

Article

Provider Fairness for Diversity and Coverage in Multi-Stakeholder Recommender Systems

Evangelos Karakolis *, Panagiotis Kokkinakos and Dimitrios Askounis

School of Electrical & Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, 15780 Zografou, Greece; pkokkinakos@epu.ntua.gr (P.K.); askous@epu.ntua.gr (D.A.)

* Correspondence: vkarakolis@epu.ntua.gr

Abstract: Nowadays, recommender systems (RS) are no longer evaluated only for the accuracy of their recommendations. Instead, there is a requirement for other metrics (e.g., coverage, diversity, serendipity) to be taken into account as well. In this context, the multi-stakeholder RS paradigm (MSRS) has gained significant popularity, as it takes into consideration all beneficiaries involved, from item providers to simple users. In this paper, the goal is to provide fair recommendations across item providers in terms of diversity and coverage for users to whom each provider's items are recommended. This is achieved by following the methodology provided by the literature for solving the recommendation problem as an optimization problem under constraints for coverage and diversity. As the constraints for diversity are quadratic and cannot be solved in sufficient time (NP-Hard problem), we propose a heuristic approach that provides solutions very close to the optimal one, as the proposed approach in the literature for solving diversity constraints was too generic. As a next step, we evaluate the results and identify several weaknesses in the problem formulation as provided in the literature. To this end, we introduce new formulations for diversity and provide a new heuristic approach for the solution of the new optimization problem.

Keywords: multi-stakeholder recommender systems; diversity; fairness; coverage; optimization

Citation: Karakolis, E.; Kokkinakos, P.; Askounis, D. Provider Fairness for Diversity and Coverage in Multi-Stakeholder Recommender Systems. *Appl. Sci.* **2022**, *12*, 4984. <https://doi.org/10.3390/app12104984>

Academic Editors: Alessandro Micarelli, Giuseppe Sansonetti and Giuseppe D'Aniello

Received: 21 March 2022

Accepted: 13 May 2022

Published: 14 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Item recommendations from traditional recommender systems are in most cases based on the highest predicted ratings for each user. Of course, this criterion is crucial for the success of a recommender system, as it proposes items that the users seem to be interested in. However, it may not be sufficient when it is the only metric taken into consideration. This is because it fails to enable the users to explore new content, and the recommendations for each user tend to be very similar to each other. As a result, relatively unknown and innovative items may never be discovered by the users. Moreover, most users are recommended items of a certain type. This problem is magnified for content-based recommender systems.

These weaknesses have been realized early by scientists and researchers, and several alternative metrics for the evaluation of recommender systems have been defined, including coverage, diversity, fairness, serendipity, and novelty. Regarding coverage, according to [1] an item is covered if it is recommended to at least a certain number of users. Diversity is the average dissimilarity between all pairs of users or items in the result set (per item and per user diversity) [2]. Serendipity in a recommender system is the experience of receiving an unexpected and fortuitous item recommendation [3]. Fairness is mostly used for group recommendations and its intuitive meaning is that a set of items needs to be recommended to a group of users, so that each group member is satisfied in a fair manner [1]. Novelty of any information is defined as the proportion of unknown and known information with respect to the user [4].

Furthermore, it is worth mentioning that, in the last few years, the focus has been shifted from users that receive recommendations to all stakeholders involved in a recommender system. Specifically, the main beneficiaries are item providers, users, and the recommendation system operator. This holistic paradigm in recommender systems, in which all stakeholders need to be jointly taken into consideration, is widely known as the Multi-Stakeholder Recommender Systems (MSRS) paradigm [5].

1.1. Objectives

The paper at hand focuses on item providers that pose requirements for fair item recommendations in terms of the coverage and diversity of users to whom each provider's items are recommended. The notion of fairness here implies that the items of each provider are recommended to a similar number of users as items of other providers, as well as that the users to whom a provider's items are recommended are as diverse as the users that items of other providers are recommended. More specifically, provider fairness (p-fairness) [6] requirements in terms of coverage and diversity are introduced as constraints that need to be fulfilled for each item provider. Hence, average user coverage and diversity per item for each provider should reach at least a certain target, while the expected satisfaction from the provided recommendations should remain as high as possible.

To this end, the presented recommender system aims at maximizing the total predicted rating of the recommendation lists, as well as satisfying the requirements for fair user coverage and diversity per item provider. Specifically, the problem is formulated as a mathematical optimization problem as proposed in the literature [7]. Moreover, the proposed methodology for the solution of the problem is followed. As the problem is not linear and hence it is not computationally feasible to be solved, and due to the fact that in the literature the solution for diversity is not adequately described, a heuristic algorithm which finds solutions that are very close to the optimal one and also fulfill the diversity constraints is introduced. As a next step, the results and the limitations of the proposed approach are presented in brief, and several weaknesses of the proposed formulation are identified. Moreover, two new mathematical definitions for provider diversity are proposed to address the identified weaknesses of the planned approach by the literature, and a high-level methodology is proposed for the solution of the new optimization problem. Finally, several limitations of the newly proposed approach are concisely presented, along with future extensions for the proposed framework.

1.2. Research Questions

The paper at hand has as its main objective to answer the following research questions:

RQ1. How provider fairness in terms of user coverage and diversity can be achieved in recommender systems?

RQ2. Are the results of the proposed methodology satisfactory?

RQ3. What are the limitations of the proposed methodologies?

RQ4. How can the aforementioned limitations be abated?

1.3. Contribution

The purpose of the publication at hand is to build a recommender system that will provide fair recommendations across the different item providers in terms of the coverage and diversity of users to whom their items are proposed. In this context, its contribution is multifold and includes:

- Review of the recent literature about provider fairness, coverage, and diversity.
- Formalization of the problem at hand, as provided in the literature.
- Application and extension of the proposed methodology to a public and well-known dataset.

- Because the proposed solution in the literature is vague in terms of solving the diversity constraints, in this publication, we present a concrete heuristic approach for solving the latter.
- Evaluation of the results and discussion on the proposed solution.
- Introduction of new definitions for quantifying diversity, as the proposed approach in the literature posed significant limitations.
- Introduction of new high-level heuristic approaches for solving the newly-defined diversity constraints as the problem for diversity remains NP-Hard.

2. Background and Related Work

There is a plethora of publications that acknowledge the need to go beyond the accuracy of recommendations in recommender systems. However, there is no commonly accepted approach for quantification of the alternative evaluation metrics of the latter. To this end, in this section we provide a literature review of the recent work in the MSRS paradigm, and the evaluation metrics of diversity, coverage, and fairness, in terms of their definitions and usage. As a next step, we present a publication that proposes a framework for providing multi-stakeholder recommendations with provider coverage and diversity constraints and the proposed solution. Finally, we discuss how the approach proposed in the publication at hand goes beyond the state-of-the-art.

2.1. Multi-Stakeholder Recommender Systems (MSRS)

The MSRS paradigm has gained significant popularity in the last few years. This paradigm, in contrast to the traditional one that puts the user (recommendation receiver) in the center of attention, also takes into consideration other stakeholders of recommender systems [5,8] such as item providers, the recommender system operator, and society in general. In this context, several recent studies focus on this paradigm. Specifically, Refs. [5,8] identify the need for examining recommender systems from the different stakeholders' perspective and introduce the multi-stakeholder recommender systems paradigm. On the other hand, Abdollahpouri et al. [9] investigated multi-sided fairness in multi-stakeholder recommendations and how different fairness concerns can be introduced in such systems. Moreover, Milano et al. [10] analyzed the ethical aspects of MSRS. Furthermore, Refs. [11,12] review the existing case studies, methods, and challenges, and propose new research directions for MSRS. Last but not least, Refs. [13,14] present applications of the MSRS paradigm. The first introduces provider constraints to the multi-stakeholder recommendation problem and formulates it as an integer programming optimization model that is solved using an approximation and can achieve satisfactory results in real use cases. On the other hand, in the second article the authors developed a multi-objective binary integer programming model to allocate sponsored recommendations. As a next step, they present an algorithm to solve the problem in a computationally efficient way. The proposed approach was applied to a real use case with good results and is easily applicable to existing recommender systems as it is applied as a form of postprocessing.

