy of the underlying market analysis. Also, such analysis needs frequent updating to be able to adopt to changes in the market, whereas auctions have such property in nature.

This paper surveys various methods used to perform pricing of digital goods from the seller's perspective, which intends to maximize his own revenue. The assumptions and notations used throughout the text are discussed in Section 2. Section 3 covers the basic random auctions that use sampling to yield optimum revenue. The problem with these auctions is that some agents may wish they were in the place of others. Auctions that try to eliminate such feeling from the agents are discussed in Section 4. A new class of auctions are presented in Section 5 that allow the sale of some goods, such as electronic books, to be continuous over time and not end as soon as the auction is over. A technique that can be used to transform any random auction for digital goods to a deterministic auction is briefly explained in Section 6, and Section 7 concludes the survey.

2 Preliminaries

In the rest of the text we assume, without loss of generality, bids are ordered in ascending order ($b_i \leq b_{i+1}$). The bid values may be divided by the smallest amount, and thus, be in the range [1,h]. R is the sum of all sale prices. Also, there is no collusion amongst the agents: none of the agents form a group to gain by changing their bidding strategies.

Given a set of bids B, selling as many items as possible is not an optimal solution. Also, it might seem tempting for the seller to maximize his revenue by selling k items at price b_{n-k} in a sense that:

$$k = argmax_{1 \leq k < n} \; k \times b_{n-k} \tag{1}$$

This is equivalent to using VCG to sell k items at the price of the $k+1$-st highest bid, where the number of items has been optimized to obtain maximum revenue. Although VCG is truthful, such usage of it is not truthful because the price each agent i pays depends on his own bid b_i.

The concept of *bid independence* introduced in [6] has been widely used. The basic idea is that in order for an auction to be truthful, it should be bid independent, i.e. when computing the price bidder i has to pay, his bid value b_i should not be used. It is also shown that any bid independent auction is truthful and vice versa.

Let F be the revenue for optimal fixed pricing. In other words, F is the revenue due to the optimal non-truthful single-price auction. F is used as a benchmark for evaluating the optimality of the revenue obtained from a specific auction. An auction is said to be β-competitive against F, if for all bids B the expected profit A(B) satisfies:

$$E[A(B)] \geq \frac{F(B)}{\beta} \tag{2}$$

3 Random Sampling Auctions

A deterministic auction assigns one specific outcome to a set of given bids and collects a known amount of revenue. On the other hand, a random auction gives a distribution over possible outcomes and the revenue is a random variable. For such auctions, the expected revenue is considered. As will be explained in section 6, deterministic auctions provide lower worst case revenue compared to random auctions, so most of the auctions presented in the context of digital good's pricing are random auctions.

3.1 Single Good Auction

Using a subset of a set to estimate certain features of the whole set is used in many fields of science. The idea here would be to use a subset of the agent's bids as data for market analysis to find an optimal price. The accuracy of the estimate depends directly on the size of the subset: the greater the subset is, the more accurate the analysis becomes. Under certain conditions, as we will see next, accuracy and revenue form a tradeoff between them. Maximizing one minimizes the other.

So doing, the multiple price random sampling auction presented in [1] works as follows:

1. Partition the agents in two groups S_1 and S_2.
2. Use Equation 1 to calculate the number of items k_1 and optimum price, p_{opt1}, for bids in S_1 & k_2 and p_{opt2} for bids in S_2.
3. For all agents in S_2, assign to each agent an item if he has bid grater than the optimum price, i.e. $b_i \geq p_{opt1}$.

4. For all agents in S_1, assign to each agent an item if he has bid grater than the optimum price, i.e. $b_i \geq p_{opt2}$.

Normally S_1 and S_2 have the same size and agents are partitioned using a random coin flip. The auction uses two different prices to assign goods to the agents, an approach that is normally not acceptable for consumers. In cases where such approach is not acceptable, in step 4, all agents in S_1 are rejected and are assigned nothing, decreasing the expected revenue by 50%.

As it can be seen, since the price each agent pays does not depend on the bid he made, this auction is truthful. As for the revenue, unfortunately it appears it is upper bounded by being at most 4-competitive [1].

3.2 Multiple Good Auction

An extension to the case of the pay per view pricing of a movie would be the pricing of pay per view of a movie in a bundle of movies. Consider the case that passengers on a plane have the chance to choose an in-flight movie from a bundle of m movies. The problem is restricted to preferring at most one movie (assume the length of the flight is short). Currently, all movies are priced equally at a fixed price by the airlines. However, having each passenger declare a different price for a ticket based on his valuation for the available in-flight movies seems somewhat interesting.

