Research Statement

Rad Niazadeh | rad@cs.stanford.edu | (607)-3795744

The advent of online marketplaces with large quantities of highly detailed user data and massive computational resources has been a blessing for operations research. Data and computation provide an immense opportunity for learning and optimization to improve such systems, in particular increasing their market efficiency, revenue, or computational efficiency. However, on the way to achieve these goals, the platforms face several barriers. For instance, how can Amazon learn the prices for its AWS cloud platform over time using past data, given that outlier users (with excessive willingness-to-pay) can mislead off-the-shelf learning algorithms used for pricing? How does uberPOOL manage to match a group of riders to drivers, given that some riders might leave the system in the near future? How can FCC guarantee the strategyproofness of its spectrum auction, given that wireless companies have (very) creative ways of strategizing to manipulate this complex market? How should Microsoft Bing sell its ad space, given that these days ads have multiple lines and configurations instead of just being three lines of text?

The above listed limitations stem from either intrinsic *uncertainties* that exist in an application domain (because of the stochastic nature or real-time aspect of the application), or having to deal with *incentives* (because of strategic users), or existing *combinatorial constraints* (because of complex structures or network effects). In a nutshell, my research identifies scenarios such as the above in various online marketplaces, and shows that despite these limitations, *there is a hope for designing algorithms/market mechanisms that perform almost as well or require almost as few resources, as if those limitations had not existed.* While current methods are broadly amplifying our understanding of many of these scenarios, I approach them with a fresh interdisciplinary perspective from computer science that comes from mixing novel ideas borrowed from machine learning, combinatorial optimization, and probability theory.

Throughout the course of my academic career, my work has led to desirable outcomes for both theoreticians and practitioners. They span from settling challenging open problems [1–3], and designing state-of-the-art approximation algorithms [4–6], to designing fast and efficient market algorithms for the industry, e.g., for the Microsoft Bing Ad Auction [7], Microsoft Azure cloud platform [3, 8], and Yahoo! Ads [9] (all with competitive performance on real data besides theoretical guarantees). Next, I elaborate on some of my research highlights, grouped into two categories based on their focus: *revenue management and market design*, and *optimization and machine learning*. Then I explain how my past work lays a foundation for my future agenda by providing details of ongoing projects, collaborations (e.g., with the Uber matching team and the Bing Ad Auction team) and a few future research directions.

Revenue Management and Market Design

Pricing in cloud computing and online retailing. The quintessential examples of market mechanisms with simple semantics, while targeting to maximize revenue, are various forms of pricing for selling items or services. Part of my research studies such mechanisms, both in the dynamic and static settings.

Dynamic pricing. Cloud platforms like AWS or Microsoft Azure have *spot markets*, i.e., a second market for selling the remaining unused resources in data centers across the globe. Because of its ad-hoc nature, the prices of spot instances can vary with time. To devise such a pricing, one prominent approach is to use no-regret online learning algorithms, either in the stochastic (e.g., Besbes and Zeevi [16]) or the adversarial (e.g., Kleinberg and Leighton [17]) settings (see [18] for other seminal work). An undesired property of these algorithms is scaling of their additive regrets with the maximum willingness-to-pay, a critical drawback in an application domain like spot instances with many high-value outlier users.

To address this issue, in collaboration with Microsoft [3], we propose a new family of adversarial online learning algorithms. We reduce dynamic pricing (and its extensions to dynamic multi-buyer auctions) to a purely learning theoretic framework, termed as *multi-scale learning*, where the reward of each action has its own maximum possible value (e.g., the action of offering price p has maximum value p). By designing a novel Online Mirror Descent (OMD) algorithm, we achieve a new style of regret bounds where the regret with respect to each action scales linearly with its own maximum value. As a result, we obtain the first learning-based dynamic pricing with multiplicative approximations (and not only additive regret). We then evaluate their *convergence rates*, i.e. the number of rounds required to obtain $(1-\epsilon)$ fraction of the revenue of the best price in hindsight. Surprisingly, we show our rates are the best one can hope for, as they match the minimum number of i.i.d. offline samples required to obtain $(1-\epsilon)$ fraction of the optimal revenue of an unknown value distribution (known as the *sample complexity of auctions* e.g., in Cole and Roughgarden [19]). On the theory side, our techniques were later found to be helpful for the famous K-server problem [20]. On the applied side, our proposed method guided Microsoft for its next generation of cloud platforms, in particular a variant where users bid and the platform runs an auction.

