Beyond Nash Equilibrium: Mechanism Design with Thresholding Agents

Wen Shen

Background: In many real-world scenarios, a principal needs to coordinate multiple self-interested agents that each has private information about their preferences to implement desirable system-wide outcomes. These agents' individual objectives are often not (fully) aligned with the principal's. For instance, a transportation regulatory authority typically aims to reduce traffic congestion caused by the wide use of private cars, while commuters often prefer to stick with their cars instead of shared rides due to concerns such as privacy, comfortability and safety [1, 2]. A powerful tool for motivating self-interested agents to cooperate is to offer incentives for their efforts (e.g., cooperation, sacrifices) according to some allocation and payment rules. The engineering approach of implementing such incentive rules is called *Mechanism Design*. Since its inception, mechanism design has witnessed great successes in many critical domains including transportation sectors [3, 1], power grids [4, 5], online marketplaces [6, 7, 8], education [9], health care [10], and social networks [11, 12].

Challenges: Despite its promising prospects in addressing some of the most challenging societal issues, mechanism design has not leveraged its full potential. There are four challenges that need to be addressed before it becomes fully fledged. First, traditional approaches to designing mechanisms typically require substantial prior knowledge about participating agents' value distributions. Nevertheless, mechanism designers usually do not have such knowledge in hand. Second, mechanism designers often must assume that the agents are fully rational. However, this assumption is often violated in many cases due to factors such as limited computational power, cognitive biases, and time constraints [13]. Third, efficient mechanism design relies on directly revealed preferences. In many real-world environments, people are reluctant to reveal their preferences directly due to privacy concerns [14]. Forth, while much research has been focused on efficient or optimal mechanisms, little attention has been given to mechanisms that are resistant to manipulations such as false-name attacks, and collusion.

Related Work: New lines of research have begun to address the four challenges. To relax the requirement of agents' value distributions, Devanur and Hartline(2009) introduced the problem of *prior-free mechanism design*. In general, there are three methods to prior-free mechanism design: deterministic empirical distribution [16], random sampling [15], and consensus estimates [17, 18]. While random sampling can provide a good approximation of optimal mechanisms in many scenarios, deterministic empirical distribution and consensus estimates may perform arbitrarily poor if the feasibility constraints or the consensus-estimate functions are not satisfied. All the three methods require direct revelation of agents' preferences, making them less appealing to be deployed in practice. To relax the rationality assumption, Ghosh and Kleinberg (2014) introduced the concept of "simple agents" where agents only reason on whether to participate in or not. Their work assumes that the agents have homogeneous valuations. It remains unknown how to extend their approach to the scenarios where agents have different value distributions. Most research on mechanism design restricts their attention to direct mechanisms where agents must explicitly reveal their preferences. In practice, however, posted-price mechanisms are more favorable due to its efficiency and ease of implementation [20]. Besides, direct mechanisms are also usually vulnerable to manipulations such as false-name attacks [21] and collusion [22]. To address the four challenges, a new unified framework to mechanism design is needed.

A Behavioral Approach to Mechanism Design: In this dissertation research, we begin to explore a behavioral approach to mechanism design that integrates a series of new techniques (e.g., indirect mechanism, cut-off policies for modeling agents' preferences, and contests) to address these challenges. The behavioral method includes the following aspects: preference modeling and reasoning, privacy-preserving mechanism design, and manipulation-resistant mechanism design.

- **Preference Modeling and Reasoning:** Previous research on mechanism design usually assumes that agents are fully rational. However, this assumption is often not grounded [23, 2]. To address this problem, we proposed to use cutoff policies to model agents' decision-making processes [8]. In doing so, each agent has a private threshold value to decide to participate in the mechanism or not. This method relaxes the assumption that each agent is fully rational and allows mechanism designers to model agents' valuations with greater flexibility instead of making strong assumptions about their rationality. In repeated interactions (e.g., repeated auctions), it is also convenient for the mechanism designers to learn and make inference about agents' valuation or types with preference data using the cutoff policies. When designing manipulation-resistant mechanisms, the mechanism designers often must estimate agents' costs of performing a specific action (e.g., false-name attacks). In many cases, however, the designers are unable to make an accurate estimation for every agent. We introduced a parameter called *cost coefficient* to help model agents' costs [12]. The cost coefficient is privately known to an agent only. This approach relieves the mechanism designers' burden of precisely calculating agents' costs of efforts. It also makes it possible for the designers to learn agents' cost functions through empirical data in repeated interactions.

