Maximizing revenue in symmetric resource allocation systems when user utilities exhibit diminishing returns

(Extended Abstract)

Roie Zivan, Miroslav Dudík, Praveen Paruchuri and Katia Sycara
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
{zivanr,mdudik,paruchur,katia}@cs.cmu.edu

ABSTRACT

Consumers of resources in realistic applications (e.g., web, multimedia) typically derive diminishing-return utilities from the amount of resource they receive. A resource provider who is deriving an equal amount of revenue from each satisfied user (e.g., by online advertising), can maximize the number of users by identifying a satisfaction threshold for each user, i.e., the minimal amount of resource the user requires in order to use the service (rather than drop out). A straightforward approach is to ask users to submit their minimal demands (direct revelation). Unfortunately, self-interested users may try to manipulate the system by submitting untruthful requirements.

We propose an incentive-compatible mechanism for maximizing revenue in a resource allocation system where users are ex-ante symmetric (same amount of revenue for any satisfied user) and have diminishing-return utility functions. Users are encouraged by the mechanism to submit their true requirements and the system aims to satisfy as many users as possible. Unlike previous solutions, our mechanism does not require monetary payments from users or downgrading of service.

Our mechanism satisfies the number of users within a constant factor of the optimum. Our empirical evaluation demonstrates that in practice, our mechanism can be significantly closer to the optimum than implied by the worst-case analysis.

Our mechanism can be generalized to settings when revenue from each user can differ. Also, under some assumptions and adjustments, our mechanism can be used to allocate resource periodically over time.

Keywords

Auction and mechanism design

1. INTRODUCTION

There are many applications where the satisfaction of users, with respect to improvements in product quality or product performance, is not linear but is governed by diminishing returns. In such applications, there is some threshold value which quantifies the quality or performance required for the satisfaction of the user: below the threshold, the user is unsatisfied; however, above the threshold, the additional satisfaction from a larger quantity or quality of a product (or resource) grows at a slower and slower rate. This latter property is often called *diminishing returns*.

Cite as: Maximizing revenue in symmetric resource allocation systems (Extended Abstract), Roie Zivan, Miroslav Dudík, Praveen Paruchuri and Katia Sycara, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. XXX-XXX.

In order to maximize the number of satisfied users (agents), the allocator needs to know their satisfaction thresholds. One way to elicit the satisfaction threshold of agents is to have them submit their demand when applying for service (direct revelation). However, in order to increase their own utility, which can be achieved by receiving a larger amount of resource (even with diminishing returns, larger amounts of resource give rise to increases in utility), self-interested agents may try to manipulate the system by submitting untruthful needs (thresholds).

In this paper we describe a method for truthful elicitation of preferences from agents with diminishing-return utility functions in resource allocation applications. The contributions of our work are as follows: First, unlike traditional mechanisms of truthful elicitation (e.g., VCG [2]), we do not require either monetary transfers, or even conversions of hypothetical payments into degradation of service (e.g., [1]), indeed, our assumption that the user utility jumps from unsatisfied to satisfied makes such conversions impossible. Second, our method ensures that a large number of agents will be satisfied. This is true even in cases where the standard VCG mechanism would assign allocations which would result in zero utility for agents (when the demands of agents are tied, but there is insufficient total resource). Third, we prove that the number of user agents satisfied by our mechanism is within a constant factor of the optimal allocation method. However, unlike our method, the optimal allocation method does not guarantee truthfulness. Our experimental comparison reveals that in practice, the number of satisfied agents is close to optimal for various distributions of agents' needs. Fourth, our method can be extended and adjusted to systems that include priorities (some agents are expected to bring higher revenue to the system and therefore are entitled for larger portions than others), and (under some restrictions) in systems where the resource is allocated periodically over time.

The goals we set for this study were most challenging considering the impossibility of payments (or payment conversions) which is one of the foundations of traditional mechanism design. When payments are part of the mechanism, an agent is indifferent between winning a resource and paying for it the appropriate amount, or not winning a resource and not having to pay. In our setup, agents are not charged anything (or possibly they are charged a flat subscription fee independently of their demand), and thus we cannot resolve tied demands by charging some agents and not charging others. A naive approach either allocates resource to all or none of the tied users, which can be very inefficient. Our work relies on the diminishing returns property to remove this inefficiency while preserving truthfulness.

