Fairness-Aware Recommendation of Information Curators

Ziwei Zhu, Jianling Wang, Yin Zhang, and James Caverlee

Texas A&M University

{zhuziwei,jwang713,zhan13679,caverlee}@tamu.edu

ABSTRACT

This paper highlights our ongoing efforts to create effective information curator recommendation models that can be personalized for individual users, while maintaining important fairness properties. Concretely, we introduce the problem of information curator recommendation, provide a high-level overview of a fairness-aware recommender, and introduce some preliminary experimental evidence over a real-world Twitter dataset. We conclude with some thoughts on future directions.

1 INTRODUCTION

Information curators serve as conduits to high-quality curated content, providing unique specialized expertise, trustworthiness in decision-making, and access to novel content. For example, in the aftermath of a natural disaster, critical citizen responders on Facebook can filter through the noise to curate high-quality, informative posts, while avoiding likely misinformation [25, 31]. During breaking news, knowledgeable locals can provide access to reputable reporting and contextual insights into the developing situation [7, 19]. And in a longer-term perspective, information curators can provide deep dives into health claims (e.g., by comparing and contrasting multiple analyses of the health benefits of new diets), products to buy (e.g., through rigorous evaluation and comparison), and insights into local governance issues (e.g., through curating analyses of proposed state amendments or local bond issues).

Successfully uncovering such information curators from the massive scale of the web and social media, reliably connecting users to the appropriate curators, and ensuring fairness-preserving properties of such curators is vitally important to prompt a trustworthy information diet, to improve the quality of user experience, and to support an informed populace. In practice, most users access information curators today via a centralized platform – like Facebook, Google, or NYTimes – meaning that the (hidden) algorithms connecting users to curators are essentially blackboxes. As a result, users have limited understanding of what factors impact what content is surfaced to them and whether or not any bias is shaping their information diet. In our context, such bias could lead to the suppression of curators by gender, race, religious beliefs, political stances, or other factors.

Yet there is a critical research gap in fairness-aware personalized recommendation of information curators at scale: First, many existing recommender systems focus on specific items – like movies, songs, or books as the basis of recommendation – rather than on information curators who organize a heterogeneous mix of high-quality curated items. And for those approaches that aim to uncover knowledgeable users in online systems – e.g., [1, 5, 11, 32, 39] – most typically do so without an emphasis on personalized recommendation. Indeed, there is a research gap in our understanding of both (i) identifying high-quality information curators who are gateways to curated content, and not just experts; and (ii) identifying
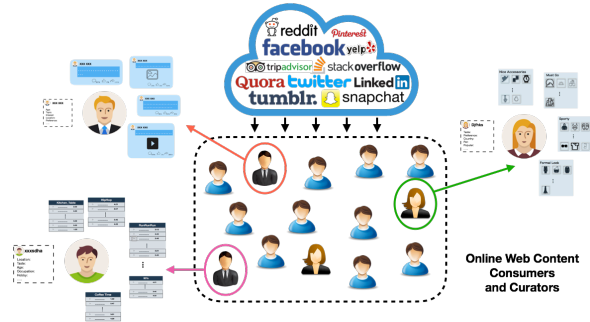

Figure 1: Information curators vary across web and social media platforms. We view curators as both high-level entities and by the curated items themselves.

personally-relevant curators, and not just highly-rated or popular ones. Second, there are typically complex and dynamic relationships between users, candidate curators, and topics of interest that manifest differently in heterogeneous environments. How to model such heterogeneity is critically important. And since user interests and information needs are inherently dynamic – that is, during a crisis, our preferences may be fundamentally different from our preferences in the long-term – there is a need for adaptable methods to handle these complex inter-relationships. Finally, as we have mentioned, most current access to information curators is mediated by centralized platforms (like search engines, social networks, and traditional news media), meaning that personal preferences may not align with the goals of these platforms, leading to potentially biased (or even limited) access to curators.