- **Privacy-Preserving Mechanism Design:** Efficient mechanism design typically requires agents to reveal their valuations or preferences directly. However, this requirement is often problematic for two reasons. First, agents may be unaware of their accurate valuations. They may overestimate or underestimate their valuations. Second, even if they know their actual values, they might be reluctant to disclose the information to the mechanism designers due to privacy concerns or for fear that the designers will use this information in an unwanted way in the future. To address this issue, the designers can implement indirect (i.e., posted-price) mechanisms [1, 12]. This method does not require agents to report their valuations. However, it does require the designers to have prior knowledge about agents' value distributions. When the designers do not have prior knowledge about agents' value distributions, it is possible to design randomized mechanisms to learn the value distributions by offering free items with monotonic increment in the agents' valuations. This method is usually applicable to online scenarios when agents' value distributions are learnable. It may result in a loss of optimality when agents' valuations are sufficiently large.

- **Manipulation-Resistant Mechanism Design:** Many online platforms (e.g., online marketplaces, crowdfunding sites, social media platforms) rely on high-quality online customer reviews to build reputations. However, online ratings are often vulnerable to manipulations. For instance, a user may post fake reviews for a seller in exchange for a free item or a discount. Such activities will undermine customers' confidence or trust in the online platforms. Another example is that a user may create multiple fake accounts to make profits without making any contributions towards promoting the task owners' goals in social media marketing [24]. To address this problem, we proposed to use multi-winner contests to design manipulation-resistant mechanisms [12]. We demonstrated that proper peer competition could significantly reduce the incentives of manipulations (e.g., false-name attacks) as well as increase agents' aggregated efforts in social network marketing. It will be interesting to explore novel contest structures to counter agents' collusive behaviors. Another direction is to design contest mechanisms to maximize agents' aggregated efforts.

Dissertation Statement: I propose a unified behavioral approach that integrates indirectly revealed preferences, threshold policies and contests to mechanism design for online platforms such as ridesharing, crowdfunding, and social networks. I exemplify the behavioral approach via a series of mechanisms that can achieve comparable or better performance than the state-of-the-art methods for the selected domains. These mechanisms relax some of the unrealistic assumptions such as (1) perfect rationality of agents, (2) direct revelation of agents' preferences, and (3) no manipulations (e.g., false-name attacks, collusion) among agents. Specifically, I have studied or will explore the following questions:

- How can we design indirect mechanisms to coordinate strategic agents for social good? To address this problem, I introduced an online mechanism that can provide real-time fare quote for autonomous mobility-on-demand systems.

- How can we model agents' preferences so that we can design effective information structures? To tackle this challenge, I proposed to use cutoff policies to model agents' preferences.

- Can we design mechanisms that counter both false-name attacks and collusion using proper contest structures? If so, to what extent the mechanisms should sacrifice the optimality? To answer this question, I will identify the properties of the contest success functions that can both prevent agents from creating false accounts and forming coalitions.

- Can we design mechanisms that counter the free-rider problems using proper contest structures? To answer this question, I will focus on multi-stage contests.

Timeline and Deliverable: I will follow the timeline described as bellows:

- Indirect mechanism that applies to on-demand ridesharing systems: results were published in IJCAI-16.

- Inform design with agents under cutoff policies: results were published in AAMAS-18.

- Multi-winner contests to counter manipulations in social networks: results were published in AAAI-19.

- In answering the third question, we will submit the results to IJCAI-19 (Deadline: Feb. 25th, 2019).

- In answering the fourth question, we will submit the results to IJCAI-19 (Deadline: Feb. 25th, 2019).
- Finish the first draft of the dissertation by April 30th, 2019.
- Take the final defense before May 30th, 2019.

