Fairness in Networks, a tutorial

Suresh Venkatasubramanian
University of Utah
Salt Lake City, USA
suresh@cs.utah.edu

Carlos Scheidegger
University of Arizona
Tucson, USA
cscheid@cs.arizona.edu

Sorelle Friedler
Haverford College
Haverford, USA
sorelle@cs.haverford.edu

Aaron Clauset
University of Colorado
Boulder, USA
aaron.clauset@colorado.edu

ABSTRACT

As ML systems have become more broadly adopted in high-stakes settings, our scrutiny of them should reflect their greater impact on real lives. The field of *fairness* in data mining and machine learning has blossomed in the last decade, but most of the attention has been directed at tabular and image data. In this tutorial, we will discuss recent advances in *network fairness*. Specifically, we focus on problems where one's position in a network holds predictive value (e.g., in a classification or regression setting) and favorable network position can lead to a cascading loop of positive outcomes, leading to increased inequality. We start by reviewing important sociological notions such as social capital, information access, and influence, as well as the now-standard definitions of fairness in ML settings. We will discuss the formalizations of these concepts in the network fairness setting, presenting recent work in the field, and future directions.

KEYWORDS

fairness, information flow, networks

ACM Reference Format:
Suresh Venkatasubramanian, Carlos Scheidegger, Sorelle Friedler, and Aaron Clauset. 2021. Fairness in Networks, a tutorial. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14–18, 2021, Virtual Event, Singapore.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3447548.3470821

1 TARGET AUDIENCE

The target audience is anyone interested in learning more about what fairness, discrimination, and inequality might mean on a social network. No specific expertise is expected beyond the familiarity that would be expected of a usual KDD attendee.

2 TUTORS

Venkatasubramanian, Friedler and Scheidegger have worked together since 2014 to develop a theoretical foundation for many problems within fairness in machine learning [11, 14]; Clauset is an expert in social network analysis and network science. Recently, their work has included examining the fairness implications of information access in social networks [12], as well as the inequity brought about in academic social networks [7, 17].

3 TUTORIAL OUTLINE.

(1) Social Science, Fairness, and Networks (45 mins)
(2) Interactive discussion and exploration (40 mins + 15 min break)
(3) Information Access and Flow Mechanisms (15 mins)
(4) Recent Research on Fairness in Networks (45 mins)
(5) Future Directions (20 mins)

3.1 Social Science, Fairness, and Networks (45 mins)

To introduce and motivate the overall tutorial from a social science perspective, we'll begin by discussing the [5] paper on The Networked Nature of Algorithmic Discrimination. This paper establishes the idea that fairness in networks is not just an allocation problem on a graph, but is about how social structures can create groups and patterns of inequality, mediated by access. A motivating example they describe, which we will also use for this tutorial, is a social network (such as LinkedIn) in the context of access to job information, where who you know can directly determine whether you receive a job [15]. Thus, they argue, your social network connections (or lack thereof) can be used to algorithmically discriminate against you in these online settings, e.g., in what job information you are shown.

This idea – the possibility of algorithmic discrimination based on social network position – is the motivating theme for this tutorial. The tutorial will begin by describing the social science rationale behind these concerns.

What are networks and how they are created. In order to understand how discrimination might manifest in social networks, we need to understand more about what these networks are and how they are created. Social networks [4], or networks where nodes represent individuals and edges represent social connections between those people, exist in both online and offline settings. Such networks universally exhibit a strong pattern of *homophily*, the tendency for people to be more likely to have ties (edges) in a social network if they share demographic characteristics and/or have common interests [19]. Thus, even without access to node attributes, homophily implies that networks themselves encode demographic

KDD '21, August 14–18, 2021, Virtual Event, Singapore

ACM ISBN 978-1-4503-8332-5/21/08.
https://doi.org/10.1145/3447548.3470821

correlations among individuals. Additionally, and especially in online settings where links between people can sometimes be the results of recommendation algorithms, these connections between similar individuals become reinforced through a process known as *contagion*.