Toward tackling these challenges, we have begun a broad research effort to create new personalized recommendation frameworks for connecting users with high-quality information curators, even in extremely sparse, heterogeneous, and fairness-aware environments.[1] In the rest of this paper, we focus on our first steps at building a fairness-aware recommender in this context. We provide a high-level overview of the approach (more details can be found in a companion piece [41]), and we introduce some preliminary experimental evidence over a real-world Twitter dataset. Our hope is this can spark a discussion about the important fairness challenges facing information curator recommendation.

2 PROBLEM SETTING

We assume there exists a set of users $U = \{u_1, u_2, ..., u_N\}$, where N is the total number of users. From this set U, there are a number of information curators denoted as $C = \{c_1, c_2..., c_M\}$, where M is the total number of information curators. We then define the problem of *personalized information curator recommendation* as: Given a user u_i, identify the top-n personally relevant curators to u_i. In practice,

---

[1]http://faculty.cse.tamu.edu/caverlee/curators.html

these information curators vary greatly. As illustrated in Figure 1, these information curators can be viewed as high-level entities (e.g., a 30-something in San Francisco with interests in entrepreneurship, a journalist in NYC focused on local governance issues) as well as by the curated items themselves (e.g., a series of blog posts analyzing a recent election, a list of product reviews). This variety and the resultant heterogeneity – in terms of content types, social relations, motivations of curators, etc. – place great demands on effective personalization. Further, existing expressed preferences for these curators in the forms of likes, following relationships, and other interactions are often sparse. Hence, a key challenge for personalized curator recommendation is tackling sparsity while carefully modeling curators in complex, noisy, and heterogeneous environments.

3 FAIRNESS-AWARE LEARNING FOR INFORMATION CURATORS

Since user preferences for curators may be impacted by many contextual factors, we propose to directly incorporate the multiple and varied relationships among users, curators, topics, and other factors directly into a *tensor-based approach*. Such a three-dimensional tensor can be naturally extended to capture higher-order factors (e.g., by adding extra dimensions for location, reputation, and so on). Of course, we can also explore traditional matrix-based approaches in comparison with these tensor-based ones.

As a first step, we can tackle the personalized curator recommendation problem with a basic recommendation framework using tensor factorization. Let $U^{(1)} \in \mathbb{R}^{N \times R}$, $U^{(2)} \in \mathbb{R}^{M \times R}$ and $U^{(3)} \in \mathbb{R}^{K \times R}$ be latent factor matrices for users, curators, and topics, respectively. The basic tensor-based curator recommendation model can be defined as:

$$\begin{aligned} \underset{U^{(n)}, \mathcal{X}}{\text{minimize}} \quad & \frac{1}{2}\|\mathcal{X} - [\![U^{(1)}, U^{(2)}, U^{(3)}]\!]\|_F^2 + \frac{\lambda}{2}\sum_{n=1}^{3}\|U^{(n)}\|_F^2, \\ \text{subject to} \quad & \Psi * \mathcal{X} = \mathcal{T} \end{aligned}$$

This basic model estimates $\hat{\mathcal{X}}$ that approximates the original rating tensor (unknown) $\mathcal{X}$ via learning optimal latent factor matrices $\{U^{(n)}, n = 1, 2, 3\}$. For each user and topic of interest, this model can recommend a ranked list of personalized curators.

3.1 Isolating Sensitive Information

However, such basic recommenders may inherit bias from the training data used to optimize them and from mis-alignment between platform goals and personal preferences. Hence, we aim to build new fairness-aware algorithms that can empower users by enhancing diversity of topics, curators, and viewpoints. As illustrated in Figure 2, we aim to augment our existing methods by isolating sensitive features through a new sensitive latent factor matrix, creating a sensitive information regularizer that extracts sensitive information which can taint other latent factors, and producing fairness-enhanced recommendation by the new latent factor matrices without sensitive information. For a fuller treatment of this approach, please refer [41].
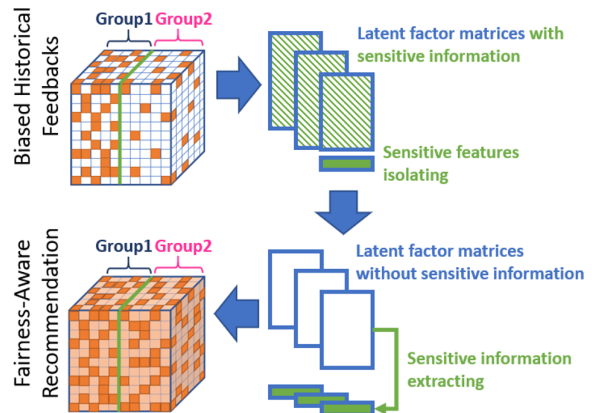