Of course, in the publication at hand, we are not examining MSRS in general; instead, we focus on the metrics of coverage and diversity across provider items and address the need for fairness across the latter. Therefore, it is worth presenting the different definitions of coverage, diversity, and fairness.

2.2. Coverage

Ge et al. [15] defined as item coverage, the proportion of items for which the system is able to generate recommendations (prediction coverage), or the proportion of available items that are recommended to a user (catalog coverage).

It needs to be noted that the previous definitions refer to coverage in items. However, a less popular term is user coverage and is defined as the proportion of users to whom specific items have been recommended [16]. User coverage poses significant interest for

lists of items, meaning that it is important for an item provider to be aware to what proportion of the users his items have been recommended. In this publication, one of the goals of the recommender system is to fulfill user coverage constraints for the different providers.

Although the coverage has been mentioned in a variety of publications, in most cases, it is calculated after the recommendations have been provided and is used only for system evaluation, without being taken into consideration during the recommendation process [15–18]. In [1], Koutsopoulos et al. include constraints for user coverage to the recommendation problem definition. Specifically, they formulate the system recommendation problem as a mathematical optimization problem with constraints for coverage. This approach is close to the approach followed in the context of this publication.

2.3. Diversity

Concerning diversity, a variety of definitions is available in the literature [16]. More specifically, it has been defined as the average pairwise distance among the proposed items, or the total pairwise distance among the recommended items [19]. The first definition is the most popular in the literature. Another interesting aspect of diversity is how the distance (or dissimilarity) is calculated. For instance, when items are modeled as content descriptors, the dissimilarity is calculated through taxonomies [19], while when they are represented as vectors of terms the dissimilarity is calculated through the complement of Jaccard [20] or cosine similarity [21]. On the other hand, when items are represented as vectors of rankings, the most suitable metrics are Hamming distance, the complement of Pearson similarity [22], or the complement of cosine similarity [21]. In our case, we use the complement of Pearson similarity, and the items are represented as vectors of rankings. Of course, there are more approaches for the calculation of distance among items, but these will not be examined in the scope of this publication [20,23].

Similar to diversity for items, diversity for users can also be defined. The idea is that an item would be worth recommending to as many differentiated users as possible. In this context, the diversity for users can be defined as the average pairwise distance among the users to whom an item is recommended. This definition may not be important for individual items or users, but can be of utmost importance for item providers who intend to maximize the diversity of users to whom their items are recommended. This type of diversity is thoroughly analyzed under the context of this paper.

2.4. Fairness

Regarding fairness, in most cases it refers to omitting or reducing prejudice from a machine learning model or a recommender system [6]. However, recommender systems in most cases provide personalized recommendations, and as a result it is difficult to omit prejudice. Moreover, in most cases recommender systems are used by different stakeholders, and fairness issues may be raised for different stakeholder groups. Therefore, a recommender system, as well as for securing fairness for customers, should also be fair for item providers. For this reason, Burke et al. [6] defined fairness for different stakeholders. Specifically, fairness is divided to customer fairness (C-fairness) and provider fairness (P-fairness). In the current publication, provider fairness implies that the items of each provider are recommended to a similar number of users as items of other providers, and that the users to whom a provider's items are recommended are as diverse as the users that items of other providers are recommended to, and this is the ultimate goal of this study.

There are a number of practical applications for addressing provider fairness in recommender systems. For instance, Borrato et al. [24] evaluate provider fairness in terms of disparities in relevance, visibility, and exposure for minority groups and propose a treatment that combines observation up-sampling and loss regularization for user-item relevance scores, with satisfactory experimental results. In comparison to the paper at hand, the aforementioned work focuses on fairness across minorities, while the current study approaches all item providers as equal and provides a recommendation strategy that

achieves at least a certain amount of average coverage and diversity per item for each provider, if such a strategy exists. On the other hand, the previous study enables the association of an item with more than one provider, which is not the case in the paper at hand, because the problem would be much more complex.

Furthermore, Sonboli et al. [25] acknowledged that individual preferences may limit the ability of an RS to produce fair recommendations. Moreover, they introduced a re-ranking approach for fairness-aware recommendations that learns users' preferences across multiple fairness dimensions instead of a single sensitive feature such as race. This approach achieved better experimental results than other approaches in the literature. Unlike our study, the aforementioned approach also focuses on re-ranking of recommended items, which is not the case in this study. Moreover, the aforementioned work pays attention to the diversity of items a user is recommended, while our approach focuses on user diversity for provider items, as well as user coverage of items for each provider.

Moreover, Gomez et al. [26] acknowledge the provider fairness in terms of geographic imbalance in educational recommender systems. Their study was based on real-world data and observe that data are highly imbalanced in favor of the United States, in terms of open courses and interactions. As traditional RSs tend to reinforce the most represented countries (or providers in general), this study proposes an approach that regulates the share of recommendations for each country and their position in the recommendation list. The definition of fairness in the aforementioned publication is the closest to the definition of the paper at hand, as it recognizes the lack of equity in recommendations across different countries in terms of visibility and exposure, while the paper at hand aims to address the imbalance (lack of equity) in the coverage and diversity of users across the different item providers.

Another notable work about fairness in recommender systems is presented in [27]. In this study, Beutel et al. evaluate algorithmic fairness in a real-world recommender system and showcase that measuring fairness based on pairwise comparisons from randomized experiments is a tractable means for reasoning fairness in rankings and propose a new regularizer that helps to significantly enhance pairwise fairness to a large-scale production RS. The proposed approach can identify and improve systematic mis-ranks or under-ranks of items of a particular group. On the other hand, the current study produces recommendations based on predefined thresholds for average diversity and coverage in users for the items of each provider, and as already mentioned does not take into consideration the rank of each item in the recommendation list.

Other practical applications for addressing fairness in recommender systems are presented in [28,29].

2.5. Problem Formalization

After presenting MSRS, as well as the main evaluation metrics that are examined in the context of this publication, alongside several applications of the latter in RS, it is worth presenting the work conducted by Koutsopoulos et al. [7]. In this publication, authors provide a mathematical formalization for the problem of fairness across item providers in terms of user coverage and diversity. Therefore, the problem at hand, along with the proposed solution as defined by Koutsopoulos et al. [7], are presented in brief.

According to Koutsopoulos et al. [7], a set of items, I , and a set of users, U , are considered along with some baseline recommendation system (in our case, item-based Collaborative Filtering) that generates recommendation lists, Lu , for each user, u . Additionally, that C is the number of providers, and each item belongs to exactly one provider, while I_c is the set of items of provider c , $c = 1, \dots, C$ for each user, u , and item, i , of a provider. Finally, riu denotes the predicted rating with the baseline RS algorithm.

It is worth mentioning that the proposed methodology can be generalized for categories or classes of items. In particular, C can also denote the number of categories or classes and not necessarily the different providers. Hence, from now on, provider and category will be used interchangeably in the problem at hand.

The output of the recommendation algorithm is a list of recommended items for each user. Each recommendation list has size L . Additionally, Lu denotes the set of items recommended to the user, u . Lu' denotes the new list of recommended items that should satisfy coverage and diversity constraints.

