Shan Jiang | Research Statement

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

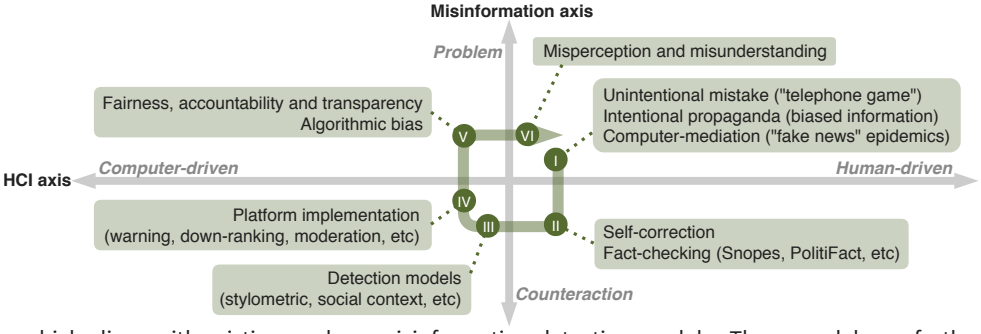
(+1) 781-502-8799 | sjiang@ccs.neu.edu | shanjiang.me | printfoo | shan-jiang

Research Interests

My research lies around social computing, computational social science and (more broadly) human-computer interaction (HCI). My recent work takes an HCI perspective on the misinformation cycle (details later). My research utilizes a diverse toolbox of statistical and causal inference, natural language processing, machine learning, etc.

An HCI Perspective on the Misinformation Cycle

In a "telephone game", players form a line and pass a message from one to another, and more than often, the message changes significantly at the last player. This game speaks about unintentional mistakes that is the nature of interpersonal communication. Such mistake is often amplified in the real world, for people sometimes intentionally promote propaganda using biased information or even twisted facts. With the emergence of computer mediation, this misinformation becomes large-scaled "fake news" epidemics. In CSCW'18a, I investigated how such misinformation affects social media users by analyzing expressed linguistic signals in their comments. I found that "faker" posts significantly discourage reasoned conversation by promoting hate speech and aggressive emotional cues in users' discussions.



Human-driven counteractions of the misinformation problem involves both self-correction and fact-checking by others. The latter evolved to specialized agencies such as Snopes and PolitiFact. In CSCW'18a, I show that though such fact-checking services have corrective effects in-the-wild, it also sometimes "backfire" on audiences who view them as biased and unreliable. Further, I used these accumulated linguistic signals to build predictive models on misinformation, which aligns with existing work on misinformation detection models. These models are further implemented by platforms to build computer-driven counteractions of the misinformation problem. Some of them are algorithmic decision-making such as warning and down-ranking, while others involves a collaboration between human and computers such as content moderation.

This reliance on algorithms raises concerns on the fairness, accountability and transparency (FAT) of these systems, as implementing such algorithms can generate misinformation per se. For example, in CSCW'18b, my collaborators and I conducted controlled audits on Google's search engine ranking algorithms. We showed that search results from politics-related queries have an overall left-leaning audience bias, except for the embedded Twitter component which have a right-leaning bias. This bias affects users perception, and could potentially shifts voting behaviors according to previous in-lab experiments. This discussion on FAT also relates to the opaque nature of many web services, For example, in FAT*'19, my collaborators and I measured online experiments running behind-the-scenes using the Optimizely platform, and we found cases of running controversial experiments such as personalized news headlines without user awareness or consent. Such bias naturally persists beyond the misinformation domain, for example, in WWW'18, I showed that ridesharing services such as Uber and Lyft provide low-quality services at low-come and high-diversity neighborhoods in San Francisco and New York City.

The concerns for FAT is well-justified, but unfortunately often misperceived and misunderstood. Accusations of hypothetical bias circles back to the human-driven misinformation problem. One claim is that search engine creates "filter bubbles" for its users, where its algorithm assume users preference and rank their results to cater users. However, in CSCW'18b, we found no significant differences in the partisan bias of the Google search results delivered to self-identified Democrats or Republicans. Another claim is that "social media is totally discriminating against Republican/Conservative voices (Tweet by Donald Trump)". In my recent work in progress (WIP'19, submitted to CHI'19), I examined the validity of such claim using YouTube as a lens. I found that though comments under right-leaning videos are moderated more heavily on a correlational level, these comments also contain more hateful content. After controlling such confounders using a causal model, no evidence is found to support such claim.

This cycle projects the misinformation problem and counteraction onto the HCI axis, from human-driven, computer-mediated communication to computer-driven decision-making and eventually circling back. My work so far has been empirical evaluation of deployed social computing systems. In future research, I plan to build creative systems to support the counteraction of misinformation.

References

CSCW'18a Shan Jiang, Christo Wilson. "Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media." *Proceedings of the ACM: Human-Computer Interaction (PACMHCI)*, 2(CSCW), Nov 2018

CSCW'18b Ronald E Robertson, Shan Jiang, Kenneth Joseph, Lisa Friedland, David Lazer, Christo Wilson. "Auditing Partisan Audience Bias within Google Search." *Proceedings of the ACM: Human-Computer Interaction (PACMHCI)*, 2(CSCW), Nov 2018

WWW'18 Shan Jiang, Le Chen, Alan Mislove, Christo Wilson. "On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi." In *Proceedings of the 27th International Web Conference (WWW 2018)*, Lyon, France, Apr 2018

FAT*'19 Shan Jiang, John Martin, Christo Wilson. "Who's the Guinea Pig? Investigating Online A/B/n Tests in-the-Wild." In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency (FAT* 2019)*, Atlanta, GA, USA, Jan 2019

WIP'19 Shan Jiang, Ronald E Robertson, Christo Wilson. A paper investigating whether YouTube comment moderation is biased. Submitted to CHI'19. The title is hidden to compile with double blind policy.