Figure 2: Overview: sensitive features are isolated (top right), then sensitive information is extracted (bottom right), resulting in fairness-aware recommendation.

For clarity, assume we have a tensor-based recommender. Such a tensor-based approach has no notion of fairness. Here, we assume that there exists a *sensitive attribute* for one mode of the tensor, and this mode is a *sensitive mode*. For example, the sensitive attribute could correspond to gender, age, ethnicity, location, or other domain-specific attributes of users, curators or topics in the recommenders. The feature vectors of the sensitive attributes are called the *sensitive features*. Further, we call all the information related to the sensitive attributes as *sensitive information*, and note that attributes other than the sensitive attributes can also contain sensitive information [17, 38]. While there are emerging debates about what constitutes algorithmic fairness [6], we adopt the commonly used notion of *statistical parity*. Statistical parity encourages a recommender to ensure similar probability distributions for both the dominant group and the protected group as defined by the sensitive attributes. Formally, we denote the sensitive attribute as a random variable S, and the preference rating in the recommender system as a random variable R. Then we can formulate fairness as $\mathbf{P}[R] = \mathbf{P}[R|S]$, i.e. the preference rating is independent of the sensitive attribute. This statistical parity means that the recommendation result should be independent to the sensitive attributes. For example, a job recommender should recommend similar jobs to men and women with similar profiles. Note that some recent works [13, 34, 35] have argued that statistical parity may be overly strict, resulting in poor utility to end users. Our work here aims to achieve comparable utility to non-fair approaches, while providing stronger fairness.

Given this (admittedly limited) notion of fairness, the intuition of the proposed framework is that the latent factor matrices of the tensor completion model contain latent information related to the sensitive attributes, which introduces the unfairness. Therefore, by *isolating* and then *extracting* the sensitive information from the latent factor matrices, we may be able to improve the fairness of the recommender itself. We propose to first isolate the impact of the sensitive attribute by plugging the sensitive features into the latent factor matrix. For instance, in our user-curator-topic example where we want to enhance the recommendation fairness for curators of both genders, we can create one vector $\mathbf{s}_0$ with 1 representing male

curator and 0 representing female curator, and another vector s_1 with 1 indicating female and 0 indicating male. s_0 and s_1 together form a matrix, denoted as *Sensitive Features* S. We put S to the last two columns of the latent factor matrix of sensitive mode (the mode for curators). In this way, we construct a new *sensitive latent factor matrix*, and we call the last two columns as *sensitive dimensions* and the others *non-sensitive dimensions*. By isolating the sensitive features, we provide a first step toward improving the fairness of the recommender. But there may still be sensitive information that resides in non-sensitive dimensions. To extract this remaining sensitive information, we propose an additional constraint that the non-sensitive dimensions should be orthogonal to the sensitive dimensions in the sensitive latent factor matrix. After the above two steps, we can get the new latent factor matrices, whose sensitive dimensions hold features exclusively related to the sensitive attributes. And their non-sensitive dimensions are decoupled from the sensitive attributes. Thus, we can derive the fairness-enhanced recommendation by reconstructing the rating tensor only by the non-sensitive dimensions.

4 EXPERIMENTS

To test such an approach, we adopt a Twitter dataset introduced in [10] that has 589 users, 252 curators, and 10 topics (e.g., news, sports). There are 16, 867 links from users to curators across these topics capturing that a user is interested in a particular curator. The sparsity of this dataset is 1.136%. We consider race as a sensitive attribute and aim to divide curators into two groups: whites and non-whites. We apply the Face++ (https://www.faceplusplus.com/) API to the images of each curator in the dataset to derive ethnicity. In total, we have 126 whites and 126 non-whites, with 11,612 positive ratings for white curators but only 5,255 for non-whites. Since this implicit feedback scenario has no negative observations, we randomly pick unobserved data samples to be negative feedback with probability of 0.113% (one tenth of the sparsity). We randomly split the dataset into 70% training and 30% testing.