Static pricing. Consider a monopolist seller, e.g., an online retailer like Amazon, who wants to sell one item to a group of buyers. The theory of microeconomics, particularly the Nobel prize winning work of Myerson [21], proves a complicated and asymmetric form for the optimal revenue auction in this setting. However, in reality, Amazon runs a simple non-discrimantory mechanism, i.e. posting a single (termed as *anonymous*) take-it-or-leave-it price, and it works very well. The mystery of this discrepancy prompts the following question: how good is the revenue of the anonymous pricing versus the optimal auction? Settling the approximation factor for this problem has remained open for the last half decade, since the initiation of the "simple vs. optimal" trend in mechanism design research [22]. In [2], we cracked this open problem by comparing the revenue of anonymous pricing to a standard upper bound on optimal revenue known as ex-ante pricing revenue. We prove the worst-case ratio between these two quantities is e and our result is tight. Our work has also important implications in mechanism design for agents with multi-dimensional preferences (e.g., for multiple items; cf. Chawla et al. [23]).

Universal recipe for reducing strategic computation to non-strategic computation. When designing systems for human users, e.g., an online market, a designer's triumph is to achieve *Incentive Compatibility (IC)*, i.e., although users could misreport their preferences, it is not in any user's best interest to do so. A fundamental question is then the following: does running an IC mechanism require more computational resources than running an algorithm for the same problem? In other words, is there a computationally efficient reduction from mechanism design to algorithm design? Such a reduction, if it existed, would be an exceedingly powerful tool. System designers could ignore issues of private information and strategic behavior, and instead focus on coming up with algorithms that achieve their objectives in an environment where all information is public. Those algorithms could then be transformed by a general-purpose reduction into mechanisms working as desired even when users of the system are strategic.

For maximizing social welfare, the celebrated result of Vickrey, Clarke and Groves [24–26] in theoretical microeconomics gives such a general purpose reduction, but only if we can solve the non-strategic welfare problem *optimally.* My research tackled the striking problem of transforming *any* arbitrary non-optimal algorithm, e.g., a heuristic algorithm working well in practice, into an incentive compatible mechanism with negligible loss in welfare. In a work with Dughmi, Hartline and Kleinberg [1], we resolve a five-year-old open question in this area: there is a polynomial time reduction from Bayesian incentive compatible mechanism design to Bayesian algorithm design for welfare maximization problems.

Unlike prior results, our reduction achieves exact incentive compatibility for problems with multidimensional and continuous type spaces. The key novel ingredient in our reduction is generalizing the literature on *Bernoulli factories* in probability theory [27, 28]. See our invited survey in SIGecom Exchanges [10] or my ACM magazine article [11] for more details and coverage around this work.

Practical revenue management for modern sponsored search. As online ad offerings become increasingly complex, with multiple size configurations and layouts available to advertisers, the sale of web advertising space faces new challenges. Standard ad auction formats, such as the Generalized Second Price (GSP) [29], do not immediately extend to these settings. Morover, truthful combinatorial auctions, such as the Vickrey-Clarke-Groves (VCG) [24–26], yield unacceptably low revenue and hence are hindering the critical goal of achieving the highest revenue in this multi-billion dollar industry.

In collaboration with the Microsoft Bing Ad auction team [7], we proposed using *core pricing*, an intriguing concept introduced by Day and Milgrom [30, 31], to increase the revenue of the FCC's spectrum auction. However, to implement this method for very time sensitive practical use cases such as real-time ad auctions, we need truly fast algorithms for finding appropriate core prices. Prior work in economics has either studied heuristics that lack performance guarantees or solutions that are based on prohibitively slow convex programs. Our main result is a fast combinatorial algorithm that surprisingly finds an approximate core pricing with quasilinear number of calls to a welfare oracle, and hence beats the running time of existing solutions in the literature. We also conclude that core pricing yields more revenue; we justify this claim experimentally using the Microsoft Bing Ad Auction platform, where we find that core pricing generates almost 85% and 18% more revenue than VCG and simple GSP, respectively.

