Shan Jiang

Ph.D. Student in Computer Science · Research Assistant

660-674, Interdisciplinary Science & Engineering Complex, 805 Columbus Ave, Boston, MA 02120

(+1) 781-502-8799 | sjiang@ccs.neu.edu | shanjiang.me | printfoo | shan-jiang

Education

Northeastern University — *Boston, MA*

Ph.D. in Computer Science — *Sep. 2016 - Present*

- Advisors: Christo Wilson and Alan Mislove

Beijing University of Posts and Telecommunications — *Beijing, China*

B.B.A. in Information Management and Information Systems — *Sep. 2012 - Jun. 2016*

- GPA: 92.5/100 Rank: 1/46

Publications

On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi — *WWW 2018*

Shan Jiang, Le Chen, Alan Mislove, and Christo Wilson

TNCs Today: A Profile of San Francisco Transportation Network Company Activity — *Report 2017*

Joe Castiglione, Tilly Chang, Drew Cooper, Jeff Hobson, Warren Logan, Eric Young, Billy Charlton, Christo Wilson, Alan Mislove, Le Chen, and Shan Jiang

Conflicts in Overlay Environments: Inefficient Equilibrium and Incentive Mechanism — *KSII TIIS 2016*

Jianxin Liao, Jun Gong, Shan Jiang, Tonghong Li, and Jingyu Wang

Interactions of Overlays and Traffic Engineering: Equilibrium and Cooperation without Payment — *GlobeCom 2015*

Shan Jiang, Jun Gong, Jingyu Wang, Jianxin Liao, and Tonghong Li

Competitive Equilibrium and Stable Coalition in Overlay Environments — *LCN 2015*

Shan Jiang, Jianxin Liao, Jun Gong, Jingyu Wang, and Tonghong Li

Combination Feature for Image Retrieval in the Distributed Datacenter — *ICPADS 2014*

Di Yang, Jianxin Liao, Qi Qi, Jingyu Wang, Haifeng Sun, and Shan Jiang

Projects

Crowd Signals for Misinformation Detection on Social Media — *Nov. 2017 - Present*

Our results show that hateful and distrustful signals from crowd responses can help detect "fake news" on social media.

- Wrote crawlers to collect factchecked news data from Snopes/Politifact, and user comments data from Facebook/Twitter/Youtube;
- Processed texts and built sentiment lexicons and clusters using natural language processing techniques;
- Performed statistical tests on the correlation between information veracity and crowd signals;
- Built predictive models based on crowd signals to detect "fake news" and misinformation on social media.

Political Bias in News Sites and Google's Search Results — *Joint work with Ronald Robertson, Sep. 2016 - Feb. 2018*

Our results show that many top-ranked news sites have liberal bias, and Google's search algorithms help alleviate such phenomena.

- Crowdsourced a search ranking dataset based on 200+ participants on Amazon Mechanical Turk;
- Analyzed 113 million Tweets using Apache Spark and computed political bias scores for more than 1 million domains;
- Performed statistical tests on the correlation between Google's ranking and personalization results and bias scores of domains.

Competition and Accessibility of Ridesharing Services — *Joint work with Le Chen, Sep. 2016 - Nov, 2017*

Our results show that ridesharing drivers are biased towards neighborhoods with high minority ratios or low incomes.

- Intercepted and inspected ridesharing mobile apps using man-in-the-middle proxy to build structured requests for data collection;
- Wrote mobile crawlers to collect Uber/Lyft data in San Fransisco/New York City for 2 months using Microsoft Azure clusters;
- Analyzed 10TB+ data using Apache Spark to discover spatiotemporal patterns of ridesharing drivers;
- Performed geo-statistical tests on the accessibility of ridesharing using a set of socio-economic and transportation features.

Competition and Cooperation among Overlay Networks — *Joint work with Jun Gong, Nov. 2013 - Oct. 2015*

Our results show that the competition of overlay networks leads to suboptimal equilibriums, while the cooperation is Pareto-efficient.

- Built game-theoretic models to understand the competition among overlay networks;
- Proposed incentive mechanisms to facilitate the cooperation of network companies, both with and without money transfer;
- Evaluated analytical models based on real-world network topologies with 100+ nodes.

Skills

Languages Python, Java, Javascript, C++, Matlab, R, SQL

Platforms Linux, Apache Spark, Tensorflow

Data Tools pandas, geopandas, scikit-learn, statsmodels, matplotlib, nltk, gensim, keras, etc