 Wei Tang | wtang.org | wt2359@columbia.edu

The goal of my research is to develop theoretically rigorous and empirically grounded algorithms for decision-making schemes in **dynamic, uncertain, and human-centric** decision-making environments. This goal is naturally interdisciplinary, and my research lies in the intersection of compute science, data science, and operations research, drawing on ideas from machine learning, algorithmic economics, behavioral (social) sciences and optimization.

Due to the uncertainty and dynamic nature of the modern decision-making environments, decision-makers often face the challenge of exploration and exploitation tradeoff. To deal with this challenge, the decision maker needs to utilize the past observations and data to understand the nature of the underlying uncertainties, and explore different options to collect information for future decisions. These observations – including human purchase decisions, online reviews, user search histories, and beyond, are often generated from the decision-maker's interactions with humans. These human involvements create unique challenges. On one hand, humans may be strategic or careless, or have behavioral biases when taking actions. On the other hand, humans may value their privacy and want to be treated fairly. Algorithms that fail to account for these behaviors or considerations can lead to substantial losses in decision-maker's utilities while also adversely affecting human welfare.

By combining theoretical analysis and empirical investigation, I pursue three complementary research themes to address the above challenges. My first research theme centers on **designing new learning algorithms** to balance exploration and exploitation in uncertain dynamic decision-making problems, especially when additional problem structures, e.g., information asymmetries, are present. My second theme focuses on **designing behavior-aware algorithms** with taking into account various human behaviors. To bridge the theory-practice gap, I run behavioral experiments to understand real human behavior in practice. My third theme focuses on **designing socially-aware algorithms**, considering societal impact when deploying the designed algorithms in human-centric environments. My vision is to **harness the capabilities of machine learning (ML) and artificial intelligence (AI) to foster more efficient, and human-friendly decision-making environments.**

Designing Novel Learning Algorithms in Online Decision Making

Motivated by applications in modern marketplaces, one major line of my research has focused on how to design *optimal* and *efficient* learning algorithms for online decision-making problems under uncertainties. These uncertainties usually root in the decision-makers' (e.g., the online platforms, online retailers/sellers) limited knowledge on the users' private preferences/valuations. On the other hand, the decision makers may also have additional information about the products/items they recommend or sell. In below, I detail my two recent works, which have applications in product recommendations and dynamic pricing, on designing novel learning algorithms on addressing the *information asymmetry* between online platforms and customers.

Online Bayesian recommendation: Thanks to the rapid advancement of modern technology, online platforms such as TikTok, Amazon Live, and Netflix can now recommend products to users in real-time. For these platforms, understanding user preferences is key to achieving optimal long-term revenue. However, an ongoing challenge for these platforms is their limited knowledge of users' preferences. Meanwhile a notable feature of these platforms is the *information asymmetry*: the platform often knows well about the recommended products while the users are uncertain about the detailed product characteristics. A crucial question then arises: whether the platforms can utilize their information advantage to better learn users' preferences and thereby achieve optimal long-term revenue?

In my recent work [1], we introduce and study the online Bayesian recommendation problem for a platform who is repeatedly interacting with a population of users through an online recommendation mechanism. We assume a natural information asymmetry between the platform and users — only the platform can privately observe the realized state of the product, whereas all users only have a prior belief about the product. For each user with her own private preference and belief, the platform commits to a recommendation strategy that utilizes his information advantage on the product state to persuade the self-interested user to follow the recommendation. The platform does not know user's preferences and beliefs, and has to use adaptive recommendation strategies to learn user's preferences and beliefs in the process. We focus on designing online learning policies with no *Stackelberg* regret for the platform, i.e., against the optimum policy in hindsight under the assumption that users will correspondingly adapt their behaviors to the benchmark policy. We show that there is an online policy that can achieve $O(\log\log T)$ regret where T is the learning horizon. To our best knowledge, this is the first online algorithm in the problem of learning with information asymmetry that has $O(\log\log T)$ Stackelberg regret. This regret is order-optimal w.r.t. horizon T but has exponential dependency on product states. We then also provide a linear programming-based algorithm that can achieve regret with polynomial dependence on the number of product states but logarithm dependence on T.

