Qingtao Cao

Complex Systems Modeling | Data Scientist | Network Science

darrylcho.14@gmail.com, (626) 628-5806

SUMMARY

I am a Sixth-year Ph.D. student in Industrial Engineering at Northeastern University, working with Professor Babak Heydari. As a research assistant in applied disciplines, I am a fast learner, questions riser, and a solid problem solver to achieve various objectives in my academic career or previous working experience.

To look for insight into big data, I am good at generating different models, from simple regression models to complex machine learning models. Moreover, I have the advanced knowledge to analyze the data from the network perspective, particularly useful for spatial data or social media data.

Apart from data analysis, I am good at developing causal models to explain the causal effect between two dependent factors or two datasets, answering the questions not only about 'HOW to make it works', but also 'WHY it works' to deeply understand the interior logic behind the complex appearance.

EDUCATION

Northeastern University (USA), College of Engineering

- Ph.D. Candidate in Industrial Engineering, MAGICS Lab — GPA: 3.92/4 — 2017 – Current (Expected Apr 2023)
  Advisor: Prof. Babak Heydari
- M.S in Operations Research — GPA: 3.92/4 — 2014 – 2015

South China Normal University (CHN), College of Science

- B.S in Applied Mathematics — GPA: 3.33/4 — 2008 – 2012

RESEARCH PROJECTS

(IAPRA Project) The Mobility Pattern Change in MA Caused by Covid-19 – *Python, SQL* — 2021 – Current

- Used detailed mobility data from SafeGraph to describe the mobility flow over multi-level networks and analyzed the change of the mobility pattern overtime.
- Predicted the further mobility flow on networks by the Dyadic Regression model and the Graph Neural Network model with demographic data.
- Ran the multivariate linear regression model to predict the number of Covid confirmed cases of towns in MA by the mobility, index of the mobility network and demographic data.

Learning Platforms: Adaptive Competition using Deep Reinforcement Learning – *Python, PyTorch* — 2021 – 2022

- Used the Deep learning Approach to solve the new challenge of the platform's governance caused by multi-level interactions between agents in the experimental complex ecosystem.
- By using Deep Reinforcement Learning (DRL), the AI platform can learn a high-performance dynamic adaptive policy to compete with another platform in a multi-sided market.
- By modularizing the strategy action space, the learned policy by DRL can reveal the ability of interpretation and, furtherly, be inferred to enhance the heuristic policy of humans.

Micro-level Social Contact Structures and the Success of COVID-19 National Policies – *Python* — 2020 – 2021

- Created an explainable epidemic model based on a new algorithm that transforms national-average data on micro-level social interaction into a small-scale social network.
- Integrated top-down (National Policies) and bottom-up (Micro-level Social Contact) factors to show how the interaction of these two factors leads to the differences in the success rate of containing the pandemic between countries by the simulation model.
- Using national-average, un-calibrated data made our model computationally more efficient, thus our model is suitable to be used in counterfactual analysis and policy scenario studies.

Theoretical Analysis of Platforms' Competition on Multi-Sided Markets – *Python* — 2018 – 2021

- Combined customers' behavior theories (e.g., bounded rationality) and the dynamic game theory to mathematically model the price competition between platforms on a multi-sided market.
- We proved that even in the presence of cross-sided network effects, the "winner-takes-all" nature of multi-sided platforms can be replaced by the possibility of multiple equilibria.
- Figured out the structure of the platform's local optimal price, consisting of the magnitude of the network effect and the customers' rationality level, in the equilibrium of the competition.

OTHER SELECTED PROJECTS

Cuebiq location data analysis – *Python, SQL* Spring 2022

- With millions of Cuebiq users' track datapoint collected from Cuebiq, we provided insights into human mobility during the pandemic and have fueled research to understand movement patterns during lockdown periods.

Data Visualization and Data Analysis – *R, SQL* Spring 2019

- Visualized and analyzed the modification of the transportation behavior (Taxi and Subway Usage) in New York City after the launch of Uber.

CONFERENCE PRESENTATIONS AND POSTERS

8th International Conference on Computational Social Science (IC2S2, 2022) July 2022 Chicago, IL

- Presentation: Micro-level Social Structures and the Success of COVID-19 National Policies
- Poster: Platform Jumping: Network Analysis of How Political Memes Seeded on Reddit Diffuse to Twitter

8th International Engineering Systems Symposium (CESUN, 2021) October 2021 Charlottesville, VL

- Presentation: Learning Platforms: Adaptive Competition Using Deep Reinforcement Learning

CESUN 2021 Virtual Staying in Touch Event February 2021 Online

- Poster: Design and Governance of Platform Systems

2020 Virtual INFORMS Annual Meeting November 2020 Online

- Presentation: Learning Platforms: Adaptive Competition Using Deep Reinforcement Learning

INDUSTRY EXPERIENCE

Operations Coordinator, 4Excelsior Inc, Los Angeles, USA 2016 – 2017

- Managed the inventory to satisfy the demand for manufacturing and operation activities.
- Reported directly to the COO for daily operations and assisted the COO to design the standard of process in the warehouse.

Financial Analyst Assistant (Internship), E Fund Management Co., Ltd. Guangzhou, China Summer 2015

- Collected the information and data of the domestic market and the global stock market, and then reported it to the senior financial analysts.
- Assisted senior financial analysts to predict the stock price by the regression model.

PUBLICATIONS

- **Cao, Q.**, & Heydari, B. (2022). Micro-level social structures and the success of COVID-19 national policies. Nature Computational Science, 2(9), 595-604.
- Chou, C. A., **Cao, Q.**, Weng, S. J., & Tsai, C. H. (2020). Mixed-integer optimization approach to learning association rules for unplanned ICU transfer. Artificial Intelligence in Medicine, 103, 101806.

WORKING PAPERS

- **Cao, Q.**, Heydari B. Competition Among Two-Sided Platforms with Myopic Agents.
- **Cao, Q.**, Chen, Q., Heydari B. Learning Platforms: Adaptive competition using Deep Reinforcement Learning.
- **Cao, Q.**, Heydari B. How openness the platform needs to be? Innovation and Competition in the multi-sided platform.
- **Cao, Q.**, & Heydari, B**.** When and how mobility networks matter in pandemic forecasting.
- Chatterjee, A., **Cao, Q.**, Sajadi, A., & Ravandi, B. From NetLogo Modeling of Deterministic Random Walk to the Identification of Asymmetric Saturation Time in Erdos-Renyi Graphs.

SKILLS AND INTERESTS

- **Concentration**: Complex Systems Analysis (Multi-sided platforms System and Epidemiology), and Data Scientist
- **Analytical Skills:** Causal Model Generation (epidemic, Economic and human-behavior), Mathematical closed-form solving, Generalized Linear Models, Machine Learning, Network Analysis, Multi-Agents Simulation, Bayesian Inference, Reinforcement Learning and Optimization
- **Computer software**: Python, PyTorch, R, SQL, AWS, NetLogo and AMPL
- **Language:** English (fluent), Cantonese (native), Mandarin (native)

AWARD

- Outstanding Graduate Research Award, College of Engineering, Northeastern University, 2022