Moreover, $x = (x_{iu} : i \in I, u \in U)$ denotes the new Boolean recommendation policy (if $x_{iu} = 0$, item i is not recommended to user u , if $x_{iu} = 1$, item i is recommended to user u) that should be found. The total deviation between the baseline recommendation lists and the new recommended ones is:

$$Cost(x) = \frac{1}{L|U|} (\sum_{u \in U} \sum_{i \in Lu} r_{iu} - \sum_{i \in I} \sum_{u \in U} r_{iu} * x_{iu}), \quad (1)$$

Regarding provider coverage, it is calculated as the sum of the item coverage values for each item of a provider or category. Thus, the average per item user coverage for items of a given provider, c , is:

$$Cov(c, x) = \frac{1}{|Ic|} \sum_{i \in Ic} \sum_{u \in U} x_{iu}, \quad (2)$$

Concerning per item diversity of users of provider, c , it is calculated for each provider, c , as follows:

$$Div(c, x) = \frac{1}{|Ic|} \sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} d_{uv} * x_{iu} * x_{iv}, \quad (3)$$

where d_{uv} is the dissimilarity between users u and v .

Instead of the absolute category diversity, the average normalized diversity (per item, per user pair) that will be used is defined as:

$$\overline{Div}(c, x) = \frac{2}{|Ic|} \frac{\sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U: u \neq v} d_{uv} * x_{iu} * x_{iv}}{(\sum_{i \in Ic} \sum_{u \in U} x_{iu}) * (\sum_{i \in Ic} \sum_{u \in U} x_{iu} - 1)}, \quad (4)$$

or simpler:

$$\overline{Div}(c, x) = \frac{2}{|Ic|} \frac{\sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U: u \neq v} d_{uv} * x_{iu} * x_{iv}}{Kc |Ic| * (Kc |Ic| - 1)}, \quad (5)$$

and the constraints for provider coverage become:

$$Cov(c, x) \geq Kc, \quad (6)$$

and for diversity:

$$Div(c, x) \geq Dc, \quad (7)$$

where Kc is the minimum user coverage for the items of a provider, c , and Dc is the minimum threshold for average per item and per user pair diversity for provider, c .

Along with the optimization problem in the same publication, a solution for the coverage constraints in polynomial time is provided, and a low complexity heuristic approach on top of the coverage solution is proposed in order to also solve the diversity constraints. The results of the proposed approach show satisfactory performance, according to the authors.

In the context of this publication, we follow the proposed methodology and introduce a heuristic algorithm for the solution of the diversity constraints. This is because the approach proposed in the aforementioned publication was too generic, suggesting the recommendation of items to users that increase diversity of a category the most and substituting item recommendations with low ranking on the condition that they do not violate coverage constraints. Because it is not easy or obvious to calculate which item recommendations should be substituted, we provide a more concrete approach, and evaluate the results. As a next step we identify several weaknesses with the current definition for diversity and provide two new definitions for it that address the identified weaknesses. Finally, we propose a new heuristic approach for the solution of the newly defined problem.

3. Heuristic Solution Approach

In this section, we present the overall methodology that has been used in order to solve the recommendation problem. It consists of three steps: (a) find a solution for the unconstrained problem, according to the best predicted ratings, (b) find a solution that deviates the least from the unconstrained solution, subject to the constraints for coverage, and (c) based on the solution that fulfills the constraints for coverage and with the minimum possible deviation from the previous solution, change the recommendation lists in a way that increases the most the diversity for the providers that do not fulfill the aforementioned constraints, on condition that the constraints for coverage and diversity that are fulfilled still hold.

3.1. Baseline Solution (Unconstrained Problem)

The first step of the methodology is to find a solution for the unconstrained recommendation problem. For this problem, item-based collaborative filtering is selected as the most proper technique. Specifically, as a first step we create the user-item matrix and predict the ratings of the users for items that they have not rated yet. To do this, we use the K nearest neighbors (KNN) technique [30]. After some experimentation, the best predictions were occurring for $K = 20$. As a similarity metric we used the Pearson similarity [22].

The result of this methodology is a list of items with the best predicted ratings for each user. The results of this approach are used to form the optimization problem of the next steps.

3.2. Addressing Coverage Constraints

Having available the solution from the unconstrained problem, we can solve the problem subject only to the coverage constraints. The optimization problem as described in [7] can be described from the following equations:

$$\max(\frac{1}{L|U|} \sum_{i \in I} \sum_{u \in U} r_{iu} * x_{iu}), \quad (8)$$

subject to:

$$Cov(c, x) = \frac{1}{|Ic|} \sum_{i \in Ic} \sum_{u \in U} x_{iu} \geq Kc, x_{iu} \in \{0,1\}, \quad (9)$$

$$\sum_{i \in I} x_{iu} = L, \text{ for all } u \in U \quad (10)$$

where all the symbols have been described in the previous section. Specifically, the goal is to maximize the total predicted rating of the new recommendation list, while average coverage for the items of each provider should be at least Kc (items of provider, c , should be recommended on average to at least Kc users. The last constraint denotes that each user receives exactly L recommendations.

As described also in [7], the optimization problem with coverage constraints is linear and hence solvable and can be solved in polynomial time. The results of the problem subject to coverage constraints are a list of recommended items for each user; these have the highest predicted rating by each user, that allow coverage constraints to be fulfilled. The aforementioned recommendation lists will be used as a starting point, in order to find a solution that also fulfills the constraints for diversity.

3.3. Addressing Diversity Constraints

Regarding the objective of diversity, the following constraints should be added to the previous model:

$$Div(c, x) = \frac{2}{|Ic|} \frac{\sum_{i \in Ic} \sum_{u \in U} \sum_{v \in U, v \neq u} d_{uv} * x_{iu} * x_{iv}}{Kc|Ic| * (Kc|Ic| - 1)} \geq D, \quad (11)$$

where all the symbols have been described in the previous section. Specifically, the constraints for diversity denote that the average user diversity per item for items of each provider should be at least D .

As the formula for computing the Diversity is in quadratic form, the problem becomes much more complex.

Specifically, the problem at hand, as also identified by Koutsopoulos in [7], is a quadratic constraint integer programming problem, which is NP-Hard. As such, instead of trying to find the optimal solution to the problem, it was decided to develop a heuristic approach that finds a solution close to the optimal, which is also the approach proposed by Koutsopoulos et al. [7]. In the aforementioned publication, the approach proposed for fulfilling diversity constraints is applying item substitutions across providers, through switching of recommended items between users with priority to switches that mostly improve diversity, after making sure that coverage constraints are still satisfied. However, this approach is too generic, as calculating provider diversity and coverage at every step is not an easy task, while it is even more complex to identify the items that increase the diversity the most and the ones that should be omitted from the recommendation list. For this reason, we investigated the definition of diversity more deeply and propose a more concrete approach.

Specifically, for the development of a heuristic solution, a thorough exploratory analysis has been performed on the data to identify how diversity is influenced by other metrics, such as dissimilarity. At first, the intuition was that diversity would increase if an item was recommended to the most dissimilar set of users. However, the initial intuition has been proven wrong.

What was observed instead is that the most influential factor for increasing or decreasing the diversity was the number of users to whom an item has been recommended. Specifically, the higher the number of users to whom the most recommended item of a provider was recommended, the higher is the diversity for the specific provider.

Therefore, a high-level algorithm has been designed in order to address diversity constraints. It uses as a starting point the solution that fulfills the coverage constraints as described in the previous subsection. As a next step, it calculates the diversity for all categories (providers) and finds the ones that do not fulfill the diversity constraints. For all the aforementioned categories, it finds their most recommended item and the users to whom it was not recommended, ordered by their predicted rating for that item, and if they have been recommended with items from categories for which diversity constraint is fulfilled, the item with the worst predicted rating that belongs to a category which fulfills the diversity constraint is no longer recommended to them. Instead, it is substituted by the most recommended item of the category that needs to fulfill the diversity constraint, on condition that diversity constraint for the category of the swapped item still holds. This procedure is applied until the diversity constraint of the specific category is fulfilled.

