Characterizing Provider Fairness in Content-Based E-Service Intelligent Recommendation

Raciel Yera[1(✉)], Bapi Dutta[1], Francisco J. Martínez[1,2], and Luis Martínez[1]

[1] Computer Science Department, University of Jaén, 23007 Jaén, Spain
ryera@ujaen.es
[2] Ayesa Servicios Digitales Avanzados, 41005 Sevilla, Spain

Abstract. Fairness is currently regarded as a relevant dimension towards the goal of reaching trustworthy artificial intelligence-based systems. In recommender systems, fairness is focused on addressing biases that may disproportionately benefit or harm certain classes of users and items. In this contribution we are interested on provider fairness, which aims at guaranteeing that the providers of the items would have the same chance for the exposure of their items in the final recommendation lists. Particularly, we will be focused on characterizing the provider fairness associated to intelligent content-based recommendation used for suggesting e-services in a marketplace environment in the region of Extremadura, Spain. Herein, the generalized cross-entropy has been used as metric for characterizing fairness associated to both basic and latent dirichlet allocation (LDA)-based content-based recommendation. As main findings, our study has indicated that the recommendation based on latent dirichlet allocation might lead to better fairness values for those providers with larger number of e-services. For the managerial viewpoint, it suggests that a higher presence in online platforms of the products, might guarantee better associated fairness values. As far as we know, this contribution presents one of the first efforts on evaluating provider fairness recommendation in a concrete e-service scenario.

Keywords: Trustworthy recommendations · provider fairness · e-services · content-based recommender systems · latent dirichlet allocation

1 Introduction

Recommender systems are relevant intelligent-based systems focused on providing users with those information items that best fit their preferences and needs in search spaces overloaded with possible options. Currently, recommender systems have been actively used in a diversity of domains, such as e-commerce, e-learning, e-tourism, e-health, and e-services [7]. For providing accurate recommendations in such contexts, two main recommendation paradigms have been

C. Kahraman et al. (Eds.): INFUS 2025, LNNS 1530, pp. 575–583, 2025.
https://doi.org/10.1007/978-3-031-98565-2_62

developed: content-based recommendation, and collaborative filtering-based recommendation [7].

Beyond recommendation accuracy, the goal of developing trustworthy recommender systems has gained increasing attention in recent years [9], aligning with the concept of trustworthy artificial intelligent highlighted by the AI Act[1]. Across this objective, fairness-aware recommendation has become one of the key dimensions for ensuring trustworthiness. Addressing this issue requires identifying and mitigating challenges such as popularity bias, the marginalization of minority preferences, and algorithmic discrimination [4]. Implementing fairness-aware models allows recommender systems to deliver balanced and inclusive suggestions while still maintaining personalization.

The implementation of fairness criteria in recommender systems has been based in two main dimensions. At first, consumer fairness has been focused on avoiding that the delivered recommendation would be sensitive to some demographic class associated to the users, such as age, nationality, or gender [10]. On the other side, provider fairness has been focused on avoiding the monopoly domination by ensuring a similar chance of exposure in the recommendation lists of all the available items disregarding their associated provider [9].

The objective of the current contribution is the exploration of provider fairness in an e-service scenario. As noted by [7], e-services have been identified as one key application of recommender systems. In recent years, numerous e-service platforms have emerged as ICT-driven solutions aimed not only at benefiting corporations and key stakeholders but also at stimulating economic growth, generating employment, supporting local enterprises, and enhancing overall quality of life. A notable example is D-Rural[2], a collaborative initiative involving multiple institutions dedicated to creating a service marketplace for rural regions across Europe. The research presented in this work will be conducted within this framework, although its findings and methodologies can be extended to other contexts. Concerning the nature of data and the lack of user preferences, our methodology will be based on the content-based recommendation approach. As far as we know, our research is one of the first efforts focused on evaluating provider fairness in real business scenarios.

The contribution has the following structure. First, Sect. 2 illustrates the related works associated to our proposal. Section 3 introduces the methodology for characterizing fairness in e-service recommendation. Section 4 performs the experiments, including dataset, protocol, and results. Section 5 concludes the contribution.

2 Related Works

The lack of fairness in contemporary recommendation systems across different stakeholders has been widely recognized by researchers [4,8]. For example, job-matching platforms often fail to present users with an unbiased selection of

[1] https://artificialintelligenceact.eu/.

[2] https://cordis.europa.eu/project/id/101017304.

employment opportunities; musicians may struggle to achieve fair exposure in music discovery services; information platforms tend to disproportionately highlight certain topics on affected communities; and credit scoring systems frequently favor particular socio-demographic groups, among other cases.

