# Analyzing fairness of recommendations in e-participatory budgeting

Marina Alonso-Cortés
*Universidad Autónoma de Madrid*
Madrid, Spain
marina.alonso-cortes@estudiante.uam.es

Iván Cantador
*Universidad Autónoma de Madrid*
Madrid, Spain
ivan.cantador@uam.es

Alejandro Bellogín
*Universidad Autónoma de Madrid*
Madrid, Spain
alejandro.bellogin@uam.es

***Abstract*—In the work presented herein, we measured and analyzed recommendation fairness metrics on a public e-participatory budgeting dataset, focusing on minority, vulnerable and NIMBY (Not In My Back Yard) groups of citizens. Hybrid recommendation algorithms exploiting user geolocation and collaborative filtering information were shown to be good candidates to address fairness concerns for the above underrepresented citizen collectives.**

***Index Terms*—Recommender systems, citizen participation, fairness, bias, minority groups**

This paper is a summary of the following publication:
M. Alonso-Cortés, I. Cantador, and A. Bellogín. "Recommendation fairness in e-participation: Listening to minority, vulnerable and NIMBY citizens," in *Proceedings of the 46th European Conference on Information Retrieval, Part IV*. Springer, 2024, pp. 420–436.
https://doi.org/10.1007/978-3-031-56066-8_31

## I. Introduction

Citizen participation entails active involvement of individuals in government decision-making processes [1]. It is progressively being conducted on the internet, through the so-called electronic participation (or *e-participation*) platforms [2]. By harnessing the power of ICT, e-participation can overcome physical barriers to participation and amplify the voices of individuals who may be marginalized or excluded from the government decision-making [3]. By contrast, e-participation has difficulties and entails certain limitations, such as unequal (biased) citizen representation, and non-in-depth deliberation and polarization [4]. These problems are originated or intensified by the overload of information and the lack of personalization in the used platforms [5], among other causes.

As a particular potential solution to these problems, recommender systems have been proposed to suggest relevant citizen proposals based on personal preferences [5], [6]. However, they have not been studied considering underlying societal issues, such as biases, fairness, privacy and transparency, as done for other domains [7], [8].

## II. Contributions

We evaluated diverse recommendation algorithms according to both accuracy and fairness metrics, on a public dataset[1] of the *Decide Madrid*[2] e-participatory budgeting platform, and for 32 general, underrepresented citizen collectives, such as people with disabilities, unemployed, immigrants and refugees, poor and people in social exclusion, the LGTBIQ+ collective, and citizens affected by NIMBY issues, e.g., related to urban planning, street garbage and noise, and neighborhood gambling.

More specifically, in the context of making personalized recommendations of citizen proposals publicly available and discussed online in Decide Madrid, we carefully identified a number of minority, vulnerable and NIMBY groups of city residents, and assigned to them the proposals that deal with their concerns, needs and problems. Then, we proposed a novel social fairness conceptualization, and measured and analyzed associated fairness metrics based on an established distribution-based evaluation metric [9] with which, for the above groups, we studied existing biases of recommendations generated by distinct algorithms.

## III. Results

We experimented with five families of recommenders, namely random (*rand*), popularity-based (*pop*), collaborative filtering (item- and user-based heuristics, *ib* and *ub*, and matrix factorization and BPR models, *mf* and *bpr*), and content-based (*cb*) and hybrid (*cbub* and *cbub*) recommenders exploiting category, topic and location information.

We analyzed the fairness achieved by each recommender according to the *Generalized Cross Entropy* (GCE) metric [9]:

$$GCE_{\beta}(A, R; p_f) = \frac{1}{\beta(1-\beta)}\left[\sum_{a \in A} p_f^{\beta}(a) \cdot p_R^{(1-\beta)}(a) - 1\right] \quad (1)$$

where $A$ is the attribute space upon which probability distributions are defined, $R$ is the recommendation algorithm whose fairness is assessed, and $p_f$ is the ideal or target fairness distribution, against which GCE will compare the estimated $p_R$ distribution from $R$ –in particular, if $p_R = p_f$ then $GCE = 0$, i.e., $R$ is considered a perfectly fair model.

Table I shows the GCE values achieved by the evaluated recommenders under five different perspectives, depending on the ideal fairness distribution considered:

[1]The dataset contained 50.2K unary ratings (comments) provided by 12.8K users (city residents) to 19.0K items (citizen proposals).

[2]Decide Madrid e-participatory budgeting platform, https://decide.madrid.es

TABLE I

AVERAGE GCE VALUES (THE CLOSER TO 0, THE BETTER) ACHIEVED BY THE RECOMMENDATION ALGORITHMS. EACH COLUMN REFERS TO A FAIRNESS PERSPECTIVE. ACCURACY VALUES (NDCG) INCLUDED FOR COMPARISON. BEST VALUES PER COLUMN IN DARKER COLORS.