A high-level but more detailed overview of the developed algorithm of the heuristic solution that was applied is presented in the following algorithm (Algorithm 1):

Algorithm 1. Heuristic algorithm for addressing diversity constraints.

```

1:  $X_{cov} = \text{calculate\_solution\_for\_Coverage}(A_{pred}, Kc)$  // as a linear programming problem through cvxpy
2:  $X_{new} = X_{cov}$  // copy Coverage solution as a starting point for the final solution
3: foreach category  $c$ :
4:    $Div[c] = \text{calculate\_Diversity\_for\_category}(X_{new}, c)$  // as described in formula (11)
5:  $categories\_ordered\_by\_Diversity = \text{argsort}(Div)$ 
6:  $categories\_ordered\_by\_Diversity\_desc = \text{reverse}(categories\_ordered\_by\_Diversity)$ 
7:  $categories\ to\ change, changeable\_categories = [], categories\_ordered\_by\_Diversity$ 
8: foreach category  $c$  in  $categories\_ordered\_by\_Diversity$ :
9:   If  $Div[c] < D$ :

```

```

10:     Categories_to_change.add(c)
11:     changeable_categories.remove(c)
12: foreach category c in categories_to_change:
13:     Most_recommended_item[c] = find_most_recommended_item_of_category(c)
14:     Users_to_recommend = argsort(A_pred[most_recommended_item[c]])
15:     Users_to_recommend_except_recommended = users not recommended with the most recommended item
16:     foreach user u in users_to_recommend_except_recommended:
17:         foreach category c2 in changeable_categories:
18:             rec_items_of_cat_to_usr = find_rec_items_of_category_to_user(c2, u)
19:             foreach item i in rec_items_of_cat_to_usr:
20:                 Xnew[u,i] = 0
21:                 Xnew[u, most_recommended_item[c]] = 1
22:                 // Calculate new Coverage for category c2 according to formula (9)
23:                 Cov_xnew[c2] = calculate_Coverage_for_category(X_new, c2)
24:                 // Calculate new Diversity for categories c2 and c according to formula (11)
25:                 Div_xnew[c2] = calculate_Diversity_for_category(Xnew, c2)
26:                 Div_xnew[c] = calculate_Diversity_for_category(Xnew, c)
27:                 // If Coverage or Diversity constraints for c2 are violated rollback
28:                 If Cov_xnew[c2] < K[c2] or Div_xnew[c2] < D:
29:                     Xnew[u,i] = 1
30:                     Xnew[u, most_recommended_item[c2]] = 0
31:                 Else if Div_xnew[c] > D:
32:                     Break (line 12)

```

For more information, the code of the experiment is available in Github [31].

The algorithm that has been developed follows a greedy approach and avoids backtracking to find a better solution, in case a good enough solution has been found at a previous step. Therefore, the problem becomes efficiently solvable in terms of time and memory. The complexity of the algorithm in the worst-case scenario is $O(L*U*C_{tc}*C_{cc})$, where C_{tc} are the providers for which the diversity constraint does not hold and C_{cc} are the providers for which the diversity constraint holds. More simply, the complexity in the worst-case scenario is $O(L*U*C^2)$. This is because, in the worst case, all items need to be accessed for all users, and the latter have to be accessed for all changeable categories and for all categories or providers that need to change. However, the algorithm completes much earlier in most cases because once a category's diversity constraint is fulfilled it proceeds with a new category of items, without accessing all the users and items.

4. Experiment Details and Results

In this section, we present the dataset that was used and the preprocessing that took place for the solution of the problem. Afterwards, we present the experiment and the results of the heuristic solution.

4.1. Dataset Overview and Preprocessing

The dataset that was used in the context of this publication is a dataset for movie recommendations, and it was taken from MovieLens [32]. The initial dataset of ratings consists of 9724 items (movies) and 610 users. As the dataset is very sparse and there were limited computational resources, the experiments were executed for a subset of items. More specifically, movies that had 10 or less ratings were filtered out, and as a result only 2121 movies were kept.

4.2. Baseline (Unconstrained) Solution

After the preprocessing phase, the user-item ratings matrix was created, based on which the analysis and modeling was performed. Regarding the baseline solution and the coverage constraints, the methodology described in [7] has been followed.

Specifically, after applying five-fold cross validation to the user item (80–20% training and test sets, accordingly), the Pearson similarity between items in the training set is computed, and after some experimentation only the 20 nearest neighbors are kept for each item. More specifically, 20 nearest neighbors led to the lowest Mean Squared Error in predicted ratings on the test set of the initial dataset. Afterwards, based on this similarity metric, ratings for items of the test set were estimated. As a next step, the performance of the item-based Collaborative Filtering (CF) based on mean squared error, which measures the squared difference between predicted and actual ratings, was evaluated. The resulting mean squared error was 0.8605 for the test set. This means that the model had a relatively satisfactory performance, as the predicted ratings deviate 0.86 stars from actual ones in the test set.

Therefore, by adopting this baseline recommendation system, the ratings that users would give to items, which they have not rated so far, are predicted (null values in the initial user-item ratings matrix). The final output is a list of 10 items for each user with the highest predicted ratings that the algorithm produces with no further constraint. The resulted average rating for recommended items for all users was 4.938.

4.3. Item Providers or Categories

The result of the item-based CF method maximizes the sum of ratings that the recommended movies provide; however, the total rating is not the only metric that should be taken into consideration as the objective of the problem. In particular, the goal of this problem is to make sure that the coverage and the diversity metrics for each item provider will also be taken into consideration. As described by Koutoupoulos et al. [7], this is achieved by setting thresholds for the value of both Coverage and Diversity for each item provider.

Because coverage and diversity are defined as provider-related metrics, different providers have been created, and each item is connected to a specific one. The experiment has been conducted for different numbers of item providers with similar results. Therefore, only the results for twenty item providers are presented. Specifically, twenty item providers were created, and only one provider was assigned to each item. The provider assignment was performed randomly, with uniform probability, and as a result the distribution of providers is almost uniform, as shown in the Table 1:

Table 1. Different providers or categories and their number of items.

Category	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Items	101	91	105	115	105	107	90	121	101	101	90	114	94	120	115	104	114	108	107	118

4.4. Coverage Solution

The coverage constrained problem was solved for discrete values of x (specifically 0 and 1) because an item is either recommended or not recommended to a user. For the solution of the problem, the cvxpy [33] python library was utilized. Although the new lists for different values of K are different from the baseline lists, Lu , the results of the total rating for different coverage thresholds found with the discrete solution are the same for different values of coverage. In particular, the total rating is the same for $K = 1, 1.5, 2, 2.5$, while for $K \geq 3$ there is no solution. This fact occurs because there are many excellent ratings (rating = 5) in the predicted rating matrix, and the items of the new lists Lu' are different from the ones of the baseline recommendation list Lu , but the new recommended

items are rated perfectly, the same as the old ones (5 stars). However, the distribution of coverage changes per item provider (category), as shown in Figures 1 and 2:

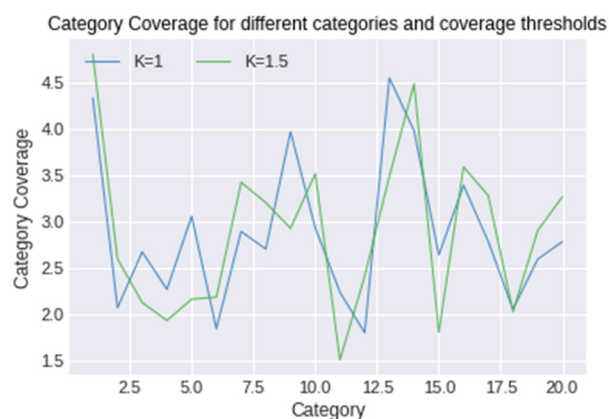


Figure 1. Coverage per provider for $K = 1$ and $K = 1.5$.