Social capital in networks, the view from sociology. Granovetter's seminal work on The Strength of Weak Ties introduced the idea that position in a network may determine an individual's access to resources and that such access may be based on meso-structures in a network beyond the scope of an individual's direct ties [15]. In a job-focused social network, such weak ties between individuals who do not share other acquaintances, allow information about jobs to flow further through the network, letting individuals hear about jobs they might not otherwise know about. This tangible importance of an individual's position in a network has been further developed in the notion of social capital.

Social capital is the idea that an individual's position in a network is a form of wealth, privilege, and power [6, 8]. The concept of social capital appears in many different contexts, and as such has a number of nuanced meanings and implications. The idea of *contagion*, mentioned above, is when individuals with social capital lead other individuals to copy their behavior, either through active influence or through passive role modeling. *Network models of prominence* assume that social capital is an indication of quality or resources. *Closure* is the related property where highly connected networks spread information widely, while *brokerage* is the ability of highly connected bridge nodes to have the social capital that comes from controlling access to information.

The tutorial will discuss these and related ideas from sociology so that participants can understand how social networks and fairness are connected.

Social capital in networks, the view from CS. Increasingly, there has been empirical technical work demonstrating the importance of network position and social capital from within computer science and other technical fields. Continuing with our running example about the way that social position in a network increases access to jobs, a line of work in [7, 17, 23] has shown that the prestige of an academic institution reflects its network position, and this position then shapes the job prospects of its PhD graduates, its faculty's productivity, and the spread of scientific ideas. We will discuss this and other examples that demonstrate empirically how social and epistemic inequality manifests from academic network position.

What is fairness? Disparate impact, error rate balance, individual fairness. Separate from the work on social networks, there is now a substantial amount of work on algorithmic fairness, with a focus on fairness in classification problems. A basic review of some of the developed fairness definitions will be useful for tutorial participants to understand how fairness could be quantified. Reviewed definitions will include disparate impact [11], error rate balance [16, 24], and individual fairness [9].

3.2 Part 2: Interactive discussion and exploration (30 mins)

In breakout rooms of 5-10 people, we will have tutorial participants discuss the question of what fairness in networks should mean, and how it could be variously defined. We will structure this effort via guided discussion questions around given case study scenarios. One such example scenario follows. Three such case studies will be distributed to the different breakout rooms, and groups will take notes on their outcomes in a Google Doc shared with the organizers.

> Suppose that you are developing algorithms that can be used by a job-related social network such as LinkedIn to help recruiters determine which people (nodes) to target for job opportunities. Consider the following questions in your discussion groups:
> (1) Who would the recruiters like to be able to reach with job opportunities? How can they be identified?
> (2) What would it mean to allow recruiters to focus on equality of access in their outreach? How can they be helped to do this?
> (3) How could equality of access be formally defined in this case?
> (4) What interventions is it possible for recruiters or the social network itself to do in order to increase equality of access?
> (5) Are there other aspects of fairness in networks you think should be considered in this job-related network setting?

15 min break.

Follow-up discussion / report back (10 mins). The organizers will synthesize the outcomes from the discussion groups to report back some ideas about fairness in networks to the full group. The goal will be to prepare the audience to evaluate current research on fairness in networks and the extent to which it reflects their own thinking from the discussions.

3.3 Information Access and Flow Mechanisms (15 mins)

We will start the second part of the tutorial with a brief introduction to the models used to capture the flow of information over a (social) network. In keeping with the translational spirit of this tutorial, we will emphasize the different ways in which similar mechanisms for propagation are described in different communities.

We will start with the basic independent cascade model (where information is transmitted by a node to any given neighbor with a fixed probability, and after which the node does not transmit again), and follow this with the (linear) threshold model (where a node that receives sufficiently many signals from neighbors transmits information outward). We will relate these to the notions of simple and complex cascades in networks, as well as the different models for information flow used in epidemiology (pointing out the connection between the independent cascade (IC) and the susceptible-infected (SI) models, for example).

We will also briefly cover generalizations of these basic models that are pertinent to information flow, such as the models where nodes are more or less likely to transmit or be persuaded about information based on node-specific characteristics [2, 3].

