My research interests fall under the broad theme of Computing and Society, spanning the research areas of Social Computing, Information Retrieval, Computational Journalism and Fairness in Machine Learning Algorithms. The problems I study usually have two distinct aspects: an algorithmic aspect and a societal aspect. A common thread in my research lies in collecting and analyzing large scale data available in online systems to understand whether there is any potential cause for concern, and then develop mechanisms to eliminate or reduce the impact. My research agenda is inherently interdisciplinary, and therefore, the tools and theories that I often use in my works come from information retrieval, machine learning, media and journalism studies, computational social choice, social welfare and behavioral economics.

1 Past and Ongoing Research Works

Major online platforms today (such as Amazon, Netflix, Spotify, Uber, LinkedIn or AirBnB) can be thought of as two-sided markets, with multiple stakeholders: (i) producers/providers of goods and services (e.g., artists on Spotify, drivers on Uber, hosts on Airbnb), (ii) customers who pay for them, and (iii) the platform which sits at the center of the ecosystem, mediating transactions between producers and customers, and essentially controlling the information access through search, recommendation or matching services. We can further categorize these platforms based on their purpose, e.g., what is being exchanged between producers and customers:

A. Ecommerce platforms (e.g., Amazon, Flipkart) for sale of products.
B. Multimedia streaming platforms (e.g., YouTube, Spotify, Netflix), news aggregation services (e.g., Google News, Dailyhunt) etc. for consuming contents.
C. Ride-hailing platforms (e.g., Ola, Uber), freelance marketplaces (e.g., Fiverr, Upwork), hotel booking platforms (e.g., Booking.com, Airbnb) etc. for availing services.
D. Charity websites (e.g., DonorsChoose, GlobalGiving), crowdfunding platforms (e.g., Kickstarter, Indiegogo) for making donations; and
E. Social media platforms, also for contents but one user can be both producer and customer.

During my PhD and Postdoc, I have been working on two-sided platforms, spanning all the categories listed above, aiming to achieve different goals for different stakeholders. While some of my works are focused on improving the efficiency of the platforms, in others, I have attempted to identify and address the concerns of bias and unfairness for both producers and customers. One key aspect of any two-sided platform is that both customers and producers have preferences, and the platform needs to take into account those preferences while developing search, recommendation or matching algorithms. My focus has been on considering these preferences in a fair manner. This is a marked departure from existing works on the emerging research domain of fairness in machine learning, where the prevalent focus is on fairness in supervised classification or regression tasks. In such tasks, an objective ground truth is present; whereas search, recommendation or matching algorithms deal with subjective preferences. Most of my works have greatly benefited from interactions with my supervisors and collaborators from different institutes. Next, I present my works by grouping them into different problem scenarios.

1.1 Achieving Platform-Level Goals While Recommending News [WWW’17, IR’19]

To ensure newsreaders returning to their sites at regular intervals, news media platforms (and news aggregators) constantly update news stories on their front(home) page round the clock. Such constant updates are hard to manage manually using only human editors, resulting in the need for automated methods for recommending front-page news stories. I considered the setting where the same front-page stories are
shown to all newsreaders visiting a website, and personalized interests are not used for selecting them. When recommending news stories, there are two basic metrics of interest – recency and relevancy. Ideally, recommender systems should recommend the most relevant stories soon after they are published. However, the relevancy of a story only becomes evident as the story ages, thereby creating a tension between recency and relevancy. I analyzed popular recommendation strategies in use today and found that they lead to poor trade-offs between recency and relevancy in practice. I proposed a new recommendation strategy which attempts to optimize on both axes. I further proposed approaches to inculcate diversity in the recommended news which can maintain a balanced proportion of news from different news sections.