4.1 Metrics

We consider metrics to capture recommendation quality, recommendation fairness, and the impact of eliminating sensitive information.

Recommendation Quality. To measure *recommendation quality*, we adopt **Precision@k** and **Recall@k**, which are defined as:

$$Precision@k = \frac{1}{|\mathbb{U}|} \sum_{u \in \mathbb{U}} \frac{|\mathbb{O}_u^k \cap \mathbb{O}_u^+|}{k},$$

$$Recall@k = \frac{1}{|\mathbb{U}|} \sum_{u \in \mathbb{U}} \frac{|\mathbb{O}_u^k \cap \mathbb{O}_u^+|}{\mathbb{O}_u^+},$$

where $\mathbb{O}_u^+$ is the set of items user u gives positive feedback to in test set and $\mathbb{O}_u^k$ is the predicted top-k recommended items. We also consider **F1@k** score, which can be calculated by $F1@k = 2 \cdot (Precision@k \times Recall@k)/(Precision@k + Recall@k)$. We set $k = 15$ in our experiments.

Recommendation Fairness. To measure *recommendation fairness*, we consider two complementary metrics. The first one is the absolute difference between the mean ratings of different groups (**MAD**):

$$MAD = |\frac{\sum R^{(0)}}{|R^{(0)}|} - \frac{\sum R^{(1)}}{|R^{(1)}|}|,$$

where $R^{(0)}$ and $R^{(1)}$ are the predicted ratings for the two groups and $|R^{(i)}|$ is the total number of ratings for group i. Larger values indicate greater differences between the groups, which we interpret as unfairness.

The second fairness measure is the Kolmogorov-Smirnov statistic (**KS**), which is a nonparametric test for the equality of two distributions. The KS statistic is defined as the area difference between two empirical cumulative distributions of the predicted ratings for groups:

$$KS = |\sum_{i=1}^{T} l \times \frac{\mathcal{G}(R^{(0)}, i)}{|R^{(0)}|} - \sum_{i=1}^{T} l \times \frac{\mathcal{G}(R^{(1)}, i)}{|R^{(1)}|}|,$$

where T is the number of intervals for the empirical cumulative distribution, l is the size of each interval, $\mathcal{G}(R^{(0)}, i)$ counts how many ratings are inside the ith interval for group 0. In our experiments, we set $T = 50$. Lower values of KS indicate the distributions are more alike, which we interpret as being more fair.

Note that we measure the fairness in terms of MAD and KS metrics across groups rather than within individuals, since absolute fairness for every individual may be overly strict and in opposition to personalization needs of real-world recommenders.

4.2 Baselines

We adopt a modified Gradient Descent algorithm to optimize the proposed model and name the framework **FT** – in comparison with two tensor-based alternatives:

Ordinary Tensor Completion (OTC): The first is the conventional CP-based tensor completion method using ALS optimization algorithm. This baseline incorporates no notion of fairness, so it will provide a good sense of the state-of-the-art recommendation quality we can achieve.

Regularization-based Tensor Completion (RTC): The second one is an extension from the fairness-enhanced matrix completion with regularization method introduced in [15, 16, 34], which adds a bias penalization term to the matrix factorization objective function. For tensor-based recommenders, we can also add a regularization term to enforce statistical parity. We use Gradient Descent to solve this optimization problem.

We also consider purely matrix-based approaches, which compute user preferences on curators for each topic independently. We consider matrix versions of our tensor based methods (named **FM**) corresponding to FT versus matrix baselines of *Ordinary Matrix Completion (OMC)* corresponding to OTC and *Regularization-based Matrix Completion (RMC)* corresponding to RTC.