Optimization and Machine Learning

Optimal approximation for continuous submodular maximization. A continuous submodular function generalizes the classic set function submodularity to continuous domains: a non-concave function whose second-order partial derivatives with respect to every two different coordinates are non-positive. This class of functions appears in every corner of operations research, as well as machine learning and finance. For instance in revenue management under network effects, the Hotels.com problem listed in the introduction paragraph can be modeled as continuous non-monotone submodular maximization over a hypercube.

In joint work with Roughgarden and Wang [4], we study the above fundamental problem. Our main result is the first 2-approximation algorithm for continuous non-monotone submodular maximization; this factor is the best possible for algorithms that only query the function at polynomially many points. This result extends the classic result of Buchbinder et al. [32] on optimal approximation for non-monotone submodular set function maximization to general submodular functions on any (continuous) conic lattice. Interestingly, unlike prior work, our key ingredient comes from bringing a game theoretic perspective to this purely algorithmic question. Roughly speaking, our algorithm goes over coordinates one-by-one, and then by solving a stylized related zero-sum game (per coordinate) finds the right value at this coordinate.

Practical algorithms for hierarchical clustering. Hierarchical Clustering (HC) is a widely used exploratory data analysis tool, ubiquitous in phylogeny, finance and social network analysis. This clustering technique represents a given dataset as a binary tree where each leaf represents an individual data point. As a result, it provides richer information at all levels of granularity simultaneously, compared to more traditional flat clustering approaches like k-means or k-median. Hence, it has been extensively used by practitioners for information retrieval, data mining, and unsupervised machine learning.

In a series of works with Charikar and Chatziafratis [5, 6, 12], we design algorithms for HC, while also analyzing the popular practical algorithms in this domain (e.g., *average-linkage* or *single-linkage*). Our approach is inspired by recent developments in viewing this unsupervised learning task as an optimization problem [33]. On the applied side, we develop a framework for incorporating structural constraints, i.e., restrictions on the final clustering tree identified by domain experts. On the theory side, we study the limitations of average-linkage as the most popular agglomerative clustering used by practitioners. Then, we propose new efficient algorithms based on semidefinite programming with provably better guarantees. More recently, in an ongoing work, we design very fast algorithm when the data has geometric structures, which would be of interest to both statisticians and biologists who deal with big data.

Stronger approximation guarantees in stochastic online optimization. Several decision making problems in operations research are modeled as stochastic online optimization. Richard Bellman's "principle of optimality" [34] suggests using a natural dynamic programming formulation to find the *online optimum* policy in these settings, which requires exponential time. One remedy is then to prove approximation guarantees by comparing an online policy to the *offline optimum*, i.e., the solution of an omniscient prophet with no uncertainties about the future, which a fortiori implies the same approximation guarantee against the online optimum policy. However, it is known that this approach has limitations in many settings. In a series of works with Amin Saberi and Uber matching team [13–15] , we show how to overcome these limitations in various domains by approximating the online optimum policy directly.

Consider a profit-maximizing planner who wants to allocate a limited supply of resources to agents arriving over time, subject to combinatorial constraints. One approach is to model this problem as a stochastic online optimization problem. If we do so, thanks to the literature on *prophet inequalities*, it is known that no better than a constant 2 approximation is possible for this problem with respect to the offline optimum [35]. In [13], we take the first stab at obtaining improved approximations with respect to the online optimum policy for this class of allocation problems. We give a Polynomial Time Approximation Scheme (PTAS) for the *laminar Bayesian online selection*, a special case that includes policy design for firms with productions over time. Our approach is based on rounding the solution of a hierarchy of linear programming relaxations that systematically strengthen the commonly used expected linear program for this problem and approximate the optimum online with any degree of accuracy.

Future Research Directions

Ride-sharing matching and stochastic online optimization. To continue my investigation on improved approximations in stochastic online optimization, I am exploring new future directions stemming from *on-demand transportation* with Amin Saberi, in collaboration with the Uber matching team.

These days, Uber strives to offer lower-priced modified services that are more efficient from the perspective of the platform. One new such modification is adding a wait time (e.g., before pooling the ride). Though these wait times help with thickening the market and overall efficiency, they can lead to computationally harder decision-making problems for the platform. As an example, imagine during the wait-time of the current rider the next rider suddenly decides to close her app; this rider is not available anymore for matching. How should the platform incorporate this future uncertainty in its algorithm?