The auction presented in [2] for this problem setting is identical to what was presented in the previous subsection, supporting both the single price and the dual price auctions. The difference, however, is the way the optimum prices are calculated. The latter appears to be more sophisticated. Let b_{ij} represent the bid agent $1 \leq i \leq n$ places for item $1 \leq j \leq m$. Without loss of generality, we can assume item m is a dummy item valued 0 by every agent. This allows *agent i being assigned item m* be interpreted as *agent i not being assigned anything.* Additionally, r_i represents the optimal price for item i, x_{ij} indicates agent i being assigned item j and, finally, p_i is the profit of agent i, the difference between the utility he has and the price he is charged. The seller's goal is to select r_i's in a way that maximizes the revenue, $\sum_j \sum_i x_{ij}.r_j$. Not delving too much into details, the following optimization problem provides enough information about how the optimal price is calculated [2]:

$$\begin{array}{cc}
max \sum_j \sum_i x_{ij}.r_j & subject\ to \\
r_m = 0 & \\
\sum_i x_{ij} \leq 1 & 1 \leq i \leq n \\
x_{ij} \geq 0 & 1 \leq i \leq n, 1 \leq j \leq m \\
p_i + r_j \geq b_{ij} & 1 \leq i \leq n, 1 \leq j \leq m \\
\sum_i p_i = \sum_j \sum_i x_{ij}.(b_{ij} - r_j) & 1
\end{array} \tag{3}$$

Walking through the listing line by line, the program maximizes the revenue (line 1) constrained by item m being priced 0 (line 2), each agent being assigned at most one item and at least none (lines 3 & 4) and each agent being assigned an item only if he bids higher than the price of that item (lines 5 & 6).

A careful reader might notice there are no restrictions on x_{ij} being integers (more accurately $x_{ij} \in \{0, 1\}$). It follows that for any fractional solution of Equation 3, there exists a solution with $x_{ij} \in \{0, 1\}$. For more details see [2]. From a performance perspective, the most expensive function used for the single item goods was sorting the bids. As for the multiple unit case, an optimization problem has to be solved which has a quadratic objective function. Performance here is definitely an issue. As for the revenue, this method is 24-competitive [2].

4 Envy-free Auctions

Consider the dual price auction presented in section 3.1. The agents in the subset paying the higher price some how wish they had been in the other subset. In other words, some agents *envy* others. Even in the single price auction, some agents are rejected an item even though their bid might have been higher than the price decided by the seller (the agents in the rejected subset). These agents all get utility of zero, whereas they could have obtained positive utility if they had been assigned a good.

An auction is said to be *envy-free* if after the auction is run, no bidder would be happier with the outcome of another agent. For digital goods, this means all the agents pay the same price, and a good is assigned to all agents that bid higher than this value. As important such property of an auction may seem, unfortunately, part of the work done in [4] is dedicated to showing that no auction can be truthful, constant-competitive and envy-free at the same time.

Fortunately, that's not all. The restrictions of being truthful or envy-free may be relaxed to achieve good results. In [4] two different auctions are presented: one which is truthful with high probability and envy-free, another which is envy-free with high probability and truthful.

An auction is truthful with probability $1-\epsilon$ if the probability that any bidder can benefit from an untruthful bid is at most ϵ. An auction is truthful with high probability if $\epsilon \to 0$ as $m \to \infty$, where m is the number of winners in the auction.

The methods described in [4] are more sophisticated to be able to cover in a short discussion, but they both extend the Consensus Revenue Estimate (CORE) class of auctions. To give a glance of how they're implemented, we start by defining the *profit extractor*. The profit extractor is a truthful envy-free auction that extracts exactly a given total revenue C, if possible. It simply finds the largest possible k such that the highest k bidders can equally share the cost C and pay C/k each. Let f be the function that maps bid values to prices for all agents. Since the profit extractor is truthful, from agent i's point of view, it can be viewed as a function $f(b_{-i}) = cs_C(b_{-i})$. Now, for any function r, if we have $r(b_{-i})$ equals to some constant c for all i, i.e r is a *consensus*, the auction defined by $f(b_{-i}) = cs_{r(b_{-i})}(b_{-i})$ is a profit extractor extracting a total revenue of exactly c and, hence, truthful and envy-free. The rest of the definition of the promised auctions revolve around selecting the functions r.

