Incentive-Aware Learning for Large Markets*

Mohammad Mahdian
Google Research
mahdian@google.com

Vahab Mirrokni
Google Research
mirrokni@google.com

Song Zuo
Tsinghua University
songzuo.z@gmail.com

Abstract

In a typical learning problem, the goal is to use training data to pick one model from a collection of models that optimizes an objective function. In many multi-agent settings, the training data is generated through the actions of the agents, and the model is used to make a decision (e.g., how to sell an item) that affects the agents. An illustrative example of this is the problem of learning the optimal reserve price in an auction. In such cases, the agents have an incentive to influence the training data (e.g., by manipulating their bids in the case of an auction) to game the system and achieve a more favorable outcome. In this paper, we study such *incentive-aware learning* problem in a general setting, and show that it is possible to approximately optimize the objective function under two assumptions: (i) each individual agent is a "small" (part of the market); and (ii) there is a cost associated with manipulation. For our illustrative application, this nicely translates to a mechanism for setting approximately optimal reserve prices in auctions where no individual agent has significant market share. For this application, we also show that the second assumption (that manipulations are costly) is not necessary, since we can "perturb" any auction to make it costly for the agents to manipulate. Finally, we prove that in the case of auction pricing, without any assumption, it is impossible to compete with the optimal revenue.

1 Introduction

Machine Learning is the science of computing a model or a hypothesis (from a fixed hypothesis space) that best describes data, or more generally, optimizes a given objective function over the data. This is done using a collection of observed historical data points, called *training data*. Many of the applications of machine learning are in environments that are *game theoretic* in nature, i.e., they involve multiple self-interested agents whose actions affect the data points that the machine learning algorithm observes, and who are affected by the outcome selected by the machine learning algorithm. A prime example of this is automated auction marketplaces such as the online advertisements market, where the agents are bidders who can influence the observed data point through the act of bidding. In this example, machine learning algorithms are often used to fine-tune various parameters in the auction (e.g., the reserve price) that have direct monetary impact on the bidders.

Learning in such game theoretic environments can be quite challenging: Agents are strategic and seek to optimize their own utility functions. If the outcome of the learning task has some impact on the utility of these agents, they might try to manipulate the training data to push the learning algorithm to make a decision to their advantage. The main challenge is to develop learning algorithms that encourage agents to behave truthfully, resulting in a correct outcome. Learning in the presence of such strategic agents has been studied the past, however known results are either for specific estimation problems [Dekel et al., 2010, Meir et al., 2012, Hardt et al., 2016], or mainly in the context of prior-free mechanism design [Lu et al., 2006, Balcan et al., 2005] which does not take advantage

*A full version of this paper could be found at https://ssrn.com/abstract=3057435.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

of any training data, or with a small number of non-strategic samples [Cole and Roughgarden, 2014]. In this paper, we study this problem in a simple (general) model of *incentive-aware learning* from a training data that strips away the learning complexity of the task to focus on the strategic component. A motivating application for this model is the problem of learning auction parameters such as the reserve price. In our model, the goal is to select one mechanism from a (potentially infinite) family of mechanisms (called the *ground set*) that optimizes an objective function (e.g., the revenue). Our main result is that it is possible to approximately optimize the objective function in this setting under the following two assumptions:

1. Individual agents are "small".[2]
2. The mechanisms in the ground set are strongly incentive compatible, i.e., not only misreporting values is not better (for the agent) than reporting truthfully, but also it is worse by a specified margin.

A formal definition of the above properties in the general case is given in the full version. Under the above conditions, we can prove that there is an approximately incentive compatible mechanism whose value of the objective function is close to that of the best mechanism in the ground set.

We argue that these assumptions are reasonable in many practical settings. The first assumption holds in large markets where no agent has significant market power. For example, online advertisement marketplaces generally satisfy this. In fact, there is a growing literature in economics on market design for large markets that designs market mechanisms under similar assumptions [Kojima and Pathak, 2009, Budish, 2011]. Furthermore, as we will prove later in the paper, in the case of auction reserve pricing, without such an assumption it is not possible to get any positive result. Specifically, we show that in this model, it is not possible to design an incentive compatible auction that collects revenue more than the vanilla second price auction with zero prior knowledge.

