KDD 2022 Workshop on Decision Intelligence and Analytics for Online Marketplaces: Jobs, Ridesharing, Retail, and Beyond

Zhiwei (Tony) Qin
Lyft
San Francisco, CA, USA
zq2107@caa.columbia.edu

Liangjie Hong
hongliangjie@gmail.com
LinkedIn Inc.
Sunnyvale, CA, USA

Rui Song
North Carolina State University & Amazon
Raleigh, NC, USA
rsong@ncsu.edu

Hongtu Zhu
University of North Carolina
Chapel Hill, NC, USA
htzhu@email.unc.edu

Mohammed Korayem
mkorayem@gmail.com
CareerBuilder
Toronto, Ontario, Canada

Haiyan Luo
itshaiyan@gmail.com
Indeed
Sunnyvale, CA, USA

Michael I. Jordan
University of California
Berkeley, CA, USA
jordan@cs.berkeley.edu

ABSTRACT

Online marketplace is a digital platform that connects buyers (demand) and sellers (supply) and provides exposure opportunities that individual participants would not otherwise have access to. The KDD-22 Workshop on Decision Intelligence and Analytics for Online Marketplaces: Jobs, Ridesharing, Retail, and Beyond brought together academics and practitioners working on a broad array of online marketplaces to appreciate the diversity in applications, to draw connections to inform decision optimization across different industries, and to discover new problems that are fundamental to marketplaces of different domains. The workshop followed a dual-track format, covering marketplaces in general and job marketplaces specifically.

ACM Reference Format:
Zhiwei (Tony) Qin, Liangjie Hong, Rui Song, Hongtu Zhu, Mohammed Korayem, Haiyan Luo, and Michael I. Jordan. 2022. KDD 2022 Workshop on Decision Intelligence and Analytics for Online Marketplaces: Jobs, Ridesharing, Retail, and Beyond. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), August 14–18, 2022, Washington, DC, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3534678.3542895

1 INTRODUCTION

The Workshop on Decision Intelligence and Analytics for Online Marketplaces: Jobs, Ridesharing, Retail, and Beyond was held at the Walter E. Washington Convention Center in Washington, DC on August 15, 2022. It was co-located with the ACM SIGKDD Conference 2022, which was the first in-person KDD conference since Anchorage in 2019.

1.1 Motivation

An online marketplace is a digital platform that connects buyers (demand) and sellers (supply) and provides exposure opportunities that individual participants would not otherwise have access to. Online marketplaces exist in a diverse set of domains and industries, for example, rideshare (Lyft, DiDi, Uber), house rental (Airbnb), real estate (Beke), online retail (Amazon, Ebay), job search (LinkedIn), and food ordering and delivery (Doordash, Meituan). Besides academia, many companies and institutions are researching on topics specific to their particular domains. The fundamental mechanism of an online marketplace is to match supply and demand to generate transactions, with objectives considering service quality, participants experience, financial and operational efficiency. It is valuable to bring together researchers and practitioners from different application domains to discuss their experiences, challenges, and opportunities to leverage cross-domain knowledge. Online marketplaces are seeing rapid growth and evolution in the recent years, penetrating every facet of our daily life. Operational efficiency has profound impact to our experiences and quality of life. Holding this workshop at KDD this year was timely and anticipating, especially with KDD being a venue for research on topics related to marketplace analytics and optimization in recent years. The goal of this workshop was to offer opportunity to appreciate the diversity in applications, to draw connections to inform decision optimization across different industries, and to discover new problems that are fundamental to marketplaces of different domains.

1.2 Theme

An online marketplace is typically a two-sided (or even three-sided) marketplace, where all participating parties are price-sensitive, and the participation of each party affects the return and experience of

KDD '22, August 14–18, 2022, Washington, DC, USA

ACM ISBN 978-1-4503-9385-0/22/08.
https://doi.org/10.1145/3534678.3542895

the other ones, exhibiting strong network effects. The core topic of the workshop is on data-driven decision-making in online marketplaces. Various issues covered by this workshop include supply-demand-ETA predictions, matching/assignment, supply planning and balancing, dynamic pricing, search/ranking, experimentation, privacy, and variants of these problems under competition. Methods of interest span over deep learning, reinforcement learning, multi-armed bandits, mathematical programming, and stochastic optimization.

This workshop followed a dual-track format. Track 1 covered the issues and algorithms pertinent to general online marketplaces as well as specific problems and applications arising from those diverse domains, such as ridesharing, online retail, food delivery, house rental, real estate, and more. Track 2 focused on the state of the art advances in the computational jobs marketplace. Interesting challenges in this domain include the drastic increase of work from home or remote work, the imbalance between the demand and supply of the job market, the popularity of independent workers, the capability of helping job seekers on their whole job seeking journey and career development, the different objectives and behaviors of all major stakeholders in the ecosystem, e.g. job seekers, employers, recruiters and job agents.