It appears that if c is chosen within a constant factor ρ of the optimum fixed price F, there exist families of r functions, most of which are a consensus. So if an r is picked randomly, the resulting auction is envy-free with high probability, $1 - \log_c \rho$ to be precise. In terms of the number of agents winning the auction, m, the probability of this auction being envy-free is $1 - \log_c \frac{m}{m-1} = 1 - \Theta(1/m)$ [4]. So this auction is truthful and constant-competitive and also envy-free with high probability.

For each function r found in the previous auction, there seems to be another function r' where the resulting auction is truthful with the same probability, $1-\Theta(1/m)$. Hence, this new auction is constant-competitive and envy-free plus truthful with high probability [4]. In the tv broadcast mass market, where m is intended to be high, this high probability seems to be very satisfactory.

5 Online Auctions

Single round auctions have a natural time requirement for the bids: all bids should be submitted by a certain deadline, after which the seller starts calculating the outcome of the auction and rejects all incoming bids. While this time requirement may not be an issue for an auction with a small number of items to sell, digital goods aim at a huge market which is, of course, hard to coordinate in terms of having every agent bind his participation in the auction to a specific time slot. Moreover, most consumers require a prompt answer to their bids of getting or not getting the good. All this put aside, most digital goods such as mp3s and electronic books are sold continuously, not just once. If a single round auction be used to sell such goods, it is very likely that most consumers will not manage to bid at the right time, some won't want to bid because they will have to wait for a long time to get an answer and the seller rejects the revenue he could have gained from agents bidding in the future.

Online auctions have been designed to accommodate such situations. The seller keeps the auction running, and as each agent places his bid, the seller responds to his bid very quickly, but not necessarily immediately, by calculating a price without waiting for the other agents who will place their bids afterwards. Due to the fact that decisions are made one by one, using sampling techniques might not seem convenient.

An online auction may be viewed as a collection of n pricing functions $s_1, ..., s_n$, where each s_i is a function of the bids made up to the point agent i make his bids. In other words, $s_i = s_i(b_1, ..., b_{i-1})$. The same proposition about truthfulness applies to online auctions as it did to single round auctions: an online auction is truthful if and only if it is bid independent [5].

A simple online auction would be for the seller to set for each agent i, $s_i = 2^k$ where $k \in \{0, ..., \lfloor \log h \rfloor\}$ is chosen at random and h is the ratio of the highest and the lowest bids, assuming it is known in advance. Since the price each agent pays is independent of his bid, this auction is truthful. In cases where h is not known in advance, the auction can start by assuming $h = 1$ and update its value at each step when a new bid comes in. Surprisingly, this auction is $\Theta(\log h)$-competitive in both cases that h is either known in advance or not. However, it feels to be more offline than online.

An online auction could be created using buckets. Given a sequence of bids, they are divided into $l = \lfloor \log h \rfloor + 1$ intervals $I_0, ..., I_{l-1}$, where $I_k = [2^k, 2^{k+1})$ and is assigned to bucket B_k. The weight of the k-th bucket, w_k, is the sum of the bids inside it: $w_k = \sum_{i \in B_i} i$. The seller selects a bucket at random from the buckets, with the probability of a bucket being selected being proportional to its weight, and charges that agent the lowest possible bid of the bucket. More formally, let $d \leq$ be some tuning parameter. The price agent i is charged, s_i, has the following probability distribution over k:

$$Pr[s_i = 2^k] = (\frac{w_k}{\sum_{r=0}^{l-1} w_r})^d \tag{4}$$

Notice that when agent i places his bid, the buckets contain the bids $b_1, ..., b_{i-1}$. The bid independence of the auction is visible, confirming that the auction is truthful. The analysis available in [5] proves this auction is $O(3^d (\log h)^{1/(d+1)})$-competative relative to the optimal fixed price revenue.

Although the auction may perform weakly at the beginning, eventually, the price becomes closer to the optimal price. Since the decisions in the auction are made at separate times, envy of different agents doesn't make much sense. Only in the case that two different agents place their bids almost simultaneously could envy become an issue. Having the auction make decisions after specific time intervals (and not instantaneously) could solve this problem.

