Impact of Recommender Systems on the Dynamics of Users' Choices

Doctoral Consortium

Naieme Hazrati
Free University of Bolzano
Bolzano
nhazrati@unibz.it

ABSTRACT

The major focus of recommender systems (RSs) research has so far been on improving the precision and quality of the recommendations. However, it is also important to understand whether the recommended items are actually chosen and how they influence users' choice making. Few studies have attempted to analyse the impact of RSs on users' choice making. In this PhD research, we aim at better understanding the impact of RSs on the evolution of the choices made by a collection of users. We propose simulation procedures where users are simulated to make choices over a period of time when they are exposed to alternative RSs. We measure several properties of the users' choices distribution, which capture the RS effect. Our goal is to understand the evolution of the choices of a collection of users as time goes; next choices are influenced by previous choices used by the RS to generate recommendations. Additionally, we propose online experiments to study the effect of RSs on real users' choices. We propose to design web-based platforms where alternative RSs recommend items to the users and study RSs impact by analysing the evolution of the choices.

KEYWORDS

simulation; decision making; recommender systems

ACM Reference Format:

Naieme Hazrati. 2021. Impact of Recommender Systems on the Dynamics of Users' Choices: Doctoral Consortium. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 3 pages.

1 INTRODUCTION

Recommender Systems (RSs) are routinely used to influence users' choices by suggesting items that the user will find interesting and of use [6]. The quality of a RS is typically measured by metrics such as precision of recommendations or Click Through Rate (CTR). Even though RSs are used widely, their impact, which can be studied at the level of an individual user or at the level of a community, has not been fully understood yet. It is therefore important to fill this research gap by identifying the key conditions and system features, such as the specific RS technique, that do affect the users' choices.

A few works that have attempted to understand the impact of RSs have measured aggregated indicators of RSs' impact, i.e., users' choice distribution [3, 5, 7]. They have followed two main approaches: *online experiments* or *simulations*. In online experiments web platforms are designed where online users select items while they are also exposed to recommendations [5]. Users may select the recommended items or not. In simulations artificial agents make choices according to a probabilistic or deterministic choice model and alternative RSs are simulated to make recommendations to these agents [3, 7]. In these works, a RS's effect on the agents (decision makers) is simulated by increasing the likelihood of the recommended item(s) to be chosen. In both approaches, choices, real or simulated, are observed for a time period and then some metrics of collective users' behaviour are measured to capture the RS effect. These approaches have specific limitations. Online experiments have discovered notable results [5], e.g., Matt et. al. [5] discovered that diversity of the choices can be rather different depending on the RS's approach. However, since online experiments require the involvement of real users, it is difficult to assess the RSs effect under different circumstances. However, previously conducted simulations have made simplifying assumptions on the environment and their choice models that reduces the significance of their results [3, 7]. Moreover, some have used synthetic data to generate the user preference model and the alternative choice options [3].

Our research goal is to improve both approaches. First, we propose a more realistic simulation of users' choices under the impact of RSs. A reliable simulation design is a prerequisite to understand the true effect of RSs and inform the decision of which RS to use in a real application. Moreover, it is necessary to improve online experiments as the results obtained so far are not sufficient. In fact, online experiments have assessed the qualities only of a few RSs, and also they have measured only the choice diversity, which is only one of the many dimensions of aggregated users' choice making. We see that it is still necessary to conduct online experiments. Indeed, we aim to improve online experiments to better understand the users' behaviours.

In this PhD research we first focus on improving the significance of the simulation of users' choices under the effect of a RS, to better portray the reality of such human computer interactions. We assess the RSs' effect on users' behaviours by measuring metrics that capture the global effect of RSs, such as Gini index and Shannon entropy, as measures of diversity [2], catalogue coverage of the choices, average quality of the choices, and recommendation acceptance. Moreover, we analyse alternative choice behaviours of users, such as, users that tend to choose popular items, recent items or highly rated items. In this work, we assess the simultaneous impact of alternative choice behaviours and RSs on users' collective choices. Secondly, we focus on conducting online experiments assessing different RSs in terms of their impact. We will design a

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.
© 2021 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

web platform where products are offered to online users. The users will browse recommendations before making choices. The log of this system can give us information on the behaviour of the users.

2 SIMULATING USERS' CHOICES

In order to address the aforementioned goals, in our first work we have proposed a novel simulation procedure of users' choices in the presence of RSs that affect such choices. We have improved previously proposed simulations on the users' iterative choice makings by designing a new simulation [4] that uses real datasets of choices to correctly define the simulation components: user preferences and choice model. We consider several datasets and RSs types in the simulation. In the following we briefly describe this work.