The second assumption holds in situations where there is enough fluctuations in the environment so that a buyer (e.g., an advertiser) cannot safely misreport their data (e.g., their bid) without affecting the outcome to his detriment. For example, in a second price auction, if there is enough variance in the highest competing bid, overbidding runs the risk of winning an item at a price higher than its value, and underbidding runs the risk of not winning an item that is priced below its value. Such variance commonly exists in practice, which may, for example, come from the uncertainty in the prediction of the click-through-rates by the search engine in cost-per-click auctions, the changes of other buyers' bidding strategies, the random throttling, etc. On the other hand, for clean theoretical models, we will prove that in the motivating application of reserve price learning for auctions, even if the mechanisms in the ground set initially do not satisfy this condition: It is possible to modify them by adding a small "perturbation" so that the resulting mechanisms satisfy the above condition and the amount of revenue loss due to the perturbation is small.

The algorithm we propose is simple and practical: instead of picking the mechanism in the ground set that maximizes the value of the objective function (which would have been the natural thing to do in absence of incentive constraints), we use a randomized algorithm that picks a "soft" maximum among the mechanisms. Intuitively, this eliminates cases where a small change in the value of the objective function for a mechanism causes a sudden change in the output of the algorithm. The bulk of the work is in analyzing this mechanism and showing that it has the required incentive property and that it achieves an objective value close to the optimal.

Our contributions and techniques In summary, our contributions in this paper are as follows:

- We formalize a framework of *incentive-aware learning* for a major class of learning problems.
- Under this framework, we introduce the novel concept of *strong incentive compatibility*, which enables us to show a general randomized learning strategy that can "learn" an approximately optimal mechanism from a given ground set of truthful mechanisms that satisfy the strong incentive compatible property. In particular, the learning process is

[2] A "small" agent might have significant impact to the learning objective by adopting different strategies, but her willingness to misreport is *relatively small* compared with the optimal objective achievable by the designer within the ground set. Intuitively, it means that the power of the designer is relatively "large" compared with agents.

guaranteed to be approximately truthful and the degree of truthfulness as well as the approximate optimality becomes stronger as the buyer market power becomes weaker.

- More specifically, we present how this learning strategy can be used in an important industry application, the reserve price optimization in second price auctions. In particular, we will manually introduce the strong incentive compatibility of second price auctions with reserves by "perturbing" the auction a little bit.
- Finally, to complement our previous results, we show that the exact truthfulness and more revenue than vanilla second price auctions with zero-prior knowledge cannot be achieved simultaneously.

From technical point of view, as discussed earlier, our general learning algorithm is based on a "soft" maximum function that selects approximately optimal mechanisms from the ground set. Although the "soft" maximum function is similar to that of McSherry and Talwar [2007] applied in the context of differential privacy, techniques to prove the approximate incentive compatibility are quite different. In particular, the approximate incentive compatibility in McSherry and Talwar [2007] is derived from the differential privacy property of the mechanism, such as digital good auctions with unlimited supply, where the effect of one's misreporting on the total revenue is bounded by some constant.

However, such differential privacy property does not hold for general learning problems, such as the illustrative example that we consider, i.e., reserve pricing for single-item auctions, where we allow each buyer to misreport in m parallel single-item auctions simultaneously and the effect of one's misreporting on the total revenue cannot be bounded. Such a difference makes the problem more challenging and does require different techniques and concepts — strong incentive compatibility. In fact, strong incentive compatibility is critical to enabling the "soft" maximum based learning algorithm to work for the problems where the differential privacy property does not hold.

In order to apply our general result to the auction pricing application, we need some assumptions that are not automatically satisfied. To resolve this problem, we introduced some novel techniques to "perturb" the auctions given in the ground set. Intuitively, high-level ideas therein are to remove the discontinuity and eliminate the weak dominance in the system while sacrificing a small fraction of revenue.