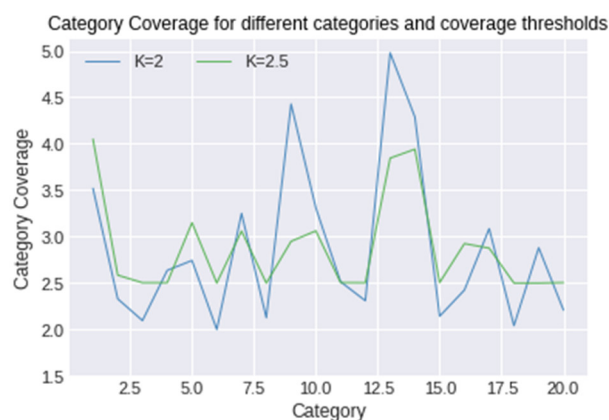


Figure 2. Coverage per provider for $K = 2$ and $K = 2.5$.

As observed from the figures above, as K increases, the coverage tends to be more uniform.

4.5. Final Solution

As already mentioned, the solution for coverage is used as a starting point in order to find a feasible solution that also fulfills the diversity constraints. As a next step, the heuristic approach of Algorithm 1 has been followed. The result of the proposed solution is shown in Figure 3:

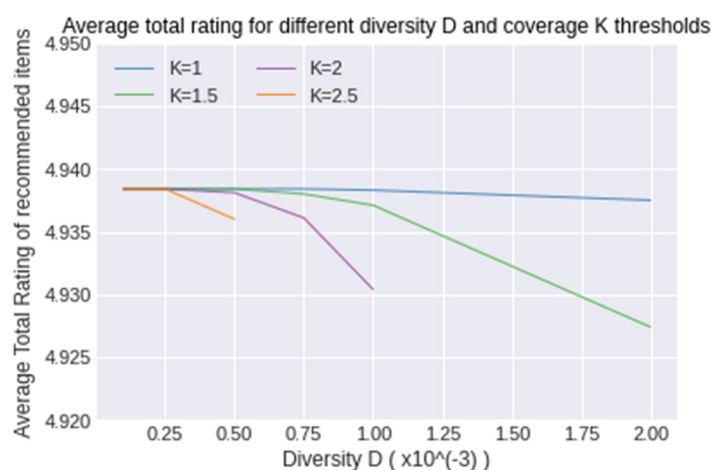


Figure 3. Average rating for different values of diversity and coverage.

As observed in the figure above, for every value of coverage, as the diversity threshold increases, the average rating of recommended items is similar or drops. Furthermore, for greater coverage thresholds, the problem becomes more difficult to solve and the average rating drops more rapidly with the increase of diversity threshold. More specifically for coverage threshold $K_c = 2.5$, the problem has no solution for a Diversity threshold greater than 0.0005. In Table 2, the total and average rating of the recommended items are presented for different values of Diversity threshold and for Coverage threshold $K_c = 1$ and $K_c = 2.5$.

Table 2. Average and total rating for different values of diversity and coverage.

Diversity Threshold D	Total Rating ($K_c = 1$)	Average Rating ($K_c = 1$)	Total Rating ($K_c = 2.5$)	Average Rating ($K_c = 2.5$)
0.0001	30,124.411	4.9384	30,124.411	4.9384
0.0002	30,124.411	4.9384	30,124.411	4.9384
0.00025	30,124.411	4.9384	30,124.411	4.9384
0.0005	30,124.411	4.9384	30,110.984	4.9362
0.00075	30,124.411	4.9384	No solution	No solution
0.001	30,124.120	4.9383	No solution	No solution
0.002	30,119.159	4.9375	No solution	No solution

As observed from the table and the figure above, the total rating of all the recommended items for all users is very high and, in most cases, it is the same as the coverage solution. This shows that the heuristic solution chosen may not be the optimal one but produces a result that is very close to the optimal.

In order to provide a better view on how diversity influences the total rating, in the following plots the total rating for coverage threshold $K = 1.5$ and $K = 2$ for different thresholds of diversity is presented (Figures 4 and 5):

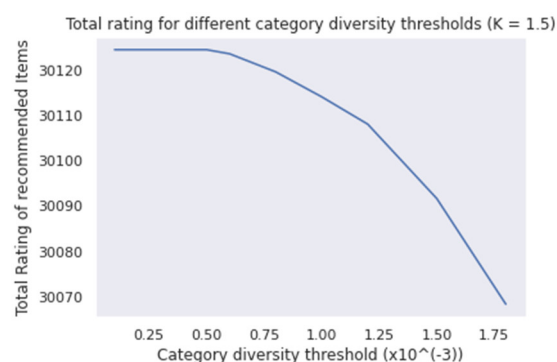


Figure 4. Total rating for different category Diversity thresholds for $K = 1.5$.

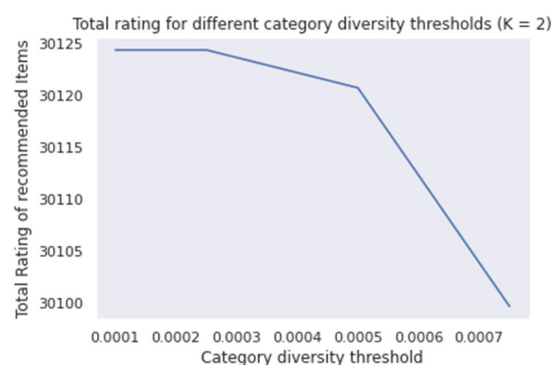


Figure 5. Total rating for different category Diversity thresholds for $K = 2$.

From the plots above (Figures 4 and 5), it is observed that, as the diversity threshold increases, the total rating decreases, while the decrease of the total rating is more important for greater diversity thresholds and not proportional. Concerning the relation between diversity and coverage, it is observed that, for a bigger coverage threshold, the total rating decreases faster with the increase of diversity. This is also what was expected, because, for a higher coverage threshold, on the one hand a stricter constraint is introduced, and on the other, less exchanges with items from other categories are allowed.

Moreover, as observed in Table 2, the results in terms of total or average rating are very close to the ones of the unconstrained solution. The total rating of the latter serves as the upper bound for the optimal solution, because the constrained problem cannot provide more relevant recommendations than the unconstrained one.

As stated before, the approach that has been followed is heuristic and does not always provide the optimal solution to the optimization problem. Instead, it always provides a solution that is close to the optimal one. Specifically, the algorithm that has been proposed follows a greedy approach and avoids backtracking to find a better solution, in case a good enough solution has been found at a previous step.

The cases where the algorithm fails to identify the optimal solution stem from the fact that when the diversity constraint is fulfilled for a certain category, the recommendations of this category can only change on the condition that the diversity and coverage constraints for this category are not violated with a certain swap. Thus, it is possible to find a different recommendation plan in which the total predicted rating would be slightly higher than the total rating of the heuristic solution. For example, the approach at hand, fails to identify cases where the most recommended item's rating of a category is not as good as the second ones, while by recommending the second most recommended item of the category fulfills the diversity constraint. Moreover, if in a certain swap a category's diversity constraint is violated, the heuristic solution will not revisit the latter category to

fix its diversity, even if there are options to do that. However, at least for the specific dataset the results are very close, not only to the optimal solution but also to the baseline solution, which is the upper bound for the total rating.