2. RESOURCE ALLOCATION MECHANISM

Our goal is to allocate an infinitely divisible but bounded resource among agents from some set $\mathcal{A}$. We assume that the total available quantity of resource is $Q > 0$. The allocation is a vector

$\mathbf{q} = (q_a)_{a \in \mathcal{A}}$ where $q_a \geq 0$ and $\sum_{a \in \mathcal{A}} q_a \leq Q$. The utility that the agent a derives from the quantity q is denoted $v_a(q)$. We assume that v_a is non-negative, i.e., $v_a(q) \geq 0$, and non-decreasing in q. We also assume that each agent has some minimum demand d_a which is of value to her and the additional benefit beyond this amount is only small (the property of diminishing returns). We discretize demands at the precision $\epsilon > 0$, i.e., we assume that d_a is a positive integer multiple of ε. The agent derives no value for an allocation smaller than $d_a - \varepsilon$. Then, within an ε amount, the agent's value dramatically increases to some value $v_a(d_a) > 0$. We formalize the diminishing returns beyond the demand d_a using a slope parameter $0 \leq \lambda < 1$:

$$\frac{v_a(y) - v_a(x)}{y - x} \leq \lambda \cdot \frac{v_a(d_a)}{\varepsilon} \text{ for } y > x \geq d_a,$$

i.e., we assume that beyond d_a, the utility grows at a rate slower by a factor of at least λ compared with the initial jump. We say that the agent a is *satisfied* if she receives an amount $q \geq d_a$.

2.1 Mechanism

Agents from the set $\mathcal{A}$ apply for an allocation. Our mechanism asks agents to submit their demands d_a. The submission of the agent a will be referred to as a *bid* and denoted b_a. We assume that the bids $\{b_a\}$ are integer multiples of ε, but possibly different from the true demands $\{d_a\}$. Our mechanism selects an allocation $\hat{\mathbf{q}}$ which assigns only three possible values to agents: $\hat{q}_a \in \{0, \hat{q}, \hat{q} + \varepsilon\}$, for some $\hat{q} \in \mathbb{R}$. The value $\hat{q}$ is the largest integer multiple of ε such that all submission $b_a \leq \hat{q}$ can be satisfied. Specifically, let $\mathcal{M}(q)$ denote the set of agents with submitted demands at most q: $\mathcal{M}(q) = \{a \in \mathcal{A} : \ b_a \leq q\}$, then: $\hat{q} = \max \{q \in \varepsilon\mathbb{Z} : \ |\mathcal{M}(q)| q \leq Q\}$.

All of the bids $b_a \leq \hat{q}$ receive $\hat{q}$ amount of the resource. Let m denote the corresponding number of satisfied submissions, i.e., $m = |\mathcal{M}(\hat{q})|$. We have an excess resource amount of $Q - m\hat{q}$. When $Q - m\hat{q} \geq \hat{q} + \epsilon$ we distribute the excess among agents with $b_a = \hat{q} + \varepsilon$ as follows. Let k denote the number of submissions with $b_a = \hat{q} + \varepsilon$, i.e., $k = \big|\{a \in \mathcal{A} : \ b_a = \hat{q} + \varepsilon\}\big|$. Let

$$\hat{k} = \min \left\{ \left\lfloor \frac{Q - m\hat{q}}{\hat{q} + \varepsilon} \right\rfloor, \left\lfloor \frac{k}{1 + \lambda} \right\rfloor \right\} \ .$$

We choose a random subset of $\hat{k}$ agents among k. Thus, each individual agent is chosen with probability $\hat{k}/k$, and each of the chosen agents receives $\hat{q} + \varepsilon$ of the resource. Note that by definition $\hat{k} \leq \left\lfloor \frac{Q - m\hat{q}}{\hat{q} + \varepsilon} \right\rfloor$ and thus we always obtain a valid allocation (we never redistribute more resource than available after giving $\hat{q}$ to the initial m agents). Since each of the k agents is chosen with probability at most $1/(1 + \lambda)$, it can be proved that agents with lower true demands have no incentive to over-report. This random distribution of excess resource among agents with $b_a = \hat{q} + \epsilon$ ensures that a constant fraction of agents is satisfied even when their bids are tied, unlike VCG and other classical solutions.