Finally, in order to prove the impossibility result for auction pricing, we start with the observation that second price auction is Pareto optimal (except for some trivial corner cases) in the single sample case (only one auction and one item sold via the auction). Then we carefully generalize it to multiple sample cases by induction.

Related Work Balcan et al. studied the problem of designing digital good auctions (hence unlimited supply) via the technique of separating the market into two parts and applying the prices learned from each part to the other [Balcan et al., 2005]. Such a technique, in fact, creates many non-strategic samples. However, it is still limited to applications in separable environments. In particular, for single item auctions, the generalization — (i) separating buyers into different groups, (ii) learning reserve prices from bids in each group, and (iii) applying the reserves for other groups — can never be better than simply running a second price auction. Because the reserve price learned from each group can never be more than the highest bids within this group and applying such a reserve to other groups cannot be better than simply putting the buyers together and run a second price auction, where the second highest bid is no less than the reserve.

Another line of studies around prior-free auctions initiated by Cole and Roughgarden consider the design of approximately optimal mechanisms with oracle accesses to non-strategic samples of the prior distributions, while the number of such accesses is limited [Cole and Roughgarden, 2014, Devanur et al., 2015, Morgenstern and Roughgarden, 2015, Huang et al., 2015, Gonczarowski and Nisan, 2016, Balcan et al., 2017].

Dekel et al. and Meir et al. study the most classic learning tasks, regression [Dekel et al., 2010] and classification [Meir et al., 2012], in game theoretic environments with some mild assumptions. In particular, the objectives in both the regression and classification problems are restricted to minimizing the loss functions. In contrast to such specific learning tasks, we consider learning with arbitrary objective with strategic agents having arbitrary utility functions.

For more related works, we refer the readers to the full version.

References

Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Learning prices for repeated auctions with strategic buyers. In *Advances in Neural Information Processing Systems*, pages 1169–1177, 2013.

Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Repeated contextual auctions with strategic buyers. In *Advances in Neural Information Processing Systems*, pages 622–630, 2014.

M-F Balcan, Avrim Blum, Jason D Hartline, and Yishay Mansour. Mechanism design via machine learning. In *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, pages 605–614. IEEE, 2005. 1, 3

Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. Sample complexity of multi-item profit maximization. *arXiv preprint arXiv:1705.00243*, 2017. 3

Eric Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011. 2

Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 243–252. ACM, 2014. 2, 3

Ofer Dekel, Felix Fischer, and Ariel D Procaccia. Incentive compatible regression learning. *Journal of Computer and System Sciences*, 76(8):759–777, 2010. 1, 3

Nikhil R Devanur, Zhiyi Huang, and Christos-Alexandros Psomas. The sample complexity of auctions with side information. *arXiv preprint arXiv:1511.02296*, 2015. 3

Hu Fu, Nicole Immorlica, Brendan Lucier, and Philipp Strack. Randomization beats second price as a prior-independent auction. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 323–323. ACM, 2015.

Yannai A Gonczarowski and Noam Nisan. Efficient empirical revenue maximization in single-parameter auction environments. *arXiv preprint arXiv:1610.09976*, 2016. 3

Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. Strategic classification. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 111–122. ACM, 2016. 1

Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 45–60. ACM, 2015. 3

Yash Kanoria and Hamid Nazerzadeh. Dynamic reserve prices for repeated auctions: Learning from bids. 2014.

Fuhito Kojima and Parag A Pathak. Incentives and stability in large two-sided matching markets. *The American Economic Review*, 99(3):608–627, 2009. 2

Pinyan Lu, Shang-Hua Teng, and Changyuan Yu. Truthful auctions with optimal profit. In *International Workshop on Internet and Network Economics*, pages 27–36. Springer, 2006. 1

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on*, pages 94–103. IEEE, 2007. 3

Reshef Meir, Ariel D Procaccia, and Jeffrey S Rosenschein. Algorithms for strategyproof classification. *Artificial Intelligence*, 186:123–156, 2012. 1, 3

Jamie H Morgenstern and Tim Roughgarden. On the pseudo-dimension of nearly optimal auctions. In *Advances in Neural Information Processing Systems*, pages 136–144, 2015. 3