References

[1] **W. Shen**, Cristina V Lopes, and Jacob W Crandall. An online mechanism for ridesharing in autonomous mobility-on-demand systems. In *Proc. of IJCAI*, pages 475–481, 2016.

[2] **W. Shen**, Rohan Achar, and Cristina V. Lopes. Toward understanding the impact of user participation in autonomous ridesharing systems. In *Proc. of the 2018 Winter Simulation Conference*. IEEE, 2018.

[3] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Mechanism design via optimal transport. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 269–286. ACM, 2013.

[4] Farhad Kamyab, Mohammadhadi Amini, Siamak Sheykhha, Mehrdad Hasanpour, and Mohammad Majid Jalali. Demand response program in smart grid using supply function bidding mechanism. *IEEE Trans. Smart Grid*, 7(3):1277–1284, 2016.

[5] Hongyao Ma, David C Parkes, and Valentin Robu. Generalizing demand response through reward bidding. In *Proc. of AAMAS*, pages 60–68, 2017.

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

[7] Dejun Yang, Guoliang Xue, Xi Fang, and Jian Tang. Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing. In *Proc. of the 18th annual international conference on Mobile computing and networking*, pages 173–184. ACM, 2012.

[8] **W. Shen**, Jacob W Crandall, Ke Yan, and Cristina V Lopes. Information design in crowdfunding under thresholding policies. In *Proc. of AAMAS*, pages 632–640, 2018.

[9] Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American economic review*, 93(3):729–747, 2003.

[10] Alvin E Roth, Tayfun Sönmez, and M Utku Ünver. Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488, 2004.

[11] Bin Li, Dong Hao, Dengji Zhao, and Tao Zhou. Mechanism design in social networks. In *Proc. of AAAI*, pages 586–592, 2017.

[12] **W. Shen**, Yang Feng, and Cristina V. Lopes. Multi-winner contests for strategic diffusion in social networks. In *Proc. of AAAI*, 2019.

[13] Gerd Gigerenzer and Daniel G Goldstein. Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4):650, 1996.

[14] Moni Naor, Benny Pinkas, and Reuban Sumner. Privacy preserving auctions and mechanism design. In *Proc. of the 1st ACM conference on Electronic commerce*, pages 129–139. ACM, 1999.

[15] Nikhil R Devanur and Jason D Hartline. Limited and online supply and the bayesian foundations of prior-free mechanism design. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 41–50. ACM, 2009.

[16] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007.

[17] Andrew V Goldberg and Jason D Hartline. Competitiveness via consensus. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 215–222. Society for Industrial and Applied Mathematics, 2003.

[18] Bach Q Ha and Jason D Hartline. Mechanism design via consensus estimates, cross checking, and profit extraction. *ACM Transactions on Economics and Computation*, 1(2):8, 2013.

[19] Arpita Ghosh and Robert Kleinberg. Optimal contest design for simple agents. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 913–930. ACM, 2014.

[20] Xin Wang, Alan Montgomery, and Kannan Srinivasan. When auction meets fixed price: A theoretical and empirical examination of buy-it-now auctions. *Quantitative Marketing and Economics*, 6(4):339, 2008.

[21] Swaprava Nath, Pankaj Dayama, Dinesh Garg, Y Narahari, and James Y Zou. Threats and trade-offs in resource critical crowdsourcing tasks over networks. In *AAAI*, 2012.

[22] Jean-Jacques Laffont and David Martimort. Mechanism design with collusion and correlation. *Econometrica*, 68(2):309–342, 2000.

[23] **W. Shen**, Alanoud Al Khemiri, Abdulla Almehrezi, Wael Al Enezi, Iyad Rahwan, and Jacob W. Crandall. Regulating highly automated robot ecologies: Insights from three user studies. In *Proc. of HAI 2017*. ACM, 2017.

[24] Vincent Conitzer, Nicole Immorlica, Joshua Letchford, Kamesh Munagala, and Liad Wagman. False-name-proofness in social networks. In *International Workshop on Internet and Network Economics*, pages 209–221. Springer, 2010.