2.2 Properties

Our mechanism has two key properties. First, it is incentive-compatible, i.e., agents have no incentives to lie. Second, it satisfies the number of agents which is at least $1/(2 + 2\lambda)$ fraction of the optimal allocation. Note that if the truthfulness is not a concern, the *smallest bid first* allocation is optimal [3]. The reduction in the number of satisfied agents compared with the optimum is the price we pay for incentive compatibility. The guarantee ranges between 25% (for $\lambda = 1$) and 50% (for $\lambda = 0$). However, since this guarantee is based on the worst-case analysis, in practice the mechanism can be much closer to the (non-truthful) optimum.

For lack of space we omitted the proofs of incentive-compatibility and of the approximation bound of the optimum. The empirical evaluation of performance in a variety of settings was omitted as well. In Figure 1, we present the performance of our mechanism relative to the optimal allocation for an increasing total amount of resource Q. The number of agents was 100; their demands were uniformly random integers between 1 and 20; we used $\epsilon = 1$ and assumed $\lambda = 0.5$. The graph shows that our mechanism satisfies a number of agents much closer to the optimum than the loose theoretical bound of 33%, which we would obtain for $\lambda = 0.5$. Similar results were obtained for a fixed quantity ($Q = 200$) and a varying number of agents from zero to 200.

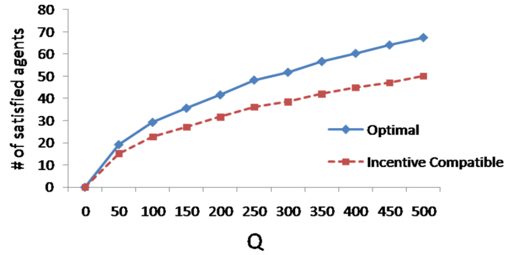

Figure 1: The number of satisfied agents for varying amount of available resource Q.

3. EXTENSIONS

In many resource allocation applications, some users should be entitled to receive larger proportions of resource than others, i.e., some users may have a higher *priority* [3]. In our settings such users are expected to bring more revenue to the service provider. We model the differential entitlement by assigning each agent a a priority $p_a \geq 1$, proportional to the expected revenue. Given a set of submitted demands $\{b_a\}$ and a set of priorities $\{p_a\}$, we require that the mechanism satisfies the submitted demand b_a only after satisfying all the submitted demands $b_{a'}$ with $\frac{b_{a'}}{p_{a'}} < \frac{b_a}{p_a}$, i.e., resource is redistributed in the order of decreasing per-unit revenue (beginning with the largest per-unit revenue). Our mechanism can be adjusted to include this form of priorities while preserving truthfulness.

Another extension to our mechanism, considers the case when the resource is allocated to user agents periodically over multiple rounds. We assume that beside the demand for an amount of resource, agents also have a time limit after which they are not willing to wait for the service (as in a real-time allocation system [3]), agents cannot manipulate their arrival times and their deadlines, and their utility is constant between the arrival and the deadline.

4. CONCLUSIONS

We propose an incentive-compatible mechanism for resource-allocation systems in which the system's expected revenue from satisfying different agents is equal. It is guaranteed to satisfy a number of agents within a constant factor of the optimal (but not necessarily truthful) allocation. Our empirical study demonstrates that the number of satisfied agents is much closer to the optimum than our theoretical bound. Our mechanism can be generalized to systems with priorities and to multi-round allocation.

Acknowledgment: This research has been supported by a MURI award through ARO W911NF-08-1-0301.

5. REFERENCES

[1] J. D. Hartline and T. Roughgarden. Optimal mechanism design and money burning. In *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 75–84, 2008.

[2] N. Nissan, T. Roughgarden, E. Tardos, and V. Vaziriani. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[3] A. S. Tanenbaum. *Modern Operating Systems*. Prentice Hall, 2'nd edition, 2001.