4.3 Experimental Results

We set 20 as the latent dimension for all the methods and fine tune all other parameters. The experiments are run three times and the averaged results are reported.
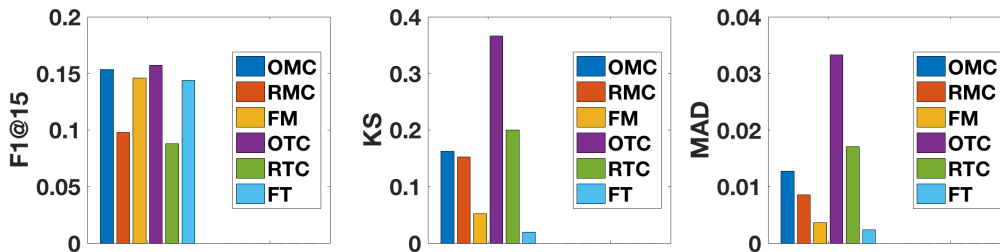
Figure 3: F1@15, KS, and MAD scores for all six models. The fairness-aware approaches (FM for matrix, FT for tensor) provide comparable quality (F1@15) relative to non-fair methods while providing much lower KS and MAD scores (which are proxies for statistical parity).

Methods	R@15	P@15	KS	MAD
OMC	0.4128	0.0942	0.1625	0.0127
OTC	**0.4384**	**0.0958**	0.3662	0.0333
RMC	0.1609	0.0702	0.1521	0.0086
RTC	0.3003	0.0515	0.2003	0.0171
FM	0.4045	0.0891	0.0523	0.0037
FT	0.4180	0.0870	**0.0195**	**0.0024**

Table 1: Comparison for recommending Twitter curators.

First, let's focus on the differences between matrix and tensor approaches as shown in Table 1. We observe that the tensor-based approaches mostly provide better recommendation quality (Precision@k and Recall@k) in comparison with the matrix-based approaches. Since the Twitter curation dataset is naturally multi-aspect, the tensor approaches better model the multi-way relationships among users, curators, and topics. We see that the fairness quality (KS and MAD) of matrix-based methods are better than tensor-based ones for the baselines methods (OMC vs OTC, and RMC vs RTC), but the fairness improves for our proposed methods when we move from matrix to tensor (FM vs FT).

Second, let's consider the empirical results across approaches as present in Figure 3. We see that: (i) the proposed methods are slightly worse than OTC from the perspective of recommendation quality, but keep the difference small, and FM methods also have comparable recommendation performance with OMC; and (ii) FT provides the best fairness enhancement results, and FM also alleviates the unfairness a lot compared with other matrix-based methods. RTC and RMC improve the fairness as well, but their effects are not as good as the proposed methods.

These results suggest the potential of such a framework toward building fairness-aware information curator recommendation.

5 RELATED WORK

Many previous works have focused on finding experts in many domains (e.g., enterprise corporate, email networks [1, 3, 12, 23, 26, 30, 39]), with a recent emphasis on social media [11, 27, 32]. Building on these efforts, we have prototyped expert recommenders on Twitter [5, 10, 24] that provide a firm foundation for the proposed research tasks here. However, there is a research gap in identifying personalized, high-quality information curators who are gateways to curated content, and not just experts or popular users. In a related direction, many works have focused on recommendation, in which user preferences may be projected into a lower dimensional embedding space [20, 24, 33, 36]. MF and BPR-based approaches

have shown good success in implicit feedback scenarios [4, 29, 40] as in our case, and in social media scenarios [21, 40]. In recent years, tensor factorization models are becoming popular and successfully applied in recommendation, including [2, 14, 18]. In contrast, this project explores personalized curator recommendation by integrating multiple, heterogeneous contexts into a unified model.