Dynamic advertising and pricing with demand learning: Dynamic pricing is a key strategy in revenue management that allows sellers to anticipate and influence demand in order to maximize revenue. When the customer valuation and demand response for a product is a priori unknown, price variation can be used to observe and learn the demand function in order to adaptively optimize price and revenue over time. It has been theoretically and empirically shown that advertisements can serve as a credible signal of the quality or characteristics of the advertised product. Sellers can use advertising to provide partial information about a product in order to better position the product in the market and potentially increase customers' chances of purchasing the product. For example, in the online used-car market, the dealer can advertise the used car by partially disclosing its history report.

In my recent work [2], we use Bayesian persuasion framework to model the effect of an advertising strategy on customers' beliefs about product quality/characteristics and consequently their purchase decisions. Our novel formulation combines a canonical information design framework with dynamic pricing and learning in order to quantify the tradeoffs between the design of the pricing and advertising strategies and their combined impact on the revenue outcomes. Without any apriori knowledge of the demand function, we design a novel and also computationally efficient learning algorithm that achieves an $O(T^{2/3}(m\log T)^{1/3})$ regret bound where m is the cardinality of the discrete product quality domain and T is the time horizon. This regret bound is order-optimal for dynamic pricing within logarithmic factors, which is a special case of our problem. Our results provide managerial implications that learning optimal advertising and pricing has the same learning complexity as in dynamic pricing problems.

Designing Behavior-aware Algorithms and Examining Human Behavior

My second research theme has focused on how to design behavior-aware learning algorithms in online decision-making problems with uncertainty. Modern user behavior has been profoundly influenced by the widespread presence of price comparison sites, user review websites, and Q&A platforms. Below I detail my recent work on designing optimal pricing policy when customer's purchase decision is affected by a reference price, and briefly mention my works on learning from human behavior, and my works on running behavior experiments to better understand human behavior in practice.

Dynamic pricing with averaging reference price: In dynamic pricing applications, the demand for a product is shaped not only by the current selling price but also by customers' price expectation, often manifested through the reference price. Customers typically form this reference price based on previously observed prices. A commonly used mechanism to model reference price is the exponential smoothing mechanism (henceforth, ESM) where the reference price is generated by exponentially weighting all past prices. However, this mechanism has one major disadvantage: it requires the customer to remember all the past prices or at least the last reference price, which may significantly increase customer's cognitive load to make the decisions.

Motivated by real-world practice of online platforms where an average of historical prices is often displayed, in my recent work [3], we adopt an alternative modeling approach where we model the reference price as platform's published averaging price (henceforth, APM). This modeling eases the customers' cognitive load since they don't need to remember all past prices. Unlike ESM which is a *time-invariant* mechanism, APM is a *time-variant* mechanism. Notably, we show that, even under this non-stationary APM, the *markdown* pricing strategy is *exactly* optimal for gain-seeking customers, and is *nearly* optimal for risk-averse customers. Our results provide managerial insights for online platforms on how to design optimal pricing policy when the averaging price is presented to the customers. When it comes to the demand learning, APM also creates unique challenges: under ESM, the fixed-price policy is shown to be approximately optimal for risk-averse customers (thus it suffices to learn the optimal fixed-price); while under APM, the fixed-price policy could be highly suboptimal, and the optimal policy is a dynamic policy. Our results present a novel and efficient *online policy learning* algorithm that can achieve *optimal* $\widetilde{O}(\sqrt{T})$ regret, irrespective of customers' risk attitude.

Learning from biased, (boundedly) rational human behavior: I have studied how to design optimal learning algorithm when integrating human bias into online decision-making [4]. I also studied boundedly rational behavior in information design problems and aim to understand how the human boundedly rational behavior impacts the optimal information policy [5]. When human behavior remains unknown to the designer, I also studied robust information design and robust decision-making schemes [5] that have provable robustness guarantees [6]. In addition, for (Bayesian) rational users , I also studied the game among the markets' competitive information design in a searching market and fully characterized the Nash equilibrium for markets' competitive information revelation [7].

Modeling human behavior via behavioral experiments: In practice, human behavior may be biased or complex. To fill in the gap, I run behavioral experiments to better understand human behavior and develop more realistic behavior model. For example, I conducted behavioral experiments with 400 participants from Amazon Mechanical Turk, to assess how they update beliefs and take actions [8]. I also studied automated information design [9] where we propose HAIDNet, a neural-network-based optimization framework that can account for various human behaviors. In addition, I also run experiments to study the benefits of peer communication in crowdsourcing [10], and how predictive information affects human ethical preferences [11].