3.4 Recent Research on Fairness in Networks (45 mins)

Armed with this mathematical framework, we will dive into a presentation of research on fairness in networks. We will start with a discussion of *access* as the quantity that needs to be equitably distributed. Access in a network has typically been measured as some utility function of the probability of receiving information via a flow mechanism such as above. We will discuss the different utility functions that have been proposed in the literature.

Equity. Following this, we will review the axiomatic frameworks used to decide what it means for access to be equitable. Different works have proposed different principled arguments for deriving different measures of equity, such as preventing the rich from getting richer, or ensuring that the least advantaged gain access (a Rawlsian maximin argument, for example). We will briefly discuss concepts from welfare theory that are used in some works. The papers that we will cover here include those by [22], [13] and [18].

These papers also represent two different frames for thinking about equity: a more "individually focused" version exemplified by [13] and one focused on equity for groups illustrated by [18, 22]. We will highlight the tensions between these perspectives.

Dynamics. Measures of equitable access can also be used to shed light on inequities in existing networks as well as be used to monitor how social phenomenon (through biases in attachment or recommendations) might increase bias in networks. We will introduce the audience to research that seeks to explore this in the context of gender discrimination [21] by suggesting mechanisms that lead to increased inequity, as well as in more general contexts with majority and minority groups [25]. This part of the tutorial will connect back with the work described above on the CS view of social capital: indeed, the work described here can be viewed as attempting to mathematically model empirically observed patterns of bias and thus provide a framework for interventions.

Interventions. If we recognize that patterns of inequity manifest in a social network and that there are ways to measure it, how might we rectify this with interventions? In the world of influence maximization, interventions correspond to *seeding* a network with carefully chosen nodes to improve access. We will review algorithms in the works above to intervene to optimize for fair access. The underlying algorithmic questions turn out to be much harder in general, not always admitting close-to-optimal solutions via submodular maximization except in special cases [1, 10, 13, 18, 22]. We will also review work [20] that seeks to design models for network formation that admit more efficient interventions that are both effective and fair under appropriate definitions of equity.

3.5 Future Directions (20 mins)

We will close with a group discussion of suggested future directions for exploration of this field. As prompts, we will encourage the participants to reflect on their interactive activity and how the questions they posed there are addressed (or not) by current research efforts.

4 PRIOR INSTANCES.

We are not aware of previous tutorials that consider the intersection of fairness and networks.

5 PEDAGOGY

We plan to alternate between lecture (with questions) and structured breakout room discussions for the tutorial. These structured breakout room discussion questions are provided within the outline above. Enough breakout rooms will be created, based on the number of people who attend the tutorial, so that 5-10 people are in each room. We will scribe the final group discussion in a document to add to what we hope will be a post-tutorial report for dissemination.

6 EQUIPMENT

We will not require any equipment beyond the ability to share our screens for presentation slide visibility with the participants and the ability to divide participants into breakout rooms. Attendees will simply need to join the video conference.

REFERENCES

[1] Junaid Ali, Mahmoudreza Babaei, Abhijnan Chakraborty, Baharan Mirzasoleiman, Krishna P Gummadi, and Adish Singla. 2019. On the fairness of time-critical influence maximization in social networks. *arXiv preprint arXiv:1905.06618* (2019).

[2] Sinan Aral and Paramveer S Dhillon. 2018. Social influence maximization under empirical influence models. *Nature human behaviour* 2, 6 (2018), 375–382.

[3] Sinan Aral and Dylan Walker. 2012. Identifying influential and susceptible members of social networks. *Science* 337, 6092 (2012), 337–341.

[4] danah boyd and Nicole B Ellison. 2007. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication* 13, 1 (2007), 210–230.

[5] danah boyd, Karen Levy, and Alice Marwick. 2014. The networked nature of algorithmic discrimination. *Data and Discrimination: Collected Essays. Open Technology Institute* (2014).

[6] Ronald S Burt. 2000. The network structure of social capital. *Research in organizational behavior* 22 (2000), 345–423.