We have modeled this problem as a two-step stochastic matching [15]. Our preliminary result shows the existence of a simple primal-dual matching algorithm that achieves the optimal approximation with respect to the offline optimum solution. Interestingly, we show that by switching to the online optimum benchmark (which we show is computationally hard to compute), better approximations are possible. We are currently pushing for testing these algorithms on Uber data and extending our model to capture more real aspects of the actual problem. As an ambitous goal, we are also trying to apply our techniques to other stochastic online optimization problems in operations research (e.g., online stochastic matching [36]).

Market design for on-demand food delivery. Beyond transportation, other forms of *on-demand economy* are on the rise. The fast-moving technology companies competing in this arena have developed new models that are transforming industries like grocery and restaurant, which have historically been slow to innovate. I envision engaging in this new emerging area in my future research plan.

As an example, consider the operations of a food delivery platform like DoorDash [37]. I believe many problems related to this platform are at a sweet spot between theory and practice to work on. At this point, the matching part of the DoorDash platform (between Dashers and the merchants) and its recommendation system (between the customers and the merchants) are run separately. How should the platform jointly operate these two? Can the recommendation system change the demand distribution, e.g., reduce the demand spikes for food delivery, and help with the efficiency of the Dasher-merchant matching? Due to the acceptance of longer delays in food delivery, DoorDash has more opportunity to batch, i.e. assign one Dasher to work on more than one service simultaneously. How to design such mechanisms for increasing market efficiency? How to use data to help with answering these questions?.

Assortment planning for virtual retailers. I would like to dig into the new challenges of the burgeoning field of *assortment planning*, a close cousin of traditional mechanism design that studies how a retailer should pick a subset of products for sale under a particular customer behavior model. Specifically , I am interested to study assortment questions related to the new generations of retailers: in what ordering a completely virtual grocery store like Amazon fresh should sort the products on its website (besides assorting a subset of them) so as to increase the revenue? In designing assortments for both the Amazon website and AmazonGo stores, what models of user behavior are credible and are justifiable through data, besides more traditional models such as logit or nested logit models? I would like to explore more around these questions by incorporating real-data and through potential collaborations with Amazon.

My Related Publications (alphabetical author ordering)

[1] Shaddin Dughmi, Jason Hartline, Robert Kleinberg, and **Rad Niazadeh**. Bernoulli factories and black-box reductions in mechanism design. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 158–169. ACM, 2017. Extended version in *Journal of the ACM (under-review).*

[2] Saeed Alaei, Jason Hartline, **Rad Niazadeh**, Emmanouil Pountourakis, and Yang Yuan. Optimal auctions vs. anonymous pricing. *Game and Economic Behavior (special issue on selected papers from STOC/FOCS/)*, 2018. Preliminary version in *Proceedings of the IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, 2015.

[3] Sébastien Bubeck, Nikhil R Devanur, Zhiyi Huang, and **Rad Niazadeh**. Multi-scale online learning and its applications to online auctions. *Journal of Machine Learning Research (accept with minor revision)*, 2018. Preliminary version in *Proceedings of the 2017 ACM Conference on Economics and Computation (EC).*

[4] **Rad Niazadeh**, Tim Roughgarden, and Joshua Wang. Optimal algorithms for continuous non-monotone submodular and DR-submodular maximization. In *Proceedings of the 32nd Conference on Neural Information Processing Systems (NIPS) (accepted for full oral presentation)*, 2018. Extended version in *Journal of Machine Learning Research (under-review).*

[5] Vaggos Chatziafratis, **Rad Niazadeh**, and Moses Charikar. Hierarchical clustering with structural constraints. In *Proceeding of the 36th International Conference on Machine Learning (ICML)*, pages 773–782, 2018.

[6] Moses Charikar, Vaggos Chatziafratis, and **Rad Niazadeh**. Hierarchical clustering better than average-linkage. In *Proceeding of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2019.

[7] Jason Hartline, Nicole Immorlica, Mohammad Reza Khani, Brendan Lucier, and **Rad Niazadeh**. Fast core pricing for rich advertising auctions. *Operations Research (revise and resubmit)*, 2018. Prelimnary version in *Proceedings of the 2018 ACM Conference on Economics and Computation (EC)*, 2018.

[8] Shuchi Chawla, Nikhil Devanur, Janardhan Kulkarni, and **Rad Niazadeh**. Truth and regret in online scheduling. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)*, pages 423–440. ACM, 2017.