5. Discussion

5.1. Evaluation

The approach that has been followed to solve the optimization problem for provider fairness in terms of coverage and diversity finds a heuristic solution to the problem that, according to the results, is very close to the optimal solution as it is very close to the results of the baseline recommender system that does not pose any constraint for coverage or diversity and hence acts as an upper bound for the proposed solution.

Concerning the comparison of our results to other studies, only Koutsopoulos et al. [7] dealt with the same problem of provider fairness in terms of user coverage and diversity. The results of the proposed algorithm are similar to the experimental results in the aforementioned work, meaning that as diversity and coverage requirements increase, the total predicted rating of the recommended items tends to drop, while diversity constraints have a larger impact to the total rating. Moreover, our approach shows smaller deviation from the baseline solution, which poses the upper bound on the total predicted rating. Of course, this is not a safe inference as the solution of [7] does not clearly describes neither the preprocessing that took place nor the hyperparameters that have been used.

In general, the advantage of treating the recommendation problem as a constraint optimization problem is that it ensures those certain criteria (the constraints) will be fulfilled. Of course, the constraints should be defined, also taking into consideration the tradeoff between the constraint requirements and the user preferences. This means that if the constraints are too stringent the relevance of the recommended items with the user may be critically low. Therefore, recommender system operators should always evaluate and update their requirements.

Consequently, the proposed approach or a similar optimization approach with constraints for equal representation of providers or other stakeholder groups can be used in order to soften “The winner takes it all phenomenon” [26], in which the most significant providers (or groups of providers) have huge advantage over the others. This is because the optimization algorithm would secure at least a certain target of exposure (coverage, diversity, or any other metric) for all groups, if such a solution existed. Moreover, the proposed approach can also ensure that minority item providers will be protected as they would be guaranteed at least a certain amount of coverage and diversity. Additionally, with a similar approach, exposure and visibility for minority groups can also be achieved.

However, the literature-based definition of the problem as an optimization one [7], poses several limitations. Firstly, it does not take into account the position of an item in the recommendation list. This is problematic because in most cases the position of the items in the user’s list of recommendations plays a vital role in users’ behavior, especially when the recommended items are numerous. For instance, a user may pay attention to the top few recommendations but not to the bottom ones.

Another limitation of the problem definition at hand is that the definition of diversity fails to capture the dissimilarity among users to whom an item is recommended. Instead, it is observed that the most recommended items of a provider influence the diversity significantly more than the dissimilarity among the users to whom an item is recommended, and this fact is exploited by the heuristic approach that has been developed.

To make this weakness more evident, we tried to approach the item’s diversity as defined in the literature, by supposing that the dissimilarity of two different users is fixed and equals the average dissimilarity between two different users. With this approximation, if d is considered as the average dissimilarity between any two different users, the formula for computing the diversity becomes:

$$Div(c, x) = \frac{2 \sum_{i \in I_c} \sum_{u \in U} \sum_{v \in U, v \neq u} d_{x i u} * x_{i v}}{|I_c| Kc |I_c| * (Kc |I_c| - 1)} \geq D, \quad (12)$$

However, if the item's total Diversity is analyzed further, Formula (12) results in:

If the item was recommended to one user, it is:

$$Div(item) = 0$$

If the item was recommended to two users 1 and 2 it is:

$$Div(item) = d(1,2) = d$$

If the item was recommended to three users 1, 2, and 3 it is:

$$Div(item) = d(1,2) + d(1,3) + d(2,3) = 2d + d = 3d$$

With the same approach for n users:

$$Div(item) = (1 + 2 + 3 + \dots + (n-1)) * d$$

Given that $(1 + 2 + 3 + \dots + n - 1)$ is a well-known sum and is equal to $\frac{1}{2}n(n-1)$, finally:

$$Div(item) = \frac{1}{2}n(n-1)d, \quad (13)$$

The item diversity for all user pairs was approximated according to the formula above, and the result is illustrated in the following figures:

Specifically, in the Figures 6 and 7, the red line illustrates the diversity approximation for different values of n (number of users to whom an item is recommended), while the blue points illustrate the actual item diversity calculated for different items and their correlation with n . Figure 6 shows the results after removing outliers, while Figure 7 illustrates all the item diversities for every different occurred value of n .

Correlation between #users to whom an item was recommended vs item diversity

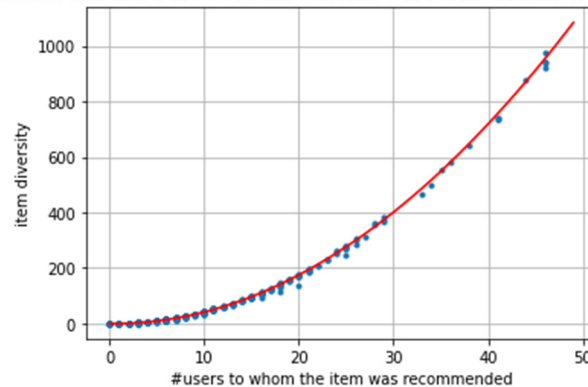


Figure 6. Relation between item diversity and number of users to whom the items were recommended (outliers removed).

Correlation between #users to whom an item was recommended vs item diversity

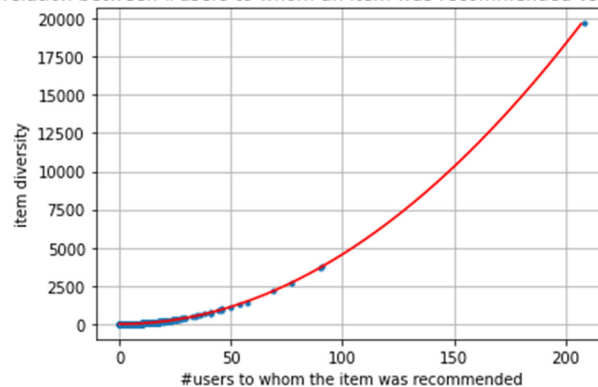


Figure 7. Relation between item diversity and number of users to whom the items were recommended.

As observed, the approximation seems to fit almost perfectly, at least for the specific dataset. In particular, the actual values of item Diversity are very close to the predicted ones, and this is the case for every item in the dataset.

At this point it is worth mentioning that several tests (e.g., the Kolmogorov–Smirnov for both uniformity and normality) were applied to check if the dissimilarity is either fixed or follows normal distribution around d . However, in all cases the null hypothesis was rejected and statistically the approximation is not valid for this dataset. However, the approximation is more than satisfactory and illustrates that instead of paying attention to user dissimilarity, the proposed definition for item diversity favors the most recommended items.

Of course, item providers would not be happy if their most recommended items were recommended to more users but their least recommended ones would be recommended to less users, even if the average diversity of their items would be significantly increased.

Hence, the definition of diversity needs to be reexamined. Moreover, new solutions should be designed to solve the problem efficiently.

5.2. Redefining User Diversity and a New Heuristic Solution Approach

In order to identify the problems of the proposed definition of provider diversity, we will revisit its definition as presented in the literature [7]:

First of all, user diversity for an item is defined as the sum of dissimilarities among users to whom the specific item is recommended:

$$Div(i) = \sum_{u \in U} \sum_{v \in U, v \neq u} duv * xiu * xiv, \quad (14)$$

while the diversity of an item provider is defined as the sum of item diversities for items that belong to the specific provider.

As a next step the category diversity is normalized by dividing with the number of pairs of users to whom items of a specific provider are recommended. As a result, provider diversity per item and user pair is defined in Formula (4) or simplified (using the coverage constraint) in Formula (5).