1.2 Fairly Representing User Opinion in Social Media Platforms

1.2.1 Fairness in Crowdsourced Recommendation [ICWSM’17, FAT’19a]

Social media users rely on crowdsourced recommendation Trending Topics to find important events and breaking news. Typically, hashtags and key-phrases are recommended as trending when their usage by the crowds suddenly jump at a time. Once a topic is selected as trending, it gets prominently displayed on the social media homepages, thus reaching a large user population. However, I found that the majority of trends on Twitter are promoted by crowds whose demographics differ significantly from Twitter’s overall user population, and certain demographic groups (e.g., middle-aged black female) are severely under-represented in the process. To make the demographic biases of Twitter trends transparent, I developed and deployed a service Who-Makes-Trends where one can find the trending topics in the US and check their promoter demographics. Going over demographic bias, crowd often consists of malicious actors trying to manipulate the recommendations. I proposed to reimagine crowdsourced recommendations as the outcomes of a multi-winner election that is periodically repeated, and then the observed biases in recommendations can be attributed to the unfairness in the electoral system. I found that today’s trending topic(s)election algorithms are vulnerable to electing fringe trends with as low as 0.001% of the electorate support. To fairly aggregate the preferences of all users, I developed an innovative mechanism to attribute preferences of silent users, and applied a voting mechanism with the fairness properties we desire. The proposed approach is shown to provide maximum user satisfaction, reduce demographic bias and remove topics disliked by most but hyper-actively promoted by only a few users.

1.2.2 Fairness in Textual Summarization [CSCW’19]

With rapidly growing user-generated textual contents, summarization algorithms are used to provide users a quick overview of the information content. Traditionally, summarization algorithms have been evaluated based on how well they match human-written summaries (e.g. as measured by ROUGE scores). However, when the data comes from different socially salient user groups, e.g. men or women, Caucasians or African-Americans, or different political groups, I proposed to evaluate summarization algorithms on whether the generated summaries fairly represent these different user groups. Experiments over real-world datasets showed that existing extractive summarization algorithms often represent the groups very differently, and some groups get far less exposure than what they would have obtained in the original data. To reduce such adverse impacts, I proposed novel fairness-preserving summarization algorithms which produce high-quality summaries while ensuring fairness among various groups.

1.3 Ensuring Two-Sided Fairness for Producers and Customers

1.3.1 Fairness in Passenger-Driver Matchings in Ride-hailing Platforms [KDD’19]

Ride hailing platforms like Uber, Lyft or Ola have traditionally focused on the satisfaction of the passengers, or on boosting successful business transactions. However, recent studies provide a multitude of reasons to worry about the drivers, where the concerns range from bad working conditions and worker manipulation to discrimination against minorities. With more drivers financially depending on such platforms to secure a living, it is pertinent to ask what a fair distribution of income is, and what power and means the platform has in shaping these distributions. I analyzed the job assignments of a major taxi company and observed that there is significant inequality in the driver incomes. I proposed a novel framework to think about fairness in the matching mechanisms of ride hailing platforms, where the fairness notion

---

1Available at https://twitter-app.mpi-sws.org/who-makes-trends
requires that, spread over time, all drivers should receive benefits proportional to the amount of time they are active in the platform. By not requiring every match to be fair, but rather distributing fairness over time, I showed that it is possible to achieve better overall benefit for the drivers and the passengers. Finally, I proposed an ILP based optimization procedure for achieving two-sided fairness, taking the first step towards rethinking ride hailing platforms with an additional emphasis on the well-being of drivers.

1.3.2 Two-Sided Fairness in Personalized Recommendations [Under Review]

Traditionally, search and recommendation services in two-sided platforms have focused on maximizing customer satisfaction by tailoring the results according to the personalized preferences of individual customers, which may lead to unfair distribution of exposure to the producers. However, while being fair to the producers, the platforms should also fairly distribute the loss in utility among all the customers. By innovatively mapping the problem of personalized recommendation to the problem of fair allocation of indivisible goods, I proposed to provide fairness guarantees for both producers and customers. More formally, the proposed algorithm guarantees at least Maxi-Min Share (MMS) of exposure for majority of the producers, and Envy-Free upto One Good (EF1) fairness for all the customers.

