FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Bc. Ladislav Maleček

Fairness in group recommender systems

Department of Software Engineering

Supervisor of the master thesis: Mgr. Ladislav Peška, Ph.D.
Study programme: Computer Science
Study branch: Artificial Intelligence

Prague 2021

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

Dedication.

Title: Fairness in group recommender systems

Author: Bc. Ladislav Maleček

Department: Department of Software Engineering

Supervisor: Mgr. Ladislav Peška, Ph.D., Department of Software Engineering

Abstract: Abstract.

Keywords: key words

Contents

1. Introduction

Most of us interact with many recommender systems daily. Even if seemingly indirectly. The proliferation of this technology is astounding. Almost every interaction with today's web is in some way personalized. From the search results, shopping, listening to music, reading news, to browsing social media, and many more. It has become quite unavoidable.

We can view recommender systems from a very simple perspective - they are algorithms that recommend items to users. Where items and users can be many different things, items, for example, being movies, news articles, a more complex object, or even entire systems. And users being, real people or other entities that exhibit some sort of preference on which the algorithm can decide.

One of the variants of recommender systems is those where the recommendation result is shared among more users based on their shared (aggregated) preferences. This is a subset called group recommender systems. They are not used as widely as the non-group variants due to the nature of the usage of most of the aforementioned technologies. We mostly use the web, listen to music and read the news as individuals. At least from the perspective of those systems. But for some of the domains, there are valid use cases. We often listen to music and watch movies in groups. Select a restaurant and other public services not just for us. And that's where the group recommenders come in handy.

tady je trochu moc velky skok - nejdriv asi neco o tom jak group RS funguji. Spis nez evaluation zminit primo fairness, nebo obecneji co ma byt cilem skupinoveho doporuceni

But how to approach the evaluation? It starts to become harder than just simply rating the results based on single feedback, now we have multiple users with possibly very different personal experiences. We want to be fair towards all individuals in the group. But the fairness property can be tricky to describe and evaluate due to the subjective nature of preference perception and distribution among the group members.

Classical recommendation systems have been studied for quite a long time, but the group variant and more soft-level (meaning evaluation on other than classical parameters) thinking about them is quite recent. With the rise of social dilemmas around recommender systems appears the fairness-ensuring topic more and more in many different shapes and sizes. And with that, there is growing popularity towards recommender systems that are trained (and therefore evaluated) with these novel requirements in mind.

1.1 Problem statement

The current research on the topic of group recommender systems is lacking. There are no standardized data sets that would offer evaluation of the research without using various methods of data augmentation and artificial data creation. And the definition of fairness is not unified. It can mean many different things and be evaluated with many different methods.

These two aforementioned problems go hand in hand with the very subjective nature of user preference.

Mozna merge obou kapitol dohromady?

1.2 Research objective

We would like to study how fairness can be defined in the context of the recommender systems, how it can be measured and eventually used to improve recommendations in the group setting. And explore different variants of fairness such as long-term fairness and different distribution of fairness among group members.

The primary goal of this thesis is to research and design a novel group recommender system algorithm that would keep fairness as its primary optimization objective. If we could adapt fairness preserving methods such as voting systems from other fields to group recommendation problem. And evaluating the new algorithm with already existing approaches in the domain of group recommender systems.

Additionally, we would like to research and contribute to data sets that could be used for the group setting. Expanding single user data sets with data augmentation that would generate synthetic groups' information and creating a web application in a movie domain that would serve as a platform for online evaluation of group recommender algorithms and provided us with real-user group recommendation data.

1.3 Thesis structure

We start with an introduction to recommender systems and specifically to group recommender systems in the chapter: 2. Then we will continue with the definitions and evaluation methods for fairness in chapter Fairness. Next, we will introduce few algorithms that are used in the group recommender field in chapter Related work. TODO: check out other works and decide what should be here. This can be nice from the reading perspective, but is it really necessary?

2. Recommender systems

In this chapter, we will briefly introduce in general what recommender systems are (hereinafter referred to as RS) and then continue with a description of the group variant of recommender systems and introduce common approaches and methods they employ.

2.1 Recommender systems

Broadly speaking, recommender systems are algorithms that are trying to suggest items to their users, or from another perspective, they aim to predict how would a user rate (like) an unseen item. They are used in a variety of settings, from e-commerce, media consumption, social networks, expert systems, search engines, and many others.

We can generally divide them by their approach as stated in [1] into:

- **Collaborative filtering** (CF)
  Solely based on ratings of items from users (user-item interactions). Trying to recommend unseen items that were liked by users who have a similar taste for other items that they both rated. And thus exploiting data of users with similar preferences.