Conceptually it can be defined as:

$$\overline{Div(c)} = \frac{2}{|Ic|} \frac{\text{sum of all diversities for items in the specific category}}{\text{all pairs of users to whom items of class } c \text{ are assigned}}, \quad (15)$$

Although this definition normalizes the category diversity according to the pairs of users its items are recommended to, it fails to capture the dynamics of most recommended items, as illustrated in the previous section. Consequently, if a provider owns x items that are recommended to y users, the provider diversity can be significantly increased if all users are recommended with the most recommended item, and this should not be happening.

As a result, the definition of provider diversity per item and user pair should either normalize the diversities per item or define a new diversity for the entire set of items for the specific provider. In the first case, the definition of provider diversity per item and user pair becomes:

$$\overline{Div'(c)} = \frac{2}{|Ic|} \sum_{i \in Ic} \frac{\sum_{u \in U} \sum_{v \in U: u \neq v} duv * xiu * xiv}{(\sum_{u \in U} xiu) * (\sum_{u \in U} xiu - 1)}, \quad (16)$$

Of course, with the previous definition for diversity, the optimization problem is not solvable, and of course the heuristic solution that was presented in Section 4 is not valid anymore.

In the second case, the provider diversity, instead of the sum of item diversities, can be calculated for each provider by considering all the items of diversity as one super item. In this case, we should create a new matrix $Y(c, u)$ ($c \in C$, $u \in U$) that denotes the Boolean recommendation policy (if $y(c, u) = 0$ no item for category c has been recommended to user u , and if $y(c, u) = 1$ one or more items of category c have been recommended to user u).

As a result, category diversity can be defined as:

$$\overline{Div''(c)} = \frac{2}{|Ic|} \frac{\sum_{u \in U} \sum_{v \in V: u \neq v} duv * ycu * ycv}{(\sum_{u \in U} ycu) * (\sum_{u \in U} ycu - 1)}, \quad (17)$$

In the previous definition, the denominator can be simplified as in [7]:

$$\overline{Div''(c)} = \frac{2}{|Ic|} \frac{\sum_{u \in U} \sum_{v \in V: u \neq v} duv * ycu * ycv}{Kc|Ic| * (Kc|Ic| - 1)}, \quad (18)$$

The last definition is much simpler (although it still places quadratic constraints on the problem, which is NP-Hard), as the number of categories is much smaller than the number of items. However, it is not suitable to be used for small systems, for which the users have rated a significant amount of items. Instead, for large-scale instances of the recommendation problem and large user-item matrices that are sparse, this definition can prove very helpful.

Regarding the solution of the newly defined diversity per provider, in both cases a heuristic approach can also solve the problem. For example, in the first case, by recommending the best rated item to users that have the greatest distance from the centroid of the users to whom the item was recommended, substituting the worst rated item (from another category) that was recommended to them, on condition that the coverage and the diversity constraints for the category of the swapped item are still fulfilled. The same approach can also be applied in the second solution, by recommending the best rated items to users that are the most distant from the centroid of the users who they had recommendations from the specific provider. Moreover, in case the number of users is very large, clustering techniques can be used to cluster the users according to the similarity between them, and the item swaps can be performed only for users from the most distant clusters.

5.3. Answers to Research Questions

After presenting the literature, the methodology, and the results and acknowledging the limitations of the proposed approach, it is time to answer the research questions that were presented in Section 1.

RQ1. How can provider fairness in terms of user coverage and diversity be achieved in recommender systems?

The literature provides a formalization of the problem at hand, alongside a solution. The literature models the problem as a quadratic constraint mathematical programming problem, which is NP-Hard and hence finding the optimal solution is not feasible. As a result they propose a heuristic solution for solving diversity constraints. However, as the latter is not adequately described, we introduce a new heuristic approach for the solution of diversity constraints.

RQ2. Are the results of the proposed methodology satisfactory?

The results of the proposed heuristic approach not only achieve the targets of sufficient user coverage and diversity per item provider, but are also very close in terms of predicted ratings with the recommendation lists provided by the solution of the unconstrained problem, which acts as an upper bound for the heuristic solution.

RQ3. What are the limitations of the proposed methodologies?

There are several limitations with the proposed approach that stem from the formalization of the problem. Specifically, the definition of diversity fails to capture the dissimilarity among users to whom an item is recommended. Instead, it is observed that the most recommended items of a provider influence the diversity significantly more than the dissimilarity among the users to whom an item is recommended, and this fact is exploited by the heuristic approach that has been developed. Hence, the problem definition and in particular the quantification of diversity should be redefined. As a result, new solutions to the problem should be developed.

Moreover, the proposed approach does not take into consideration the position of an item in the recommendation list, which is of utmost importance in synchronous recommender systems.

RQ4. How can the aforementioned limitations be abated?

In order to address the problematic definition of diversity, we redefine it under the context of Section 5.2. Specifically, two new definitions are introduced, discussing their pros and cons. Moreover, new high-level heuristic approaches are proposed for the solution of the newly-defined diversity constraints.

6. Conclusions and Future Extensions

The contribution of this publication is multifold. Specifically, the methodology proposed by the literature is applied to solve the problem of provider fairness for user diversity and coverage in Multi-stakeholder recommender systems. However, as the problem at hand is NP-Hard and the methodology proposed in the literature is too generic, a new heuristic algorithm for solving diversity constraints is proposed, along with its complexity and limitations. The results of the proposed solution are presented, and they are very close to the optimal solution for the specific dataset.

As a next step, the results were evaluated and discussed thoroughly, and several weaknesses to the proposed problem definition were identified. Specifically, the proposed approach is not taking into consideration the ranking of items in the recommendation list. Furthermore, after digging further to the definition of diversity to explore how the latter is influenced by dissimilarity and other metrics, it was discovered that, with the current definition, the most influential factor for provider diversity is the number of users to whom the most diverse item of a provider has been recommended instead of the dissimilarity among the users to whom it was recommended. To this end, in this publication, we analyzed the proposed definition for diversity and identified why this weakness appears. Finally, we proposed two new definitions for provider diversity as well as a new high-level heuristic solution approach.

The next step forward regarding the proposed approach is to take into consideration the ranking of recommendations in the recommendation problem, as the position of an item in a recommendation list is very important for the recommender system users. This will be carried out using more evaluation metrics such as the Normalized Discounted Cumulative Gain (NDCG), which measures the quality of ranking of the recommended items for each user.

Furthermore, minority representation will be examined in terms of user coverage and diversity, against other well-known approaches (e.g., [24]). Moreover, different metrics will be examined in order to be formalized as constraints in the optimization problem.

Finally, in the near future, we intend to apply the proposed approach to several test datasets, as well as to a real-life recommender system. Finally, the results will be evaluated in terms of coverage, diversity, and satisfaction by real users including item providers and simple users of the system.

Author Contributions: Conceptualization, E.K.; methodology, E.K.; software, E.K.; validation, P.K. and D.A.; formal analysis, E.K.; investigation, E.K.; data curation, P.K.; writing—original draft preparation, E.K.; writing—review and editing, E.K. and P.K.; visualization, P.K. and E.K.; supervision, D.A.; project administration, D.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS, ICCS.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset that has been used for this publication has been taken from MovieLens and is available at [32].

