

Shan Jiang | Curriculum Vitae

660-674 (#86), Interdisciplinary Science and Engineering Complex, 805 Columbus Ave, Boston, MA 02120

☎ (+1) 781-502-8799 | ✉ sjiang@ccs.neu.edu | 🏠 shanjiang.me | 🖨️ printfoo | 🌐 [shan-jiang](https://shan-jiang.github.io)

Education

Northeastern University

Ph.D. in Computer Science

- Advisor: Christo Wilson

Boston, MA

Sep 2016 - 2021 (Expected)

Beijing University of Posts and Telecommunications

B.B.A. in Management Information Systems

- Rank: 1/46 GPA: 92.5/100

Beijing, China

Sep 2012 - Jul 2016

Experience

Google

Software Engineer Intern @ Fact Check Team, Google AI

New York, NY

Jun 2019 - Aug 2019

- Project: ClaimReview markup (e.g., claim, claimant and verdict) extraction from fact-check articles.
- Hosts: Simon Baumgartner, Abe Ittycheriah and Cong Yu.
- Explored several task formulation possibilities, e.g., language generation under encoder-decoder frameworks, and eventually formulated the task as a sequence tagging problem and conducted several experiments by modifying and fine-tuning BERT models.
- [Productionized](#) codebase with test files and technical documentation.
- Prepared a paper with additional data exploration and model comparison to be submitted to [WWW'20](#).

Dataminr

Research Intern @ AI and Data Science Team

New York, NY

Feb 2019 - Apr 2019

- Project: Crisis sub-event (e.g., burning road after a wildfire) detection on social media for emergency management.
- Hosts: William Groves, Sam Anzaroot and Alejandro Jaimes.
- Built a pipeline model that first scans the Twitter firehose and parses Tweets to dependency trees, then traverses to extract connected noun-verb pairs (e.g., home-burn, house-destroy), and finally clusters similar pairs as sub-events.
- Case-studied California wildfires to understand the temporal cascading (e.g., fire→smoke→pollution) of sub-event networks;
- Published a paper at the AI for social good workshop, i.e., [AISG'19@ICML](#).

Northeastern University

Research Assistant @ Khoury College of Computer Sciences

Boston, MA

Sep 2016 - Present

- Research areas: computational journalism, computational social science, algorithm auditing, information quality.
- Collaborators: Christo Wilson, Alan Mislove, Ronald E Robertson, Kenneth Joseph, etc.
- Collected and analyzed TB-sized data under Spark/Hadoop frameworks and applied statistical and causal models for hypothesis testing.
- Project: Bias in content moderation - Collected a dataset of moderated YouTube comments with their auto-labeled partisan bias and used propensity score models to estimate causal effects. Natural experiments showed that the claim of biased content moderation is yet a mis-perception from correlation to causation. Published a paper at [ICWSM'19](#) and won an [outstanding analysis paper](#) (1/238) award.
- Project: Effects of "fake news" and fact-checking - Collected 5K+ fact-checks from Snopes and PolitiFact and 2M+ comments from Facebook, Twitter and YouTube, and then analyzed linguistic differences in user comments between truthful/fake news and before/after fact-checks. Also leveraged empirical observations to build predictive models on misinformation detection. Published a paper at [CSCW'18](#).
- Project: Audience bias in Google search - Recruited 200+ participants to install browser extensions that collect search data from their computers and estimated audience bias of a website based on its Twitter sharers. Built a visualization system at polarshare.shanjiang.me and published a paper at [CSCW'18](#) that won an [honorable mention](#) (30/1,106) award for best papers.
- Project: Ridesharing competition and accessibility - Intercepted Uber and Lyft's mobile traffic using man-in-the-middle proxy and built structured requests for data collection, and then analyzed 10TB+ data to understand spatiotemporal patterns and inequality of accessibility. Collaborated with SFCTA on a visualization system at tncstoday.sfcta.org and published a paper at [WWW'18](#).