1.3 Relevance to KDD

The operations in online marketplaces are increasingly data-driven. Data mining and knowledge discovery (KDD) is instrumental to the understanding and prediction of supply and demand which an online marketplace is to match. Methods and algorithms for KDD also power the prescriptive analytics (decision optimization) for marketplace operations, for example, how demand and supply should be matched, how supply should be planned in anticipation of demand, and how prices should be set with respect to the dynamic nature of supply and demand. Hence, this workshop has been of significant interest to the SIGKDD community.

2 PARTICIPATION

The main audience of the workshop were researchers and practitioners who were interested in data mining and analytics for marketplace optimization or were from the industries where those online marketplaces operate. They typically came from academia, industry (e.g., ridesharing, online retailing, on-demand delivery), and government agencies (e.g., transportation board). We saw strong interests from the KDD, machine learning, and operations research communities. At peak, the workshop attracted more than 100 participants, making it one of the best attended workshops at KDD 2022.

3 KEYNOTE SUMMARIES

Susan Athey. Susan and her collaborators considered the problem faced by an online marketplace seeking to balance fairness and efficiency in an environment where platform participants may have preferences characteristics of potential partners that may be revealed in photographs. They considered the setting of a microlending marketplace, Kiva, where lenders select among borrowers based on the borrower profiles. They began by using observational data to construct estimates of the effect of features detected in images on funding rates and defaults. They found that lenders preferred images that depict female borrowers, smiling borrowers, and that were not body shots; but these characteristics were not strongly correlated with repayment rates. Although their estimation approach adjusted flexibly for a range of features detected using state-of-the-art computer vision techniques, there was still potential for bias due to information that lenders might be able to discern from photographs that were captured.

To address the potential for confounding, Susan and collaborators conducted a survey experiment with recruited subjects to isolate the impact of image characteristics on lender choices. Experimental subjects chose between fabricated profiles generated using Generative Adversarial Networks that differed in the chosen features. The results of the experiment were consistent with the observational estimates. Finally, they observed that low-performing borrowers and females were more likely to smile and less likely to use body shots, and these choices exacerbated the gender gap on the platform. They simulated outcomes under various counterfactual policies aimed at improving fairness and efficiency by encouraging different choices about profiles. They showed that policies based on image profile recommendations could boost outcomes of the least popular campaigns and reduce overall inequity without sacrificing the number of transactions.

Abraham Bagherjeiran. In over 190 markets, eBay helps millions buyers find unique items from small and large sellers across the world. Empowering buyers to find that perfect item among billions of possibilities anywhere in the world is a difficult challenge. In this talk, Abraham discussed some of eBay's uniquely challenging problems and some of the innovative machine learning solutions his team had built to address them.

Ido Bright. Marketplace companies rely heavily on experimentation when making changes to the design or operation of their platforms. The workhorse of experimentation is the randomized controlled trial (RCT), or A/B test, in which users are randomly assigned to treatment or control groups. However, marketplace interference causes the Stable Unit Treatment Value Assumption (SUTVA) to be violated, leading to bias in the standard RCT metric. In this talk, Ido talked about a new technique for platforms to run standard RCTs and still obtain meaningful estimates despite the presence of marketplace interference. They specifically considered a matching setting, in which the platform explicitly matches supply with demand via a matching algorithm. Their proposed technique is quite simple: instead of comparing the total value accrued by the treatment and control groups, they instead compare each group's average shadow price in the matching linear program. At the heart of their result is the idea that it is relatively easy to model interference in matching-driven marketplaces since, in such markets, the platform intermediates the spillover.

They have proved that, in the fluid limit, their proposed technique corresponds to the correct first-order approximation (in a Taylor series sense) of the value function of interest. They then used this result to prove that, under reasonable assumptions, their estimator is less biased than the RCT estimator. They also presented an extension of this result to supply-chains.

Daniel Hewlett. LinkedIn Talent Solutions provides a variety of applications to job seekers and hirers, including Job Search and LinkedIn Recruiter. In this talk, Daniel described how they had designed LinkedIn's AI systems to support the full range of job and member matching problems across search and recommendations, and both sides of the hiring marketplace. He discussed how they measured hiring outcomes and enabled their models to effectively optimize hiring outcomes. They used transfer learning to provide ML models across all our applications with a shared understanding of member and job content such as profiles and job postings, as well as job-seeking activities.