1.3.3 Incremental Fairness during Platform Updates [AAAI’20]

To supposedly improve customer utility, different platforms frequently update their underlying algorithms. Such updates are often very rapid and immediate, leaving no room for the producers to adjust to the change. I investigated the fairness issues arising out of such frequent updates, and argued for incremental updates of the platform algorithms so that the producers have enough time to adjust (both logistically and mentally). However, naive incremental updates may become unfair to the customers. Focusing on recommendations deployed on two-sided platforms, I formulated an ILP based online optimization to deploy changes incrementally in \( n \) steps, ensuring smooth transition of the exposure of items while guaranteeing a minimum utility for every customer.

1.3.4 Fairness in Related Item Recommendations [Under Review]

Related item recommendations in two-sided platforms not only help customers compare items related (or similar) to a given item, but also play a major role in bringing traffic to individual items. With growing number of producers (of items) depending on such platforms to earn their livelihood, it is important to understand whether the recommendations are leading to fair exposure of different items, or whether undeserving items get more exposure than deserving items. To measure exposure of an item due to related item recommendations, I proposed a novel mechanism involving random walks over the network induced by such recommendations, and observed that the state-of-the-art related item recommendation algorithms do not render items their deserved exposure, leading to exposure bias. I further proposed three types of fairness-inducing interventions in the recommendation pipeline, which significantly reduce exposure bias, while preserving the underlying relatedness of recommendations to a high extent.

1.4 Platforms Taking Paternalistic Roles

1.4.1 Revealing Biases of Social Media News Outlets [ICWSM’18]

In social media, anyone can register as a news publisher without any upfront cost. Worryingly, consumers of such news are often not aware of the biases of these outlets, which is in sharp contrast to the traditional news channels (print or TV), where at least well-informed consumers are aware of the biases of different news publishers. To measure the biases of social media news outlets at large scale, I leveraged Facebook advertisement APIs to get detailed insights into the demographics of different news outlets’ audience on Facebook. I empirically showed that the ideological (liberal or conservative) leaning of a news source can be accurately estimated by the extent to which liberals or conservatives are over-/under-represented in the source’s audience. To help users know the biases of social media news publishers, I developed and deployed a web-based service Media Bias Monitor which quantifies and exposes the ideological biases of 20,000+ news outlets in Facebook. Additionally, it also shows the audience demographics along five other dimensions: gender, income level, racial affinity, national identity and age.

Available at https://twitter-app.mpi-sws.org/media-bias-monitor
1.4.2 Detecting and Preventing Proliferation of Clickbaits [ASONAM’16, CSCW’18]

Today, news outlets face cut-throat competition for reader attention and they need to make money from their clicks. Thus, to attract readers to visit a media site and click on an article, they employ various techniques, such as using catchy headlines to accompany article links, which can lure the readers to click on the links. Such headlines are known as Clickbaits, which exploit curiosity gap, where the headlines provide forward referencing cues to generate enough curiosity such that the readers feel compelled to click on the link to fill the knowledge gap. While these baits may trick them into clicking, in the long-run, clickbaits usually don’t live up to the expectation of the readers, and leave them disappointed. I developed the first automated classifier to detect whether a headline is clickbait or not. Then, with my collaborators, I built a chrome extension, Stop Clickbait, which warns the readers about the possibility of being baited by clickbaits, and offers the readers an option to block certain types of clickbaits they would not like to see during future encounters. The extension has been actively used by 2,000+ Chrome users till date.