[7] Aaron Clauset, Samuel Arbesman, and Daniel B Larremore. 2015. Systematic inequality and hierarchy in faculty hiring networks. *Science advances* 1, 1 (2015), e1400005.

[8] James S Coleman. 1988. Social capital in the creation of human capital. *American journal of sociology* 94 (1988), S95–S120.

[9] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference.* 214–226.

[10] Golnoosh Farnad, Behrouz Babaki, and Michel Gendreau. 2020. A Unifying Framework for Fairness-Aware Influence Maximization. In *Companion Proceedings of the Web Conference 2020* (Taipei, Taiwan) *(WWW '20)*. Association for Computing Machinery, New York, NY, USA, 714–722. https://doi.org/10.1145/3366424.3383555

[11] Michael Feldman, Sorelle A Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. 2015. Certifying and removing disparate impact. In *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining.* 259–268.

[12] Benjamin Fish, Ashkan Bashardoust, danah boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2019. Gaps in Information Access in Social Networks. In *The World Wide Web Conference.* 480–490.

[13] Benjamin Fish, Ashkan Bashardoust, Danah Boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2019. Gaps in Information Access in Social Networks. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 480–490. https://doi.org/10.1145/3308558.3313680

[14] Sorelle A Friedler, Carlos Scheidegger, Suresh Venkatasubramanian, Sonam Choudhary, Evan P Hamilton, and Derek Roth. 2019. A comparative study of fairness-enhancing interventions in machine learning. In *Proceedings of the conference on fairness, accountability, and transparency.* 329–338.

[15] Mark S Granovetter. 1973. The strength of weak ties. *American journal of sociology* 78, 6 (1973), 1360–1380.

[16] Moritz Hardt, Eric Price, and Nati Srebro. 2016. Equality of Opportunity in Supervised Learning. In *NeurIPS.*

[17] Allison C Morgan, Dimitrios J Economou, Samuel F Way, and Aaron Clauset. 2018. Prestige drives epistemic inequality in the diffusion of scientific ideas. *EPJ Data Science* 7, 1 (2018), 40.

[18] Aida Rahmattalabi, Shahin Jabbari, Himabindu Lakkaraju, Phebe Vayanos, Eric Rice, and Milind Tambe. 2020. Fair Influence Maximization: A Welfare Optimization Approach. *CoRR* abs/2006.07906 (2020). arXiv:2006.07906 https://arxiv.org/abs/2006.07906

[19] Cosma Rohilla Shalizi and Andrew C Thomas. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research* 40, 2 (2011), 211–239.

[20] Ana-Andreea Stoica, Jessy Xinyi Han, and Augustin Chaintreau. 2020. Seeding Network Influence in Biased Networks and the Benefits of Diversity. In *Proceedings of The Web Conference 2020* (Taipei, Taiwan) *(WWW '20)*. Association for Computing Machinery, New York, NY, USA, 2089–2098. https://doi.org/10.1145/3366423.3380275

[21] Ana-Andreea Stoica, Christopher Riederer, and Augustin Chaintreau. 2018. Algorithmic Glass Ceiling in Social Networks: The Effects of Social Recommendations on Network Diversity. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 923–932. https://doi.org/10.1145/3178876.3186140

[22] Alan Tsang, Bryan Wilder, Eric Rice, Milind Tambe, and Yair Zick. 2019. Group-Fairness in Influence Maximization. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 5997–6005. https://doi.org/10.24963/ijcai.2019/831

[23] Samuel F Way, Allison C Morgan, Daniel B Larremore, and Aaron Clauset. 2019. Productivity, prominence, and the effects of academic environment. *Proceedings of the National Academy of Sciences* 116, 22 (2019), 10729–10733.

[24] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, and Krishna P Gummadi. 2017. Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *Proceedings of the 26th international conference on world wide web*. 1171–1180.

[25] Yiguang Zhang, Jessy Xinyi Han, Ilica Mahajan, Priyanjana Bengani, and Augustin Chaintreau. 2021. Chasm in Hegemony: Explaining and Reproducing Disparities in Homophilous Networks. *arXiv preprint arXiv:2102.11925* (2021).