| | nDCG | $p_u$ | $p_t$ | $p_m$ | $p_n$ | $p_{m+n}$ |
|---|---|---|---|---|---|---|
| *rand* | 0.004 | -1.264 | -0.007 | -3.806 | -5.151 | -2.604 |
| $pop_u$ | 0.096 | -1.932 | -0.031 | -10.243 | -2.605 | -3.817 |
| $pop_c$ | 0.078 | -1.005 | -0.011 | -4.716 | -2.702 | -2.125 |
| *ib* | 0.020 | -1.034 | -0.004 | -3.974 | -3.626 | -2.181 |
| *ub* | 0.030 | -0.954 | -0.005 | -3.888 | -3.237 | -2.034 |
| *mf* | 0.068 | -2.071 | -0.018 | -7.292 | -6.393 | -4.077 |
| *bpr* | 0.035 | -1.270 | -0.028 | -4.482 | -4.496 | -2.611 |
| $cb_{cat}$ | 0.008 | -1.600 | -0.007 | -6.388 | -4.543 | -3.219 |
| $cb_{top}$ | 0.013 | -1.757 | -0.010 | -6.399 | -5.450 | -3.505 |
| $cb_{loc}$ | 0.010 | -1.157 | -0.006 | -3.549 | -4.774 | -2.407 |
| $cbib_{cat}$ | 0.008 | -1.633 | -0.008 | -6.617 | -4.509 | -3.280 |
| $cbib_{top}$ | 0.020 | -1.864 | -0.012 | -6.605 | -5.873 | -3.701 |
| $cbib_{loc}$ | 0.017 | -1.225 | -0.020 | -2.989 | -5.726 | -2.529 |
| $cbub_{cat}$ | 0.028 | -1.352 | -0.007 | -6.228 | -3.242 | -2.764 |
| $cbub_{top}$ | 0.037 | -1.319 | -0.007 | -6.035 | -3.237 | -2.702 |
| $cbub_{loc}$ | 0.027 | -1.387 | -0.018 | -6.780 | -2.882 | -2.824 |

- $p_u$ encodes *fairness as equality*; where *ub* and $pop_c$ achieved the best (highest, closer to 0) results, which means that their recommendations are uniform for the considered fairness attribute values.
- $p_t$ simulates a *test perspective*, that is, how close the generated recommendations are to the item distribution observed in test; where *ib* and *ub* heuristic collaborative filtering algorithms, those based on content (in particular, $cb_{loc}$), and the hybrid algorithms achieve the best results.
- $p_m$, $p_n$, and $p_{m+n}$ encode perspectives *biased towards the discriminated groups* (minority, NIMBY, or both, respectively); where the algorithms exploiting location information –in particular, $cbib_{loc}$ for $p_m$ and $cbub_{loc}$ for $p_n$– and the heuristic collaborative filtering algorithms stand out. It is worth noting the negative bias that the $pop_u$ algorithm has on the minority groups, moving far away from the idea of fairness defined by $p_m$; this is probably linked to the fact that the minority proposals are not popular on the platform.

## IV. CONCLUSIONS

Our empirical results confirmed suspicions regarding a strong popularity bias on e-participation platforms that affects the recommendation algorithms. They also showed that those recommendations generated by exploiting the users' geolocation information in a collaborative filtering fashion are less affected by such bias for underrepresented citizen collectives.

We should conduct more exhaustive experiments to further corroborate and generalize our findings and conclusions. Specifically, we plan to thoroughly test further hyperparameter settings and recommendation algorithms, and use additional datasets, such as those published in [6] from the e-participatory budgets of New York City, Miami, and Cambridge. Moreover, we should investigate ad hoc, fairness-aware recommendation algorithms and mitigation techniques. We believe diversification [10] could be an effective approach for such purpose.


## ACKNOWLEDGMENT

This work was supported by grants PID2019-108965GB-I00 of the Spanish Ministry of Science and Innovation, and PID2022-139131NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe."


## REFERENCES


[1] A. Macintosh, "E-democracy and e-participation research in Europe," in *Digital Government: e-Government Research, Case Studies, and Implementation*. Springer, 2008, pp. 85–102.

[2] ——, "Characterizing e-participation in policy-making," in *Proceedings of the 37th Hawaii International Conference on System Sciences*. IEEE, 2004, pp. 10–pp.

[3] Y. Zheng and H. L. Schachter, "Explaining citizens' e-participation use: the role of perceived advantages," *Public Organization Review*, vol. 17, pp. 409–428, 2017.

[4] I. Cantador, M. E. Cortés-Cediel, and M. Fernández, "Exploiting open data to analyze discussion and controversy in online citizen participation," *Information Processing & Management*, vol. 57, no. 5, p. 102301, 2020.

[5] I. Cantador, A. Bellogín, M. E. Cortés-Cediel, and O. Gil, "Personalized recommendations in e-participation: offline experiments for the 'Decide Madrid' platform," in *Proceedings of the International Workshop on Recommender Systems for Citizens*. ACM, 2017, pp. 1–6.

[6] I. Cantador, M. E. Cortés-Cediel, M. Fernández, and H. Alani, "What's going on in my city? Recommender systems and electronic participatory budgeting," in *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM, 2018, pp. 219–223.

[7] M. D. Ekstrand, M. Tian, I. M. Azpiazu, J. D. Ekstrand, O. Anuyah, D. McNeill, and M. S. Pera, "All the cool kids, how do they fit in?: popularity and demographic biases in recommender evaluation and effectiveness," in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, vol. 81. PMLR, 2018, pp. 172–186.

[8] Y. Wang, W. Ma, M. Zhang, Y. Liu, and S. Ma, "A survey on the fairness of recommender systems," *ACM Transactions on Information Systems*, vol. 41, no. 3, pp. 1–43, 2023.

[9] Y. Deldjoo, V. W. Anelli, H. Zamani, A. Bellogín, and T. Di Noia, "A flexible framework for evaluating user and item fairness in recommender systems," *User Modeling and User-Adapted Interaction*, vol. 31, p. 457–511, 2021.

[10] S. Vargas and P. Castells, "Rank and relevance in novelty and diversity metrics for recommender systems," in *Proceedings of the 5th ACM Conference on Recommender Systems*. ACM, 2011, pp. 109–116.