An Optimal Bidimensional Multi-Armed Bandit Auction for Multi-unit Procurement

(Extended Abstract)

Satyanath Bhat
Indian Institute of Science
Bangalore 560012, India
satya.bhat@gmail.com

Shweta Jain
Indian Institute of Science
Bangalore 560012, India
jainshweta@csa.iisc.ernet.in

Sujit Gujar
École polytechnique fédérale de Lausanne
Lausanne, Switzerland
sujit.gujar@epfl.ch

Yadati Narahari
Indian Institute of Science
Bangalore 560012, India
hari@csa.iisc.ernet.in

ABSTRACT

We study the problem of a buyer (aka auctioneer) who gains stochastic rewards by procuring multiple units of a service or item from a pool of heterogeneous strategic agents. The reward obtained for a single unit from an allocated agent depends on the inherent quality of the agent; the agent's quality is fixed but unknown. Each agent can only supply a limited number of units (capacity of the agent). The costs incurred per unit and capacities are private information of the agents. With known qualities, a) we provide the characterization for any Bayesian incentive compatible (BIC) and Individually rational (IR) mechanism, and b) we propose an optimal, truthful mechanism 2D-OPT. To learn the qualities in addition, a) we provide sufficiency conditions for an allocation rule to be stochastic BIC and IR, and b) we design a novel learning, stochastic BIC and IR mechanism, 2D-UCB.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents; I.2.6 [**Learning**]: Parameter Learning

Keywords

Optimal mechanism design, multi-armed bandit mechanism, multi-unit procurement, strategic agents

1. INTRODUCTION

Consider a hospital (auctioneer) interested in procuring several units of a generic drug from various pharmaceuticals who can supply limited quantities at different costs. The quality of the procured drug from a supplier is inherent to the supplier. Motivated by this, we consider a procurement scenario where a buyer wishes to procure multiple units of a service or item from a pool of heterogeneous agents with unknown qualities, privately held costs, and privately held limited capacities. Our goal is to design a procurement auction that learns the qualities of the agents, elicits true costs and capacities from the agents, and maximizes the expected utility of the auctioneer. In a non-strategic version, the classical Multi-Armed-Bandit (MAB) techniques can be used [4]. On the other hand, if all the agents have the same known quality (homogeneous) but with strategic costs and capacities, the auctioneer can deploy the techniques in [2, 3].

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

2. NOTATION AND PRELIMINARIES

Let L denote the number of units to be procured from an agent pool $N = \{1, 2, \ldots, n\}$. Let $q_i \in [0, 1]$, $c_i \in [\underline{c}_i, \overline{c}_i]$ and $k_i \in [\underline{k}_i, \overline{k}_i]$ represent the quality, the true cost and the true capacity of agent i. Auctioneer obtains an expected reward of Rq_i on procuring an unit from agent i where R is a fixed positive real number. Let $\hat{c}_i \in [\underline{c}_i, \overline{c}_i]$ and $\hat{k}_i \in [\underline{k}_i, k_i]$ denote the reported cost and the reported capacity respectively. We assume that the agent cannot over-report his capacity, because it is detected eventually when the agent fails to deliver. If all the parameters are known, then one can solve the following optimization problem which maximizes the utility of the auctioneer:

$$\max \sum_{i=1}^{n} \left(x_i R q_i - t_i \right) \text{ s.t. } x_i \in \{0, 1, \ldots, \hat{k}_i\} \ , \ \sum_i x_i \leq L, \quad (1)$$

where, x_i represents the number of units that are procured from an agent i and t_i denotes the payment given to an agent i. We assume an independent private value model. Let $f_i(c_i, k_i)$ and $F_i(c_i, k_i)$ denote the joint probability density function and cumulative density function respectively which are common knowledge. Let X and T denote the expected allocations and expected payments when expectation is taken over bids of other agents.

3. AUCTION WITH KNOWN QUALITIES

We begin with the characterization for any BIC, IR mechanism with heterogeneous known qualities.

Theorem 1 *A mechanism is BIC and IR iff* $\forall i \in N$,
1. $X_i(\hat{c}_i, \hat{k}_i; q)$ *is non-increasing in* $\hat{c}_i$, $\forall q$ *and* $\forall \hat{k}_i \in [\underline{k}_i, k_i]$.
2. $\rho_i(\hat{c}_i, \hat{k}_i; q) \geq 0$, *and non-decreasing in* $\hat{k}_i$ $\forall q$ *and* $\forall \hat{c}_i$
3. $\rho_i(\hat{c}_i, \hat{k}_i; q) = \rho_i(\overline{c}_i, \hat{k}_i; q) + \int_{\hat{c}_i}^{\overline{c}_i} X_i(z, \hat{k}_i; q) dz$

Using Theorem 1, we arrive at an allocation and payment rule which ensures optimality, BIC and IR.