One last issue of online auctions would be the freedom each agent has to place more than one bid. While he cannot gain by bidding untruthfully in each bid, he can gain if he floods the auction with low bids until he is assigned an item. This doesn't contradict the no collusion assumption we made previously, because no other agent is assisting this malicious agent by changing his strategy. Designing an auction that prevents such behavior may be an interesting research topic.

6 De-randomization of Auctions

It might be noticed all the auctions presented so far were random auctions. This is in fact due to the analysis done in [1], which proves no deterministic auction can be competitive. This being true, interestingly, [3] presents a general technique for turning any random auction into a deterministic auction which collects approximately the same total revenue. This technique does not contradict the analysis mentioned previously, since the new deterministic auction is asymmetric, i.e the outcome of the auction depends on the order the bids are placed. In other words, different permutations of the same bids have different outcomes. The technique uses an analogy between bid independent auctions and the hat problem.

The hat problem, initially studied in coding theory, is a setting where there are a total of n hats and each hat could be any of the k possible colors. Each hat is placed on the head of one agent, and agents get to observe the color of other agent's hats, but not their own. The problem is to devise a strategy to maximize the number of agents that are able to guess their own hat color correctly.

The basic solution to come in mind is for each agent to guess his own hat color randomly from the set of k possible colors. In the uniform case, the expected number of correct guesses would be $1/k$, but this is a random strategy. The derandomization technique constructs a directed graph based on the observation each agent has of other agent's hat colors and proposes a strategy for each agent to declare his hat color based on the maximum flow of the graph. This strategy has an expected success rate of $1/k$, equal to that of the random strategy, but is deterministic.

How is the hat problem related to an auction? The seller must calculate the price each agent pays based on the bids other agents make, b_{-i}. If the price calculated by the seller is less than what agent i has bid, he is assigned an item. So the price an agent pays can be interpreted as his hat color, and the hat color of the rest of the agents are the bids they made. Agent i correctly guessing his hat color is thus equivalent to the seller placing a price for agent i less than his bid.

The expected revenue of the resulting deterministic auction is closely related to the expected revenue of the original random auction. If the expected revenue of a random auction is $E[A(b)]$, then the expected revenue of the deterministic auction constructed from it using this technique is at least $E[A(b)]/4-2h$, where h is the ratio of the highest and lowest bids [3].

As much as the bound on the revenue may seem appealing, the computational complexity is not. The technique has to solve an exponential size flow problem. For the case $k = 2$, the paper provides a polynomial time algorithm for finding the flow. Existence of such algorithms for cases where $k \geq 3$ is left as an open problem.

7 Conclusions

This paper surveyed auctions used for pricing digital goods which were designed in a manner to be truthful and optimize the total revenue of the seller. Although it may be proven that none of these auctions can collect a revenue equal to that of an optimal non-truthful auction, some analysis shows these auctions provide a revenue that is some factor smaller than optimum, even in the worst case.

Nearly all the auctions presented were random, giving a distribution over outcomes and a revenue which is a random variable. As discussed earlier, some auctions require very complex computations to calculate the price each agent has to pay. These complexities can have great impact on applicability of the auction.

The auctions presented here are not just used for unlimited supply. In cases where the number of items k is less than the number of agents n, these auctions can be used if the seller simply rejects the $n-k$ lowest bidders.

References

[1] Andrew V. Goldberg, Jason D. Hartline and Andrew Wright, *Competitive Auctions and Digital Goods*, In Proc. 12th Symp. on Discrete Algorithms, pages 735-744, ACM/SIAM, 2001.

[2] Andrew V. Goldberg and Jason D. Hartline, *Competitive Auctions for Multiple Digital Goods*, In ESA, pages 416-427, 2001.

[3] Gagan Aggawal, Amos Fiat and Andrew V. Goldberg, *Derandomization of Auctions*, In STOC, pages 619-625, 2005.

[4] Andrew V. Goldberg and Jason D. Hartline, *Envy-Free Auctions for Digital Goods*, ACM Conference on Electronic Commerce, pages 29-35, 2003.

[5] Ziv Bar-Yossef, Kirsten Hildrum and Felix Wu, *Incentive-Compatible Online Auctions for Digital Goods*, In Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithm (SODA), pages 964-970, 2002.

[6] Andrew V. Goldberg, Jason D. Hartline, A. Karlin, M. Saks and Andrew Wright, *Competitive Auctions and Digital Goods*, Games and Economic Behaviour, 2002.