Fairness, accountability, and transparency are critical issues. Friedman [9] defined that a computer system is biased *"if it systematically and unfairly discriminate[s] against certain individuals or groups of individuals in favor of others."* For information curator recommenders, algorithmic bias could lead to the suppression of curators by gender, race, religious beliefs, political stances, or other factors. Such algorithmic bias encodes unwanted biases in the daily experiences of millions of users, and potentially violates discrimination laws. Indeed, researchers across communities have begun actively investigating evidence of bias in existing systems and in methods towards revealing and/or overcoming this bias [8, 22, 28, 37, 38]. In the recommender space, recent work has led to notions of group fairness and individual fairness in recommendation [38], and fairness-criteria for top-k ranking recommendation [37]. Kamishima summarized that recommender systems should be in adherence to laws and regulations, should be fair to all the content providers, and should exclude the influence of unwanted information [15].

6 CONCLUSION AND FUTURE WORK

This paper highlights our initial efforts at building information curator recommenders that incorporate a simple notion of fairness. In our continuing work, we are interested in generalizing our framework to consider alternative notions of fairness beyond statistical parity. By extending our framework in this direction, we can provide a more customizable approach for defining and deploying fairness-aware methods. We are also interested in exploring how to incorporate real-valued features into the framework for recommenders with explicit ratings, and in running user studies on the perceived change of fairness for our methods.

REFERENCES

[1] Krisztian Balog, Leif Azzopardi, and Maarten De Rijke. 2006. Formal models for expert finding in enterprise corpora. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 43–50.

[2] Preeti Bhargava, Thomas Phan, Jiayu Zhou, and Juhan Lee. 2015. Who, what, when, and where: Multi-dimensional collaborative recommendations using tensor factorization on sparse user-generated data. In *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 130–140.

[3] Christopher S Campbell, Paul P Maglio, Alex Cozzi, and Byron Dom. 2003. Expertise identification using email communications. In *Proceedings of the twelfth international conference on Information and knowledge management*. ACM, 528–531. https://doi.org/10.1145/956863.956965

[4] Kailong Chen, Tianqi Chen, Guoqing Zheng, Ou Jin, Enpeng Yao, and Yong Yu. 2012. Collaborative personalized tweet recommendation. In *SIGIR*.

[5] Zhiyuan Cheng, James Caverlee, Himanshu Barthwal, and Vandana Bachani. 2014. Who is the barbecue king of texas?: a geo-spatial approach to finding local experts on twitter. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. ACM, 335–344.

[6] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. Algorithmic decision making and the cost of fairness. *arXiv preprint arXiv:1701.08230* (2017).

[7] Dharma Dailey and Kate Starbird. 2014. Journalists as crowdsourcerers: Responding to crisis by reporting with a crowd. *Computer Supported Cooperative Work (CSCW)* 23, 4-6 (2014), 445–481.

[8] Ayman Farahat and Michael C Bailey. 2012. How effective is targeted advertising?. In *Proceedings of the 21st international conference on World Wide Web*. ACM, 111–120.

[9] Batya Friedman and Helen Nissenbaum. 1996. Bias in computer systems. *ACM Transactions on Information Systems (TOIS)* 14, 3 (1996), 330–347.

[10] Hancheng Ge, James Caverlee, and Haokai Lu. 2016. TAPER: A Contextual Tensor-Based Approach for Personalized Expert Recommendation.. In *RecSys*. 261–268.

[11] Saptarshi Ghosh, Naveen Sharma, Fabricio Benevenuto, Niloy Ganguly, and Krishna Gummadi. 2012. Cognos: crowdsourcing search for topic experts in microblogs. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*. ACM, 575–590.

[12] Ido Guy, Uri Avraham, David Carmel, Sigalit Ur, Michal Jacovi, and Inbal Ronen. 2013. Mining expertise and interests from social media. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 515–526.

[13] Moritz Hardt, Eric Price, Nati Srebro, et al. 2016. Equality of opportunity in supervised learning. In *Advances in neural information processing systems*. 3315–3323.

[14] Balázs Hidasi and Domonkos Tikk. 2012. Fast ALS-based tensor factorization for context-aware recommendation from implicit feedback. *Machine Learning and Knowledge Discovery in Databases* (2012), 67–82.

[15] Toshihiro Kamishima and Shotaro Akaho. 2017. Considerations on Recommendation Independence for a Find-Good-Items Task. In *FATREC Workshop on Responsible Recommendation Proceedings*.