Theorem 2 *Suppose the allocation rule maximizes*

$$\sum_{i=1}^{n}\int_{\underline{c}_1}^{\bar{c}_1}\cdots\int_{\underline{c}_n}^{\bar{c}_n}\int_{\underline{k}_1}^{\bar{k}_1}\cdots\int_{\underline{k}_n}^{\bar{k}_n}\left(Rq_i-\left(c_i+\frac{F_i(c_i|k_i)}{f_i(c_i|k_i)}\right)\right)$$
$$x_i(c_i,k_i,c_{-i},k_{-i})f_1(c_1,k_1)\dots f_n(c_n,k_n)\,dc_1\dots dc_n\,dk_1\dots dk_n$$

subject to conditions 1 and 2 of Theorem 1. Also suppose

$$T_i(c_i,k_i;q)=c_iX_i(c_i,k_i;q)+\int_{c_i}^{\bar{c}_i}X_i(z,k_i;q)dz \qquad (2)$$

then such a payment scheme and allocation scheme constitute an optimal auction satisfying BIC and IR.

Analogous to the literature on optimal auction [2, 3], we assume regularity on our type distribution as follows.

Definition 1 (Regularity) *We define the virtual cost function $\forall i \in N$ as $H_i(c_i,k_i) := c_i + \frac{F_i(c_i|k_i)}{f_i(c_i|k_i)}$. We say that a type distribution is regular if $\forall i$, H_i is non-decreasing in c_i and non-increasing in k_i.*

2D-OPT mechanism: Based on Theorem 2 and using the assumption of regularity, we propose an optimal, DSIC and IR mechanism, 2D-OPT [1], which allocates units based on non-increasing order of virtual costs and the payments are externality like which computes eq. (2).

4. AUCTION WITH UNKNOWN QUALITIES

We now discuss a set of natural properties which a learning mechanism ideally has. It also turns out that these properties are sufficient to ensure BIC and IR.

Definition 2 (Well-Behaved Allocation Rule) *An allocation rule x is called a Well-Behaved Allocation if:*

1. *Allocation to any agent i for the unit being allocated in round j, x_i^j, depends only on the agent's bids and the reward obtained for j units that are procured by the auctioneer so far and is non decreasing in terms of costs.*
2. *For the allocation in round j and for any three distinct agents $\{\alpha,\beta,\gamma\}$ such that j^{th} round unit is allocated to β. A change of bid by agent α should not transfer allocation of j^{th} round unit from β to γ if other quantities are fixed.*
3. *x_i is non-decreasing with increase in capacity k_i*

Property 1 states that the allocation should not depend on any future rewards which are not observed. Property 2 is similar to Independent of Irrelevant Alternatives (IIA) property i.e. if an agent i changes his bid then it should not affect the allocations of other agents. Property 3 states the allocation rule doesn't penalize an agent with higher capacity.

Theorem 3 *For every well-behaved allocation rule, there exists a transformation that produces the transformed allocation ($\tilde{x}$) and payment ($\tilde{t}$) such that the resulting mechanism $\mathcal{M}=(\tilde{x},\tilde{t})$ is stochastic BIC and IR.*

An example of such a transformation is Algorithm 2. The transformation mechanism ensures truthfulness as it can be shown that the expected payment satisfies eq. (2) [1].

We now propose mechanism 2D-UCB (Algorithm 3),under regularity assumption, which procures one unit at a time, learns the quality and makes the allocation similar to 2D-OPT on the basis of learnt qualities so far. The payment is computed with the help of transformed mechanism given in Algorithm 2.

Theorem 4 *2D-UCB is stochastic BIC and IR.*

ALGORITHM 1: Self-resampling Procedure

Input: bid $\hat{c}_i \in [\underline{c}_i, \bar{c}_i]$, parameter $\mu \in (0,1)$
Output: (α_i,β_i) such that $\bar{c}_i \ge \alpha_i \ge \beta_i \ge \hat{c}_i$