One of the earliest efforts to address fairness was proposed by Felfernig et al. [5], who introduced a fairness-aware approach aimed at reducing the dominance of specific categories and ensuring a more balanced representation. It involves first ranking items within their respective categories, before merging the individual lists through a turn-based strategy that distributes exposure more equitably.

In recent years, research has increasingly focused on provider fairness (P-fairness). Boratto et al. [2] analyzed disparities among item providers in terms of relevance, visibility, and exposure by simulating different levels of minority representation in item catalogs and their interactions. Their proposed approach integrates observation upsampling with loss regularization to refine item rating predictions, ultimately enhancing fairness by improving the visibility, exposure, and representation of underrepresented groups.

Gomez et al. [6] address the issue of under-represented providers in online platforms. By enriching item metadata with the continent of origin, their proposed method enhances equity through a re-ranking strategy that balances both the proportion of recommendations allocated to items from each continent and their placement within the recommendation list.

The conducted analysis has highlighted various research initiatives aimed at improving provider fairness (P-fairness) in recommender systems, most of which are tailored to specific domains. In the context of e-services, where a medium-sized item dataset is expected, there is a need for a transparent characterization of the associated fairness. However, the reviewed approaches do not fully address these specific requirements.

3 The Framework for Characterizing Fairness in E-Service Intelligent Recommendations

This section will present the framework to be used for characterizing p-fairness in the e-services associated to the D-Rural project. It will be composed of the following stages (Fig. 1): 1) Information modeling, 2) Content-based recommendation, and 3) Fairness evaluation.

Fig. 1. The framework for characterizing fairness in e-service intelligent recommendations.

3.1 Information Modeling

This stage will be focused on identifying the information that will be used for modeling e-services. Starting for the information available in the D-Rural platform, we will model e-services through the following attributes:

– IdService: E-service identifier.
– ServiceTitle: E-service name.
– Description: E-service description, mentioning the exact facility that is provided, and their main features and strengths.
– Details: Further description, that sometimes is filled and sometimes missing.
– Category: Key topic where the service is associated, such as health, leisure, technological, and so on.
– Provider: Corporation/agency that manage the service.
– Address: Place where the service is located.

Next stages will use these attributes.

3.2 Content-Based Recommendation

Taking into account that the e-service modeling is based on text and item attributes, we will use two content-based recommendation approaches for generating recommendations in this scenario (Fig. 2), instead of the collaborative filtering-based that depends on user ratings [7].

The first one represents a basic approach that builds TF-IDF vectors using the terms associated to the e-service descriptions [7], and is composed of the following stages, which are not developed in detail due to space limitations.

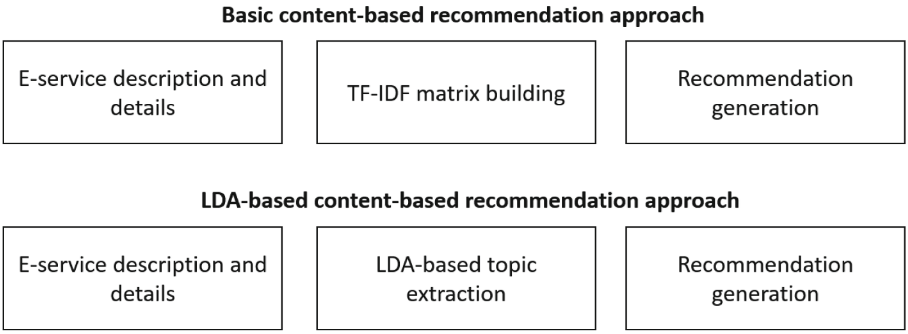

Fig. 2. Schemes of the used content-based recommendation approaches.

1. Use the content associated to the fields Description and Details of the e-service, for reaching a TF-IDF matrix associated with the associated terms.

2. Use the services consumed by the active user, for building its profile also based on the extracted terms. Build the profile of each current user, based on the e-service TF-IDF vectors that he has preferred/consumed.
3. For the active user, obtain the score used for generating recommendations using the linear kernel similarity between the user and e-service profiles.
4. Retrieve the recommendation list with the top scored services.

The second one employs latent dirichlet allocation (LDA) [1] for extracting the topics associated to the descriptions, using them for composing the item profiles. In contrast to the previous approach, stages 1 and 2 would have the following shape:

1. Use the content associated to fields Description and Details of the e-services, for reaching relevant semantic topics using latent dirichlet allocation.
2. Use the services consumed by the active user, for building its profile also based on the extracted terms. Build the profile of each current user, based on the e-service topic vectors that he has preferred/consumed.