Designing Socially-aware Algorithms

Beyond understanding human behavior, algorithms must also incorporate social norms such as fairness and privacy considerations for ensuring sustainable growth and complying with technology-related laws/regulations. With these concerns, a natural question arises: how to design desired market mechanism that obey important social norms, while still allowing us to harness the benefits of machine learning and artificial intelligence? In this line of research, I have studied **privacy preservation** and **fairness concerns** in online decision-making problems.

- In privacy preservation, I studied **how to preserve privacy** in contextual dynamic pricing [12] where I adopt differential privacy as privacy measure to explore the design of differentially private pricing algorithms. The algorihms aim to minimize the regret w.r.t a policy that knows buyers' preferences, while satisfying a privacy guarantee. In this work, I proposed an algorithm that is ε-differentially private and achieves regret $\widetilde{O}(\sqrt{dT}/\varepsilon)$, where d is the feature dimension and T is time horizon. To the best of our knowledge, this work is *the first to address the privacy issues in dynamic pricing problems*. I also studied the **secure** stochastic convex optimization, where the learner aims to optimize the accuracy, i.e., obtain an accurate estimate to the optima, while securing her privacy, i.e., preventing an adversary from inferring what she learned. We use PAC-style notions to formalize accuracy and privacy and characterize the lower/upper bounds of the query complexity for this secure learning problem [13].
- In fairness concerns, I examined the **long-term impact of actions** informed by the consequential decisions, which are often coming up when the human well-being is involved [14]. I also studied social prediction where the data distribution itself changes in response to the deployment of a model [15].

My Research Agenda Going Forward

Building upon my previous research, I plan to expand my efforts to further explore the human-centered algorithm design for the modern dynamic, uncertain decision-making environments. Below, I highlight potential directions for future research that align with my research vision.

Pushing the frontier of online decision-making and machine learning: Following my previous research, I plan to further push the frontier of online decision-making and machine learning by exploring the design of new behavior- and socially-aware learning algorithms for the applications in online marketplaces. Below I mention my two ongoing works in this line of research:

- Personalized pricing, which specifies prices based on individual characteristics, has been commonly observed in modern marketplaces. Yet, such personalization/contextualization has invited customers to strategically manipulate their features to get a lower price. In my ongoing work [16], I have started exploring the design of the robust dynamic pricing strategies for strategic customers. I believe that this work will present valuable insights on how to design methods for uncertain environments to guarantee robustness in the face of strategic manipulation.
- In my ongoing work [17], I study the impact of pricing decisions on both online seller's (long-term) revenue and customers' welfare. In practice, sellers might offer varied prices for people from different social groups. But what if people can see prices for other people? How will this affect people's perceptions of whether they are benefited or discriminated against, and subsequently, their buying decisions? To answer these questions, I plan to develop new learning algorithm around these dynamics and understand its resulting society impact.

Algorithm design for decision-making at the onset of generative AI: The rise of generative AI, such as large language models (LLMs), is significantly reshaping the manner in which we approach problems, and the interest in discerning its impact on various decision-making scenarios escalates. In this direction, I will study the theoretic and applied aspects on how to effectively leverage the benefits of the generative AI to tackle the challenges in decision-making problems. In particular, I will study the following questions:

- *Integrating the* **prediction power** *of generative AI into the algorithm design for the uncertain and dynamic decision-making environments:* The rapid development of generative AI has enabled the decision maker to readily leverage the machine learned advice/predictions to improve their understanding of environmental uncertainties. A salient example is the integration of ChatGPT into e-commerce platforms, where the real-time user interaction data is processed to develop predictive models that can further forecast market trends and customer behaviors.[1] However, many challenges arise. For example, the AI advice/predictions are often not free, but costly. Secondly, obtaining heightened prediction accuracy often necessitates more disclosed data to the AI system. These lead to the questions of how to design efficient learning algorithm to balance the cost-benefit tradeoff, and balance the privacy-accuracy tradeoff for the online marketplaces using the predictions from the generative AI? How these tradeoff behave in dynamic environments?
- *Integrating the* **behavior power** *of generative AI into the algorithm design for the human-centric environments:* Algorithms designed for the human-centric environments rely on a good understanding on human behavior. However, it is often expensive and time consuming to conduct a market research to understand the underlying customer behavior. The generative AI provides an alternative way to learn human behavior as it can be used as proxies for a diverse set of humans.[2] Notably, models like LLMs have demonstrated to exhibit a range of biases. With this perspective, I plan to use a behavioral approach, viewing generative AI as a behavioral agent. This will facilitate a deeper understanding and integration of its behavior into the algorithmic design framework.