[16] Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. 2013. Efficiency Improvement of Neutrality-Enhanced Recommendation.. In *Decisions@ RecSys*. 1–8.

[17] Toshihiro Kamishima, Shotaro Akaho, and Jun Sakuma. 2011. Fairness-aware learning through regularization approach. In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*. IEEE, 643–650.

[18] Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. 2010. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 79–86.

[19] Andrea Kavanaugh, Steven Sheetz, Rodrigo Sandoval Almazan, John Tedesco, and Edward Fox. 2016. Media use during conflicts: Information seeking and political efficacy during the 2012 Mexican elections. *Government Information Quarterly* 33 (02 2016).

[20] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009).

[21] Artus Krohn-Grimberghe, Lucas Drumond, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2012. Multi-relational Matrix Factorization Using Bayesian Personalized Ranking for Social Network Data. In *WSDM*.

[22] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P Gummadi, and Karrie Karahalios. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. *arXiv preprint arXiv:1704.01347* (2017).

[23] Xiaoyong Liu, W Bruce Croft, and Matthew Koll. 2005. Finding experts in community-based question-answering services. In *Proceedings of the 14th ACM international conference on Information and knowledge management*. ACM, 315–316. https://doi.org/10.1145/1099554.1099644

[24] Haokai Lu and James Caverlee. 2015. Exploiting geo-spatial preference for personalized expert recommendation. In *RecSys*.

[25] David Maxwell, Stefan Raue, Leif Azzopardi, Chris Johnson, and Sarah Oates. 2012. Crisees: Real-time monitoring of social media streams to support crisis management. *Advances in Information Retrieval* (2012), 573–575.

[26] David W McDonald and Mark S Ackerman. 2000. Expertise recommender: a flexible recommendation system and architecture. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM, 231–240.

[27] Aditya Pal and Scott Counts. 2011. Identifying topical authorities in microblogs. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 45–54.

[28] Dino Pedreshi, Salvatore Ruggieri, and Franco Turini. 2008. Discrimination-aware data mining. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 560–568.

[29] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*. AUAI Press, 452–461.

[30] Pavel Serdyukov, Henning Rode, and Djoerd Hiemstra. 2008. Modeling multi-step relevance propagation for expert finding. In *Proceedings of the 17th ACM conference on Information and knowledge management*. ACM, 1133–1142.

[31] Kate Starbird and Leysia Palen. 2011. Voluntweeters: Self-organizing by digital volunteers in times of crisis. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1071–1080.

[32] Jianshu Weng, Ee-Peng Lim, Jing Jiang, and Qi He. 2010. Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining*. ACM, 261–270.

[33] Baoguo Yang and Suresh Manandhar. 2014. Tag-based expert recommendation in community question answering. In *Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on*. IEEE, 960–963.

[34] Sirui Yao and Bert Huang. 2017. Beyond Parity: Fairness Objectives for Collaborative Filtering. *arXiv preprint arXiv:1705.08804* (2017).

[35] Sirui Yao and Bert Huang. 2017. New Fairness Metrics for Recommendation that Embrace Differences. *arXiv preprint arXiv:1706.09838* (2017).

[36] Hsiang-Fu Yu, Cho-Jui Hsieh, Si Si, and Inderjit Dhillon. 2012. Scalable coordinate descent approaches to parallel matrix factorization for recommender systems. In *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. IEEE, 765–774.

[37] Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. 2017. FA* IR: A Fair Top-k Ranking Algorithm. *arXiv preprint arXiv:1706.06368* (2017).

[38] Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. 2013. Learning fair representations. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*. 325–333.

[39] Jun Zhang, Mark S Ackerman, and Lada Adamic. 2007. Expertise networks in online communities: structure and algorithms. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 221–230.

[40] Wei Zhang, Jianyong Wang, and Wei Feng. 2013. Combining latent factor model with location features for event-based group recommendation. In *SIGKDD*.

[41] Ziwei Zhu, Xia Hu, and James Caverlee. 2018. Fairness-Aware Tensor-Based Recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM)*.