3.3 Fairness Evaluation

Deldjoo et al. [3] introduce the use of generalized cross entropy for evaluating fairness in individual recommender systems, which will be the measure used in this work for characterizing fairness of the e-service recommendation. For the discrete attribute case, authors define the generalized cross entropy (GCE) for some parameter $\alpha \neq 0, 1$ as:

$$I(M, a) = \frac{1}{\alpha(1-\alpha)} \sum_{a_j} p_f^{\alpha}(a_j) p^{(1-\alpha)}(a_j) - 1 \tag{1}$$

Here, $a \in A$ is the attribute for which the fairness is evaluated. p and p_f are the actual probability distribution of the attribute a across the recommendation output; and the desired, fair, probability distribution.

p_f is usually formalized as uniform distribution, e.g. regarding that fairness means equality. Deldjoo et al. [3] defines p, for the attribute a_j as:

$$p(a_j) = \frac{\sum_{i \in a_j} rg_i}{Z} \tag{2}$$

where $Z = \sum_i rg_i$. Furthermore, rg_i, which is the recommendation gain for i, is formalized as:

$$rg_i = \sum_{u \in U} \phi(i, Rec_u^K) g(u, i, r) \tag{3}$$

where $\phi(i, Rec_u^K) = 1$ if item i is in the retrieved recommendation list. $g(u, i, r)$ is the gain of recommending for the user u the item i with the ranking r, and is calculated through the discounted cumulative gain (DCG), keeping $rel(u, i) = 1 :$ $g(u, i, r) = \frac{2^{rel(u,i)-1}}{log_2(r+1)}$. The obtained DCG values are finally normalized into the NDCG values, which are used for calculating rg_i

4 Experiments

This section will be focused on executing the presented framework for characterizing the provider fairness associated to an e-service recommendation dataset. Section 4.1 presents the dataset and the experimental protocol that will be used. Section 4.2 presents and discusses the obtained results.

4.1 Dataset and Experimental Protocol

As state above, we will use a dataset associated to a marketplace of e-services deployed at the Extremadura region, in Spain. From the attributes identifying the dataset pointed out in Sect. 3, we will be focused on the company/firm that provides the service (i.e. the service provider).

Overall, our dataset contains items from 31 different providers. However, we have also detected providers with a very reduced number of items, where it has no sense the fairness characterization. Therefore, we will limit our analysis to providers having at least three associated services, which will be Activa-mente Extremadura (7 services), BidInn (3 services), Gc Genomics (5 services), takinto (3 services), BWell Labs (8 services), and ADI&SALU (5 services).

Our experimental protocol is based on three steps:

1. Assume that each user has consumed one specific e-service. Then, we will assume the presence of 62 users, each one evaluating each of the 62 services.
2. For each user, we will apply the procedure pointed out at Sect. 3, for generating their associated recommendation and its fairness value according to each of the specified providers.
3. For each provider it is obtained the average fairness values across all the recommendation lists, which are the values reported in this work.

Particularly it will be considered three recommendation approaches: 1) the simple content-based approach, 2) the LDA-based content-based approach with 5 factors, and 3) the LDA-based content-based approach with 15 factors. These numbers of factors have been set up based on the average length of the descriptions. Furthermore, it will be considered top 5 and top 10, as sizes of the recommendation lists. Our experimental analysis will be mainly focused on answering the following questions: 1) Is there some relationships between the number of e-services linked to the provider, and the associated fairness?, and 2) What is the effect of using a more sophisticated recommendation approach, over the fairness values?

4.2 Results

Tables 1 and 2 illustrate the results for the three considered recommendation approaches, for top 5 and top 10 recommendation lists.

The first important finding to notice, is that for the basic approach it was not identified a clear correlation between the number of services of each provider, and

Table 1. Provider fairness values associated to the discussed content-based recommendation approaches. Top 5 recommendations. Larger values suggest better fairness.

Provider	Basic content-based approach	LDA-based approach (5 factors)	LDA-based approach (15 factors)
Activa-mente Extremadura (7 serv)	–0.3586	–0.3549	–0.3214
BidInn (3 serv)	–0.3296	**–0.0577**	–0.5021
Gc Genomics (5 serv)	–0.2129	–0.5965	–0.3247
takinto (3 serv)	**–0.2125**	∞	**–0.2247**
BWell Labs (8 serv)	–0.3036	–0.4133	–0.2813
ADI & SALU (5 serv)	–0.3393	–0.4303	–0.3392

the corresponding fairness values. As example, while for top 5 recommendations providers with a smaller number of e-services (Gc Genomics and takinto) obtain the best fairness values (-0.2129 and -0.2125), for top 10 it is interesting that the best fairness values were obtained for two providers with respectively the largest and the smallest number of e-services (BWell Labs and takinto).