I believe that answering these questions will have great implications on better designing decision-making algorithms that attempts to harness the benefits of modern ML/AI technologies.

[1] https://www.nytimes.com/2023/03/08/technology/chatbots-disrupt-internet-industry.html

[2] https://www.microsoft.com/en-us/research/blog/using-generative-ai-to-imitate-human-behavior/

References

[1] (α–β) Yiding Feng, **Wei Tang**, and Haifeng Xu. Online bayesian recommendation with no regret. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC)*, pages 818–819, 2022. *Under review at Operations Research.*

[2] (α–β) Shipra Agrawal, Yiding Feng, and **Wei Tang**. Dynamic pricing and learning with bayesian persuasion. *Under Review, and To be submitted to Operations Research*, 2023.

[3] (α–β) Shipra Agrawal and **Wei Tang**. Dynamic pricing and demand learning with averaging reference effect. *Working Paper*, 2023. *To be Submitted to Operations Research.*

[4] **Wei Tang** and Chien-Ju Ho. Bandit learning with biased human feedback. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1324–1332, 2019.

[5] (α–β) Yiding Feng, Chien-Ju Ho, and **Wei Tang**. Rationality-robust information design: Bayesian persuasion under quantal response. *arXiv preprint arXiv:2207.08253*, 2022. *To be submitted to Operations Research.*

[6] **Wei Tang**, Chien-Ju Ho, and Yang Liu. Linear models are robust optimal under strategic behavior. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 2584–2592. PMLR, 2021.

[7] (α–β) Bolin Ding, Yiding Feng, Chien-Ju Ho, **Wei Tang**, and Haifeng Xu. Competitive information design for pandora's box. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 353–381, 2023.

[8] **Wei Tang** and Chien-Ju Ho. On the bayesian rational assumption in information design. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*, pages 120–130, 2021. **Best Paper Honorable Mention**.

[9] Yu Guanghui, **Wei Tang**, Saumik Narayanan, and Chien-Ju Ho. Haidnet: Encoding human behavior in automated information design. *Under Review*, 2023.

[10] **Wei Tang**, Chien-Ju Ho, and Ming Yin. Leveraging peer communication to enhance crowdsourcing. In *The World Wide Web Conference (WWW)*, pages 1794–1805, 2019.

[11] Saumik Narayanan, Guanghui Yu, **Wei Tang**, Chien-Ju Ho, and Ming Yin. How does predictive information affect human ethical preferences? In *ACM Conference on AI, Ethics, and Society (AIES)*, 2022.

[12] **Wei Tang**, Chien-Ju Ho, and Yang Liu. Differentially private contextual dynamic pricing. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1368–1376, 2020.

[13] **Wei Tang**, Chien-Ju Ho, and Yang Liu. Optimal query complexity of secure stochastic convex optimization. *Advances in Neural Information Processing Systems (NeurIPS)*, pages 9771–9781, 2020.

[14] **Wei Tang**, Chien-Ju Ho, and Yang Liu. Bandit learning with delayed impact of actions. *Advances in Neural Information Processing Systems (NeurIPS)*, pages 26804–26817, 2021.

[15] Yatong Chen, **Wei Tang**, Chien-Ju Ho, and Yang Liu. Performative prediction with bandit feedback: Learning through reparameterization. *arXiv preprint arXiv:2305.01094*, 2023.

[16] (α–β) Yiding Feng and **Wei Tang**. Dynamic pricing with strategic customers. *Ongoing work*, 2023.

[17] (α–β) Shipra Agrawal and **Wei Tang**. Fair price discrimination with reference effect. *Ongoing work*, 2023.