- **with probability** $(1-\mu)$, $\alpha_i \leftarrow \hat{c}_i$, $\beta_i \leftarrow \hat{c}_i$
- **with probability** μ
  - Pick $\hat{c}'_i \in [\hat{c}_i, \bar{c}_i]$ uniformly at random.
  - $\alpha_i \leftarrow recursive(\hat{c}'_i)$, $\beta_i \leftarrow \hat{c}'_i$
- **function** Recursive($\hat{c}_i$)
- **with probability** $(1-\mu)$, return $\hat{c}_i$
- **with probability** μ, pick $\hat{c}'_i \in [\hat{c}_i, \bar{c}_i]$ uniformly at random, return Recursive($\hat{c}'_i$)

ALGORITHM 2: Mechanism Transformation

Input: $\forall i$, bids $\hat{c}_i \in [\underline{c}_i, \bar{c}_i]$, $\hat{k}_i \in [\underline{k}_i, k_i]$, parameter $\mu \in (0,1)$, allocation rule x
Output: Allocation rule $\tilde{x}$ and the payment rule $\tilde{t}$

- Obtain modified bids as (α,β) using Algorithm 1
- Allocate according to $\tilde{x}(\hat{c},\hat{k}) = x(\alpha(\hat{c}),\hat{k})$
- $\tilde{t}_i(\hat{c},\hat{k}) = \hat{c}_i\tilde{x}_i + P_i$, $P_i = \begin{cases} \frac{x_i(\alpha(\hat{c}),\hat{k})}{\mu\mathcal{F}'_i(\beta_i(\hat{c}_i),\hat{c}_i)}, & \text{if}\beta_i(\hat{c}_i) > \hat{c}_i \\ 0, & \text{otherwise.} \end{cases}$

ALGORITHM 3: 2D-UCB Mechanism

Input: $\forall i \in N$, $\hat{c}_i \in [\underline{c}_i, \bar{c}_i]$, $\hat{k}_i \in [\underline{k}_i, k_i]$, $\mu \in (0,1)$, R
Output: A mechanism $\mathcal{M} = (x,t)$

- $\forall i \in N$, $\hat{q}_i^+ = 1$, $\hat{q}_i^- = 0$, $n_i = 1$
- Obtain modified bids as (α,β) using Algorithm 1
- Allocate one unit to all agents to estimate quality $\hat{q}$
- $\hat{q}_i = \tilde{q}_i(i)/n_i$, $\hat{q}_i^+ = \hat{q}_i + \sqrt{\frac{1}{2n_i}ln(t)}$
- **for** $t = n$ *to* L **do**
  - Compute $H_i = \alpha_i + \frac{F_i(\alpha_i|\hat{k}_i)}{f_i(\alpha_i|\hat{k}_i)}$
  - $i = \arg\max_{\{js.t.k_j > n_j\}} R\hat{q}_j^+ - H_j$ and $\hat{G}_i = R\hat{q}_i^+ - H_i$
  - **if** $\hat{G}_i > 0$ **then**
    - Procure from i, update $\hat{q}_i$, $\hat{q}_i^+ = \hat{q}_i + \sqrt{\frac{2}{n_i}ln(t)}$
  - **else**
    - break \\ Don't allocate future units to anyone
- $\tilde{T}_i = \hat{c}_i n_i + P_i$, where, $P_i = \begin{cases} \frac{1}{\mu}n_i(\bar{c}_i - \hat{c}_i), & \text{if}\beta_i > \hat{c}_i \\ 0, & \text{otherwise.} \end{cases}$

---

5. SIMULATIONS

We evaluate 2D-UCB via simulations and compare the expected utility per unit given by 2D-UCB against the optimal benchmark 2D-OPT which is fully aware of underlying quality. We also compare against an $\varepsilon-$separated mechanism which allocates first εL units to all the agents irrespective of their bids. The rest are allocated using 2D-OPT. The simulations indicate that all the mechanisms yield average utilities per unit which asymptotically converge to 2D-OPT [1]. The performance of 2D-UCB however is superior in the sense that it approaches 2D-OPT faster.

REFERENCES

[1] S. Bhat, S. Jain, S. Gujar, and Y. Narahari. An optimal bidimensional multi-armed bandit auction for multi-unit procurement. *CoRR*, abs/1502.06934, 2015.

[2] S. Gujar and Y. Narahari. Optimal multi-unit combinatorial auctions. *Operational Research*, 13(1):27–46, 2013.

[3] G. Iyengar and A. Kumar. Optimal procurement mechanisms for divisible goods with capacitated suppliers. *Review of Economic Design*, 12(2):129–154, 2008.

[4] L. Tran-Thanh, S. Stein, A. Rogers, and N. R. Jennings. Efficient crowdsourcing of unknown experts using bounded multi-armed bandits. *Artificial Intelligence*, 214(0):89 – 111, 2014.