National University of Singapore

Research Assistant @ School of Computing

Singapore

Dec 2015 - May 2016

- Project: Economic modeling of Bitcoin mining under risk aversion.
- Collaborator: Richard TB Ma.

Beijing University of Posts and Telecommunications

Research Assistant @ State Key Lab of Networking and Switching Technology

Beijing, China

Oct 2013 - Dec 2015

- Project: Game-theoretic modeling of overlay networks and traffic engineering.
- Collaborators: Jingyu Wang and Jun Gong.
- Published papers at computer network and system conferences, e.g., [GlobeCom'15](#), [LCN'15](#), [ICPADS'14](#).

Skills

Languages

Python, Java, Scala, C/C++, R, Matlab, SQL, HTML/CSS, JavaScript, Bash

Platforms

Linux, Spark, Hadoop, Git, TensorFlow, PyTorch

Publications

Bias Misperceived: The Role of Partisanship and Misinformation in YouTube Comment Moderation Shan Jiang, Ronald E Robertson and Christo Wilson	ICWSM'19 <i>outstanding analysis paper: 0.4%</i> acceptance rate: 21%
Crisis Sub-Events on Social Media: A Case Study of Wildfires Shan Jiang, William Groves, Sam Anzaroot and Alejandro Jaimes	AISG'19@ICML oral presentation: 18%
Auditing Autocomplete: Suggestion Networks and Recursive Algorithm Interrogation Ronald E Robertson, Shan Jiang, David Lazer and Christo Wilson	WebSci'19 acceptance rate: 24%
Auditing the Partisanship of Google Search Snippets Desheng Hu, Shan Jiang, Ronald E Robertson and Christo Wilson	WWW'19 acceptance rate: 18%
Who's the Guinea Pig? Investigating Online A/B/n Tests in-the-Wild Shan Jiang, John Martin and Christo Wilson	FAT*19 acceptance rate: 24%
Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media Shan Jiang and Christo Wilson	CSCW'18 acceptance rate: 26%
Auditing Partisan Audience Bias within Google Search Ronald E Robertson, Shan Jiang, Kenneth Joseph, Lisa Friedland, David Lazer and Christo Wilson	CSCW'18 <i>honorable mention: 2.7%</i> acceptance rate: 26%
On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi Shan Jiang, Le Chen, Alan Mislove and Christo Wilson	WWW'18 acceptance rate: 15%
Conflicts in Overlay Environments: Inefficient Equilibrium and Incentive Mechanism Jianxin Liao, Jun Gong, Shan Jiang, Tonghong Li and Jingyu Wang	KSII-TIIS'16 impact factor: 0.61
Interactions of Overlays and Traffic Engineering: Equilibrium and Cooperation without Payment Shan Jiang, Jun Gong, Jingyu Wang, Jianxin Liao and Tonghong Li	GlobeCom'15 acceptance rate: 35%
Competitive Equilibrium and Stable Coalition in Overlay Environments Shan Jiang, Jianxin Liao, Jun Gong, Jingyu Wang and Tonghong Li	LCN'15 acceptance rate: 30%
Combination Feature for Image Retrieval in the Distributed Datacenter Di Yang, Jianxin Liao, Qi Qi, Jingyu Wang, Haifeng Sun and Shan Jiang	ICPADS'14 acceptance rate: 30%

Honors and Awards

Outstanding Analysis Paper	for the top analysis paper at ICWSM'19 (1/238)	2019
Honorable Mention	for top papers at CSCW'18 (30/1,106)	2018
Dean's Fellowship	for Ph.D. students at Northeastern University	2016
Outstanding Undergraduate	for top undergraduate students in the city of Beijing	2016
National Scholarship	for top 1% students at Beijing University of Posts and Telecommunications	2014
First-Class Scholarship ×2	for top 2% students at Beijing University of Posts and Telecommunications	2013, 2015

Service

Program Committee	ASONAM (Multidisciplinary Track)	2019
Reviewer	CSCW, ICWSM	2020
	CSCW, ICWSM, CHI	2019
	CSCW, WWW	2018