[9] Christopher A Wilkens, Ruggiero Cavallo, and **Rad Niazadeh**. Gsp: The cinderella of mechanism design. In *Proceedings of the 26th International Conference on World Wide Web*, pages 25–32. International World Wide Web Conferences Steering Committee, 2017.

[10] Shaddin Dughmi, Jason Hartline, Robert Kleinberg, and **Rad Niazadeh**. Bernoulli factories and black-box reductions in mechanism design. *ACM SIGecom Exchanges*, 16(1):58–71, 2017.

[11] **Rad Niazadeh**. Algorithms versus mechanisms: how to cope with strategic input? *XRDS: Crossroads, The ACM Magazine for Students*, 24(1):20–23, 2017.

[12] Vaggos Chatziafratis, **Rad Niazadeh**, Moses Charikar, and Grigory Yaroslavtsev. Fast algorithms for hierarchical clustering on big geometric data. *ongoing work*, 2018.

[13] Nima Anari, **Rad Niazadeh**, Amin Saberi, and Ali Shameli. Nearly optimal pricing algorithms for production constrained and laminar bayesian selection. *arXiv preprint arXiv:1807.05477*, 2018.

[14] **Rad Niazadeh**, Amin Saberi, and Ali Shameli. Prophet inequalities vs. approximating optimum online. In *Proceedings of the 14th International Conference on Web and Internet Economics (WINE)*, 2018.

[15] **Rad Niazadeh** and Amin Saberi. Prophet vs. mortal: Ride-sharing matching with stochastic riders. *ongoing work*, 2018.

Other References

[16] Omar Besbes and Assaf Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, 57(6):1407–1420, 2009.

[17] Robert Kleinberg and Tom Leighton. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, page 594. IEEE Computer Society, 2003.

[18] Arnoud V den Boer. Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in operations research and management science*, 20(1):1–18, 2015.

[19] Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 243–252. ACM, 2014.

[20] Sébastien Bubeck, Michael B Cohen, Yin Tat Lee, James R Lee, and Aleksander Madry. k-server via multiscale entropic regularization. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 3–16. ACM, 2018.

[21] Roger B Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[22] Jason Hartline and Tim Roughgarden. Simple versus optimal mechanisms. In *Proceedings of the 10th ACM conference on Electronic Commerce*, pages 225–234. ACM, 2009.

[23] Shuchi Chawla, Jason Hartline, and Robert Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM conference on Electronic Commerce*, pages 243–251. ACM, 2007.

[24] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

[25] Edward H Clarke. Multipart pricing of public goods. *Public choice*, 11(1):17–33, 1971.

[26] Theodore Groves. Incentives in teams. *Econometrica: Journal of the Econometric Society*, pages 617–631, 1973.

[27] MS Keane and George L O'Brien. A Bernoulli factory. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 4(2):213–219, 1994.

[28] Şerban Nacu and Yuval Peres. Fast simulation of new coins from old. *The Annals of Applied Probability*, 15(1A):93–115, 2005.

[29] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *The American economic review*, 97(1):242–259, 2007.

[30] Robert Day and Paul Milgrom. Core-selecting package auctions. *International Journal of Game Theory*, 36(3-4):393–407, 2008.

[31] Lawrence M Ausubel, Paul Milgrom, et al. The lovely but lonely vickrey auction. *Combinatorial Auctions*, 17:22–26, 2006.

[32] Niv Buchbinder, Moran Feldman, Joseph Seffi, and Roy Schwartz. A tight linear time (1/2)-approximation for unconstrained submodular maximization. *SIAM Journal on Computing*, 44(5): 1384–1402, 2015.

[33] Sanjoy Dasgupta. A cost function for similarity-based hierarchical clustering. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 118–127. ACM, 2016.

[34] Richard Bellman. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences*, 38(8):716–719, 1952.

[35] Brendan Lucier. An economic view of prophet inequalities. *ACM SIGecom Exchanges*, 16(1):24–47, 2017.

[36] Vahideh H Manshadi, Shayan Oveis Gharan, and Amin Saberi. Online stochastic matching: Online actions based on offline statistics. *Mathematics of Operations Research*, 37(4):559–573, 2012.

[37] DoorDash Food Delivery. `https://www.doordash.com/`.