- **Content-based filtering** (CB)
  Uses item features or item descriptions to recommend items similar to those that the user liked or interacted with. We are essentially building a model of preference for users and exploiting domain knowledge about items that match the users' model.

The popularity of these two approaches varies from domain to domain. Some domains naturally contain item-specific data which allows using the *content-based filtering* for example product parameters in e-shops, but other domains do not. Then it is more beneficial to use *collaborative filtering* techniques or a mix of the two.

There are benefits and drawbacks for both, CF is able to extract latent meaning from the data that would remain inaccessible to CB that relies on items' features. But at the same time, it can cause problems to rely only on user-item interactions because we need a lot of data in order to make a precise recommendation. There will be nothing to recommend if we cannot find similar enough other users that already rated some unseen items. This problem is called a *cold-start problem.*

Some of the classical and more advanced methods include:

- User-based and item-based nearest neighbor similarity [2][3][4]
- Matrix Factorisation techniques[5]
- Deep Collaborative filtering [6][7]
- Deep Content extraction

2.2 Group recommender systems

Introduction

So far we have discussed only recommender systems, where the object of a recommendation is a single user. But what do we do, when we have a group of users that we want to recommend to? For example, a group of friends selecting a movie that they want to watch together or a group listening to music?

Group recommender systems (group RS) are an interesting subarea of recommender systems, where the object of a recommendation is not just a single user but multiple individuals forming a group. Results of a recommendation for the group do need to reflect and balance individuals' preferences among all members.

Challenges

Common approaches

According to [8] there are two main directions we can take:

- **Group aware RS approach**
  Builds a model for the group based on the preferences of all of its members. Either directly by creating a model of preference for the group or by aggregating models of individual users together and then recommending items for the group as a single entity.

- **RS aggregation approach**
  Use single-user RS to recommend to each individual of the group and then aggregate the results together to create the final recommendation for the group.

In the RS aggregation approach we further distinguish between situations where we have predictions for all possible items and therefore can do aggregation directly on the ratings of all items or if only have a list of recommendation for each user - subset of all the items. These two can function very differently, for example taking in context only the position in the recommended list or position and the rating. They are mentioned separately in [8], but the approaches are very similar, they only differ in the availability of provided results from the underlying RS, so we group them together under one main direction.

Further, both group aware RS and aggregation approaches do both have some advantages and disadvantages. One of advantages of the Aggregation approach is that we can use the same RS as we would use for an individual recommendation, either as a black box aggregating directly the top items provided, or in more involved way by utilizing the predicted ratings. On the other hand, the aggregation strategies do rely on single-user RS so there is not much that can be done in order to extract some hidden latent preferences of the group, which in case of the first method, the group aware approach, can potentially be extracted.

We will go in-depth to discuss techniques used in the latest literature in chapter 4.

At the same time, we need to define what does it even mean to recommend something to a group. Do we measure it by fairness, overall user satisfaction, or by the least satisfied member of the group? We will go in depth to describe common approaches to these problems in chapter 3.

3. Fairness

3.1 General

3.2 Long-term fairness

3.3 Evaluation

4. Related work

4.1 Aggregating member preferences

4.2 Aggregation methods

4.3 Direct model methods

5. Datasets

5.1 Main datasets

5.2 Group datasets

5.3 Creating of artificial groups

6. Our work

6.1 EP-FuzzyD'Hondt

7. Offline experiments

7.1 Our work

7.2 Proceedings

8. Application

8.1 Design requirements

8.2 Architecture and design choices

8.3 Cold start problem

8.4 User manual

9. User study

9.1 Methodology

9.2 Results

9.3 Discussion

10. Conclusion

Future work

Bibliography

[1] Francesco Ricci, Lior Rokach, and Bracha Shapira. Introduction to recommender systems handbook. In *Recommender systems handbook*, pages 1–35. Springer, 2011.

[2] Will Hill, Larry Stead, Mark Rosenstein, and George Furnas. Recommending and evaluating choices in a virtual community of use. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 194–201, 1995.

[3] Upendra Shardanand and Pattie Maes. Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 210–217, 1995.

[4] Marko Balabanović and Yoav Shoham. Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40(3):66–72, 1997.

[5] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[6] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pages 173–182, 2017.

[7] Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*, pages 191–198, 2016.

[8] Anthony Jameson and Barry Smyth. Recommendation to groups. In *The adaptive web*, pages 596–627. Springer, 2007.

List of Figures

List of Tables