1.4.3 Utilizing Truth Perception of Users to Fact Check News Stories [FAT’19b]

Social media sites have been severely criticized by policy makers and media watchdog groups for allowing fake news stories to spread unchecked on their platforms. In response, these sites are encouraging their users to report any news story that they perceive as fake. Stories reported as fake by a large number of users are prioritized for fact checking by (human) experts at fact checking organizations like Snopes and PolitiFact. Thus, social media sites today are relying on their users’ perceptions of the truthfulness of news stories to select stories to fact check. Along with my collaborators, I made the first attempt to understand how users perceive truth in news stories, and how biases in their perceptions might affect detecting and labeling fake news stories. We designed a novel test for users to rapidly assess how truthful or untruthful the claims in a news story are. We proposed specific goals for prioritizing stories to fact check and developed mechanisms to achieve those goals using such truth perceptions.

2 Future Research Plans

My long-term research goal is to build automated systems that conform to social, political and legal norms, such as, bringing more fairness, transparency and accountability in algorithmic systems. Although I have attempted to cover many such aspects in two-sided platforms, due to the sheer vastness of the problem space, there are still a lot of unexplored avenues which I want to investigate in the future.

Tackling Missing Preference from Users: One of the key challenges in utilizing user preferences is that their complete preferences are never known, and one needs to predict them in order to proceed. However, in reality, the prediction accuracy may differ for different individuals, thereby creating the possibility of unfair treatment due to incorrect predictions. It is important to note that the prediction accuracy often improves with more data, and thus, individuals who share less with the platform due to privacy reasons may be at the receiving end of the unfair treatment. Thus, a platform may have to deal with an unfortunate tradeoff between privacy and fairness, which I plan to explore in the future.

Decisions Involving Human-in-the-loop: In many platforms (e.g., online charities), the final decision rests with the human users, and human biases have a big role to play in the resulting (un)fairness of outcomes. Thus, it is important to understand how the algorithmic biases and human biases interact. Moreover, while it is possible to imagine that individual users have some true latent preference (possibly biased), multiple works in behavioral economics have shown how people often take decisions based on their intuitions (and not through rational thinking), which are often shaped by the choice architecture in which the decisions are made. I plan to explore how one might use the platform design to modify the choice architecture for achieving fairness. But, at the same time, I plan to deter the use of similar techniques for malicious purposes (e.g., preventing the use of dark patterns).

Fair Grouping of Users: For executing several tasks at scale (e.g., providing recommendations, targeting advertisements, or sharing physical resources), users are grouped into different clusters and then decisions are made at the group level. Traditional clustering algorithms try to group users such that the average within-group user similarity is maximized. However, there can be possible fairness issues in such approaches, where every user may not get similar utility derived from the clustering (e.g., some users may

[https://chrome.google.com/webstore/detail/stop-clickbait/iffolfpdcmehbghhamkgobjjdeejinma](https://chrome.google.com/webstore/detail/stop-clickbait/iffolfpdcmehbghhamkgobjjdeejinma)
not belong to their preferred clusters). I plan to explore fairness in clustering, and propose methods such that all the users are similarly satisfied from the clusters they are assigned to.

**Algorithmic Transparency:** Algorithmic transparency is increasingly becoming a common regulatory requirement, such as the European Union’s new General Data Protection Regulation providing users with a ‘right to explanation’ about algorithmic decisions made about them. However, providing transparency is not straightforward – giving information at a wrong level can confuse users, complicating their interaction with the system. Moreover, revealing too much details can make the system vulnerable to manipulations. I plan to understand what level of transparency provides users with actionable insights, yet doesn’t reveal too much about the underlying trade secrets.

As mentioned earlier, since my research agenda is inherently multidisciplinary, none of the above plans would materialize without collaborating with and learning from researchers from different disciplines and backgrounds. I have been fortunate to collaborate with many researchers from different institutes in the past, and I hope that I can continue and expand this tradition as I move forward. In the long run, my research interests are framed by the real-world practical and ethical challenges of online systems. To these challenges, I bring a computational lens to operationalize the abstract ethical concepts, and a technical approach to develop systems with the social desiderata.

**References**