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Koutsopoulos, I.; Halkidi, M. Efficient and Fair Item Coverage in Recommender Systems. In Proceedings of the IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, Athens, Greece, 12–15 August 2018; pp. 912–918.
2. Kunaver, M.; Požrl, T. Diversity in recommender systems—A survey. *Knowl.-Based Syst.* **2017**, *123*, 154–162. <https://doi.org/10.1016/j.knosys.2017.02.009>.
3. McNee, S.M.; Riedl, J.; Konstan, J.A. Being accurate is not enough: How accuracy metrics have hurt recommender systems. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*; Association for Computing Machinery: New York, NY, USA, 2006. <https://doi.org/10.1145/1125451.1125659>.
4. Saranya, K.G.; Sudha Sadasivam, G. Personalized News Article Recommendation with Novelty Using Collaborative Filtering Based Rough Set Theory. *Mob. Netw. Appl.* **2017**, *22*, 719–729. <https://doi.org/10.1007/s11036-017-0842-9>.
5. Abdollahpouri, H.; Burke, R.; Mobasher, B. Recommender systems as multistakeholder environments. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, Bratislava, Slovakia, 9 July 2017. <https://doi.org/10.1145/3079628.3079657>.
6. Burke, R.; Sonboli, N.; Ordoñez-Gauger, A. Balanced Neighborhoods for Multi-sided Fairness in Recommendation. In Proceedings of the Conference on Fairness, Accountability and Transparency, New York, NY, USA, 23–24 February 2018.
7. Koutsopoulos, I.; Halkidi, M. Optimization of Multi-stakeholder Recommender Systems for Diversity and Coverage. In Proceedings of the IFIP Advances in Information and Communication Technology, Crete, Greece, 25–27 June 2021; Volume 627. https://doi.org/10.1007/978-3-030-79150-6_55.
8. Burke, R.; Abdollahpouri, H.; Malthouse, E.C.; Thai, K.P.; Zhang, Y. Recommendation in multistakeholder environments. In Proceedings of the 13th ACM Conference on Recommender Systems, Copenhagen, Denmark, 16–19 September 2019. <https://doi.org/10.1145/3298689.3346973>.
9. Abdollahpouri, H.; Burke, R. Multi-stakeholder recommendation and its connection to multi-sided fairness. In Proceedings of the CEUR Workshop, Lviv, Ukraine, 16–17 May 2019; Volume 2440.
10. Milano, S.; Taddeo, M.; Floridi, L. Ethical aspects of multi-stakeholder recommendation systems. *Inf. Soc.* **2021**, *37*, 35–45. <https://doi.org/10.1080/01972243.2020.1832636>.
11. Abdollahpouri, H. Multistakeholder recommendation: Survey and research directions. *User Model. User-Adapt. Interact.* **2020**, *30*, 127–158. <https://doi.org/10.1007/s11257-019-09256-1>.
12. Zheng, Y. Multi-stakeholder recommendations: Case studies, methods and challenges. In Proceedings of the 13th ACM Conference on Recommender Systems, Copenhagen, Denmark, 16–20 September 2019. <https://doi.org/10.1145/3298689.3346951>.
13. Sürer, Ö.; Burke, R.; Malthouse, E.C. Multistakeholder recommendation with provider constraints. In Proceedings of the 12th ACM Conference on Recommender Systems, Vancouver, BC, Canada, 2 October 2018. <https://doi.org/10.1145/3240323.3240350>.
14. Malthouse, E.C.; Vakeel, K.A.; Hessary, Y.K.; Burke, R.; Fudurić, M. A multistakeholder recommender systems algorithm for allocating sponsored recommendations. In Proceedings of the CEUR Workshop, Lviv, Ukraine, 16–17 May 2019; Volume 2440.
15. Ge, M.; Delgado-Battenfeld, C.; Jannach, D. Beyond accuracy: Evaluating recommender systems by coverage and serendipity. In Proceedings of the Fourth ACM conference on Recommender Systems, Barcelona, Spain, 26–30 September 2010. <https://doi.org/10.1145/1864708.1864761>.
16. Kaminskas, M.; Bridge, D. Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-Accuracy objectives in recommender systems. *ACM Transactions on Interactive Intelligent Systems. ACM Trans. Interact. Intell. Syst.* **2016**, *7*, 1–42. <https://doi.org/10.1145/2926720>.
17. Rahman, M.; Oh, J.C. Graph bandit for diverse user coverage in online recommendation. *Appl. Intell.* **2018**, *48*, 1979–1995. <https://doi.org/10.1007/s10489-017-0977-1>.
18. Hammar, M.; Karlsson, R.; Nilsson, B.J. Using maximum coverage to optimize recommendation systems in E-commerce. In Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, China, 12–16 October 2013. <https://doi.org/10.1145/2507157.2507169>.
19. Ziegler, C.-N.; McNee, S.M.; Konstan, J.A.; Lausen, G. Improving recommendation lists through topic diversification. In Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, 10–14 May 2005. <https://doi.org/10.1145/1060745.1060754>.
20. Vargas, S.; Castells, P.; Vallet, D. Intent-oriented diversity in recommender systems. In Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, Beijing, China, 24–28 July 2011. <https://doi.org/10.1145/2009916.2010124>.
21. Ekstrand, M.D. Collaborative Filtering Recommender Systems. In *The Adaptive Web*; Springer: Berlin/Heidelberg, Germany, 2011.
22. Kelly, J.P.; Bridge, D. Enhancing the diversity of conversational collaborative recommendations: A comparison. *Artif. Intell. Rev.* **2006**, *25*, 79–95. <https://doi.org/10.1007/s10462-007-9023-8>.
23. Yu, C.; Lakshmanan, L.V.S.; Amer-Yahia, S. Recommendation diversification using explanations. In Proceedings of the 2009 IEEE 25th International Conference on Data Engineering, Shanghai, China, 29 March–2 April 2009. <https://doi.org/10.1109/ICDE.2009.225>.
24. Boratto, L.; Fenu, G.; Marras, M. Interplay between upsampling and regularization for provider fairness in recommender systems. *User Modeling User-Adapt. Interact.* **2021**, *31*, 421–455. <https://doi.org/10.1007/s11257-021-09294-8>.

25. Sonboli, N.; Eskandarian, F.; Burke, R.; Liu, W.; Mobasher, B. Opportunistic Multi-aspect Fairness through Personalized Re-ranking. In Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, Genoa, Italy, 12–18 July 2020. <https://doi.org/10.1145/3340631.3394846>.
26. Gómez, E.; Shui Zhang, C.; Boratto, L.; Salamó, M.; Marras, M. The Winner Takes it All: Geographic Imbalance and Provider (Un) fairness in Educational Recommender Systems. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Online, 11–15 July 2021. <https://doi.org/10.1145/3404835.3463235>.
27. Beutel, A.; Chen, J.; Doshi, T.; Qian, H.; Wei, L.; Wu, Y.; Goodrow, C. Fairness in recommendation ranking through pairwise comparisons. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019. <https://doi.org/10.1145/3292500.3330745>.
28. Lin, X.; Zhang, M.; Zhang, Y.; Gu, Z.; Liu, Y.; Ma, S. Fairness-aware group recommendation with pareto-efficiency. In Proceedings of the Eleventh ACM Conference on Recommender Systems, Como, Italy, 27–31 August 2017. <https://doi.org/10.1145/3109859.3109887>.
29. Serbos, D.; Qi, S.; Mamoulis, N.; Pitoura, E.; Tsaparas, P. Fairness in package-to-group recommendations. In Proceedings of the 26th International Conference on World Wide Web, Perth, Australia, 3–7 April 2017. <https://doi.org/10.1145/3038912.3052612>.
30. Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001. <https://doi.org/10.1145/371920.372071>.
31. Provider fairness for coverage and diversity experiments Github repository. Available online: https://github.com/vkarakolis-epu/recsys_provider_fairness_optimization (accessed on 10 May 2022).
32. Movielens datasets. Available online: <http://files.grouplens.org/datasets/movielens/ml-latest-small.zip> (accessed on 10 May 2022).
33. Cvxpy Python library, convex optimization, for everyone. Available online: <https://www.cvxpy.org/> (accessed on 10 May 2022).