Haixun Wang. The effectiveness of Search has a substandtial impact on the revenue and growth of an e-commerce business. In this talk, Haixun discussed the current status and challenges of product search. In particular, he highlighted the enormous effort it had taken to create a high-quality product search engine using classical information retrieval methods. Then, he discussed how recent advances in NLP and deep learning, particularly the introduction of large pre-trained language models, may alter the status quo. While embedding-based retrieval has the potential to improve classical information retrieval methods, developing a machine learning-based, end-to-end system for general-purpose web search is still extremely difficult. Nevertheless, Haixun argued that product search for e-commerce may prove to be an area where deep learning could create the first disruption to classical information retrieval systems.

Hongtu Zhu. In this talk, Hongtu introduced a general analytical framework for large scale data obtained from two-sided markets, especially ride-sourcing platforms like DiDi. This framework integrates classical methods including Experiment Design, Causal Inference and Reinforcement Learning, with modern machine learning methods, such as Graph Convolutional Models, Deep Learning, Transfer Learning and Generative Adversarial Network. The goal was to develop fast and efficient approaches to address four major challenges for ride-sharing platform, ranging from demand-supply forecasting, demand-supply diagnosis, RL-based policy optimization, to A-B testing. Each challenge requires substantial methodological developments and inspires many researchers from both industry and academia to participate in this endeavor. All the research accomplishments presented in this talk were based on a series of joint works by a group of researchers at Didi Chuxing and his collaborators.

4 ACCEPTED PAPERS

Submissions of the workshop were solicited from the following (but not limited to) areas:

- supply-demand prediction,
- ETA (expected time-of-arrival) prediction,
- dynamic matching and assignment,
- supply planning and rebalancing,
- dynamic pricing,
- search, ranking, recommendation, and personalization,
- growth and monetization,
- simulation, evaluation and experimentation,

Ten entries were accepted as full papers and presented as 10-minute spotlights during the workshop program. In addition, there are five extended abstracts, which were presented as posters. The full papers were all peer-reviewed by at least two invited program committee members, while extended abstracts were not be reviewed.

The full papers generally fall into these themes:

Ridesharing. **Fairness and Anarchy in Ride-Sharing**. Ian Kash (University of Illinois at Chicago)*; zhongkai wen (The University of Illinois at Chicago); Lenore Zuck (The University of Illinois at Chicago). **Monitoring Disintermediation: Actions Matter More Than Words**. Yingkang Xie (Northwestern University)*; Jingyi Wang (Huolala); Huaiyu Zhu (Huolala).

Recommendation systems. **A Simulator-based Decision-Making Approach to Sequential Recommender Systems with Application in Ride-hailing Platform**. Xiong-Hui Chen (Nanjing University)*; Yang Yu (Nanjing University); Qingyang Li (Didi Research America); Bowei He (City University of Hong Kong); Zhiwei (Tony) Qin (Lyft); Wenjie Shang (AI Labs, Didi Chuxing); Jieping Ye (Didi Chuxing). **An Industry Based Evaluation of Recommender System Performance**. Mohamed H Latif (University College of London)*. **Parametric Empirical Bayes for Predicting Quality in Rating Systems**. Thomas Ma (Stanford University)*; Ramesh Johari (Stanford University); Nikhil Garg (Cornell Tech); Michael S Bernstein (Stanford University)

Search & ranking. **From Pairwise Ranking to Click Prediction on ZipRecruiter Jobs**. Ritvik Kharkar (ZipRecruiter)*. **Leveraging Taxonomies for Solving Named Entity Recognition in Job Search Queries**. Ishan Shrivastava (ZipRecruiter)*; Hengda Shi (ZipRecruiter). **The Unified Experience of Search Ads and Display Ads at Indeed.com**. Riley Ji (Indeed); Ziying Liu (Indeed); Shichuan Ma (Indeed); Yu Sun (Indeed); Xingang Huang (Indeed); Lihong Pei (Indeed); Shu Zhang (Indeed); Haiyan Luo (Indeed)*. **Alternative Job Discovery with Attentive Neural Re-ranking**. Liang Wu (LinkedIn)*; Liangjie Hong (LinkedIn).

Causal inference. **GCF: Generalized Causal Forest for Heterogeneous Treatment Effect Estimation in Online Marketplace**. Shu Wan (Didi Chuxing)*; Chen Zheng (Didi Chuxing, Beijing, China); Zhonggen SUN (Didi Chuxing); Mengfan Xu (Northwestern University); Xiaoqing Yang (Didi Chuxing, Beijing, China); Hongtu Zhu (UNC Chapel Hill); Jiecheng Guo (Didi Chuxing).

All accepted papers can be accessed at https://sites.google.com/view/kdd22onlinemarketplaces/program/accepted-papers

5 CONCLUSION

The Decision Intelligence and Analytics for Online Marketplaces Workshop was among the first KDD workshops dedicated to cross-domain topics on two-sided marketplaces. We have received strong interests and intend to continue this workshop series at future KDD and ACM conferences.