Table 2. Provider fairness values associated to the discussed content-based recommendation approaches. Top 10 recommendations. Larger values suggest better fairness.

Provider	Basic content-based approach	LDA-based approach (5 factors)	LDA-based approach (15 factors)
Activa-mente Extremadura (7 serv)	–0.8074	**–0.1776**	–0.5561
BidInn (3 serv)	–0.9608	–0.2616	–1.1692
Gc Genomics (5 serv)	–0.9153	–0.8457	–0.7374
takinto (3 serv)	–0.6989	–1.3796	–0.7378
BWell Labs (8 serv)	**–0.5812**	–0.3871	**–0.5417**
ADI & SALU (5 serv)	–0.9666	–1.0565	–1.0215

Furthermore, it can be identified that for the scenarios with the larger number of e-services, such as BWell Labs and Activa-mente Extremadura, the use of a more complex recommendation approach like the LDA-based, tends to lead to an improvement of the fairness values. As example, in top 5 for Activa-mente Extremadura the fairness is improved from -0.3586 to -0.3214 for LDA-based (15 factors), and for BWell Labs it is improved from -0.3036 to -0.2813 also for LDA-based (15 factors). In the case of top 10 recommendations, at Activa-mente Extremadura it is improved from -0.8074 to -0.1776 (5 factors) and at BWell Labs from -0.5812 to -0.3871 (5 factors).

On the other side, for those providers with a smaller number of e-services, the use of a more complex recommendation approach was not necessarily associated to an improved fairness. Particularly, in scenarios such as BidInn and takinto, it obtained a worse fairness values. We think that this behavior should be connected to the fact that the reduced number of services as well as their limited descriptions, do not allow to connect them (through the recommendation approach) to other services belonging to the same provider. Herein, it is interesting the case of takinto for top 5 recommendations and 5 factors, where LDA-based does not retrieve any service belonging to the same provider.

5 Conclusions

The current contribution has been focused on introducing a framework for characterizing provider fairness in e-service intelligent recommendation. It is composed of three stages focused on information modeling, content-based recommendation, and fairness evaluation. The generalized cross entropy measure was used for characterizing fairness. As relevant finding, it has been proved that the recommendation based on latent dirichlet allocation might lead to better fairness values for those providers with larger number of e-services. The next future works will be center on evaluating fairness in group recommendations, as well as proposing methods for improving the obtained fairness in the current context.

Acknowledgement. This project is supported by the European Union's Horizon Europe research and innovation program under the Marie Sklodowska-Curie grant agreement number 101106164. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. We thank the support of Manuel Gimenez from Ayesa.

References

1. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of Machine Learning Research **3**(Jan), 993–1022 (2003)
2. Boratto, L., Fenu, G., Marras, M.: Interplay between upsampling and regularization for provider fairness in recommender systems. User Model. User-Adap. Inter. **31**(3), 421–455 (2021). https://doi.org/10.1007/s11257-021-09294-8
3. Deldjoo, Y., Anelli, V.W., Zamani, H., Bellogin, A., Di Noia, T.: A flexible framework for evaluating user and item fairness in recommender systems. User Modeling and User-Adapted Interaction pp. 1–55 (2021)
4. Deldjoo, Y., Jannach, D., Bellogin, A., Difonzo, A., Zanzonelli, D.: Fairness in recommender systems: research landscape and future directions. User Model. User-Adap. Inter. **34**(1), 59–108 (2024)
5. Felfernig, A., Tintarev, N., Tran, T.N.T., Stettinger, M.: Explanations for groups. In: Group Recommender Systems: An Introduction, pp. 105–126. Springer (2018)
6. Gómez, E., Boratto, L., Salamó, M.: Provider fairness across continents in collaborative recommender systems. Information Processing & Management **59**(1), 102719 (2022)
7. Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G.: Recommender system application developments: a survey. Decis. Support Syst. **74**, 12–32 (2015)
8. Sonboli, N., Burke, R., Ekstrand, M., Mehrotra, R.: The multisided complexity of fairness in recommender systems. AI Mag. **43**(2), 164–176 (2022)

9. Wang, Y., Ma, W., Zhang, M., Liu, Y., Ma, S.: A survey on the fairness of recommender systems. ACM Transactions on Information Systems **41**(3), 1–43 (2023)
10. Yera, R., Barranco, M., Martínez, L.: A novel approach for measuring demographic parity fairness in group recommendation. In: FLINS-ISKE 2024. pp. 195–202. World Scientific (2024)