We use four datasets *MovieLens* $100k$ [1] and 3 *Amazon* datasets: *Kindle*, *Apps* and *Games* [2]. We simulate the choices of the users considering the presence of alternative RSs. Our simulation leverages the knowledge of the users' preferences (derived from the dataset) and simulates repeated choices. The choices are made within monthly time intervals according to a probabilistic multinomial choice model [3]. The higher the estimated utility of a user for an item, the higher is the probability that the item will be chosen. Additionally, we assume that users are not aware of the entire catalogue of the items and can only choose items in a set called "Awareness Set" [3]. We build an initial awareness set for each user. Moreover, if an item is recommended, it enters the user's awareness set. Plus, the utility of the recommended item is increased so it's more likely that this item is chosen. We compared five RSs including personalized and non-personalized ones. We train the RSs by using the choices actually present in the dataset, up to a certain date, and incrementally, month by month, by using the simulated users' choices. For each month of simulated choices we calculated the Gini index [2] over the choices made by the users up to that month. Additionally, we consider other metrics such as catalogues coverage, chosen items' average utility, Shannon entropy, popularity of the chosen items and the ratio of choices for the recommended items.

We have discovered that RSs have different impact on users' choice diversity. In fact, personalized RSs lead to a lower Gini index (equivalent to higher diversity) compared to non-personalized ones. Moreover, even among the personalized or non-personalized RSs the effect on Gini index varies with the specific RS approach. Additionally, we discovered that with personalized RSs, the users select items with higher utility, which means that users should be more satisfied with personalized RSs. We observed that users choices popularity differs with the RS. Finally, we discovered that the size of the awareness set of the users is very important to modulate the acceptance of the recommendations.

The obtained results are important because they show that, depending on the system goal, a specific RS can be selected for a target context, e.g., if a company aims at increasing the sales of less popular items, it is possible to apply a personalized RS that has shown to increase the diversity in the simulations.

In a second analysis, we have firstly assessed (data analysis) that users' choices are influenced by distinguished properties of the items, s.a., their popularity or age (time from their first introduction). Then, we have measured global properties of users' choices (e.g., their diversity) when simulated choices are made among recommended items, but also assuming that users are influenced in their choices by the above mentioned properties of the items; for instance, they tend to prefer more recent items. In order words we have tested alternative assumptions related to the choice model. We have found some interesting results showing the importance to analyse the joint effect of the RS and the user choice model on alternative scenarios (dataset). We discovered, for instance, that the diversity of the choices seems not to be influenced by the choice model. But, the choice model and the RS may have a big impact on the average rating and age of the chosen items. Moreover, a non-personalised RS that recommends items with highest average rating may produce, under a choice model that prioritize more recent items, in one dataset (Kindle), choices distributed similarly to the actual choices, while not in another (Movielens).

[1]https://grouplens.org/datasets/movielens/100k
[2]http://jmcauley.ucsd.edu/data/amazon/

3 FUTURE WORK

Firstly, we plan to improve the choice model. So far, the adopted choice model is a multinomial logit choice model. This has been also used in previous studies [1, 3]. However, a true user's behaviour is more complex and diverse, e.g., in the music domain, the choices of a user are strongly dependent to the previous ones, which is not captured by the multinomial logit model very well.

Additionally, it is worth to emphasize that the aim of simulating the users' choices is to predict the future choice behaviours. Therefore, it is necessary to use only the choices observed up to a certain point in time to predict the future ones, i.e., not using any information derived from the knowledge of the future real choices, as we still did (e.g., to determine how many choices a user is doing in a simulated month). We should instead be able to fully predict the number and the type of choices of the users will do in the future. Applying the simulation in different domains is another critical facet in this PhD research. Different domains have different characteristics, e.g., in the music domain the users may listen to a song multiple times, while in the book domain, it is less likely that they read a book even twice.

Finally, we aim at conducting online experiments studying the choice behaviours of users in the presence of a RS. We plan to build a platform where the users are asked to rate/like items and then a range of RSs are used to generate relevant suggestions for them. Thereafter, we will ask users to select other items while we offer recommendations to them. We will also ask users to rate the chosen items after consuming them. The choices and the feedback can help us understand the RSs and their impact more thoroughly.

REFERENCES

[1] William A Brock and Steven N Durlauf. 2002. A multinomial-choice model of neighborhood effects. *American Economic Review* 92, 2 (2002), 298–303.
[2] Robert Dorfman. 1979. A formula for the Gini coefficient. *The review of economics and statistics* (1979), 146–149.
[3] Daniel Fleder and Kartik Hosanagar. 2009. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science* 55, 5 (2009), 697–712.
[4] Naieme Hazrati, Mehdi Elahi, and Francesco Ricci. 2020. Simulating the Impact of Recommender Systems on the Evolution of Collective Users' Choices. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media*. 207–212.

[5] Christian Matt, Thomas Hess, and Christian Weiß. 2013. The differences between recommender technologies in their impact on sales diversity. *aisel.aisnet.org* (2013).

[6] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2015. Recommender systems: introduction and challenges. In *Recommender systems handbook*. Springer, 1–34.

[7] Zoltán Szlávik, Wojtek Kowalczyk, and Martijn Schut. 2011. Diversity measurement of recommender systems under different user choice models. In *Fifth International AAAI Conference on Weblogs and Social Media*.