Fairness In Reciprocal Recommendations: A Speed-Dating Study

Yong Zheng
Illinois Institute of Technology
Chicago, Illinois, USA
yong.zheng@iit.edu

Tanaya Dave
Illinois Institute of Technology
Chicago, Illinois, USA
tdave@hawk.iit.edu

Neha Mishra
Illinois Institute of Technology
Chicago, Illinois, USA
nmishra8@hawk.iit.edu

Harshit Kumar
Illinois Institute of Technology
Chicago, Illinois, USA
fharshitkumar@hawk.iit.edu

ABSTRACT

Traditional recommender systems suggest items by learning from user preferences, but ignore other stakeholders in the whole system. Actually, not only the receiver of the recommendations, but also other stakeholders may come into play, such as the producers of items or those of the system owners. Reciprocal recommender system in dating or job recommendations is one of these examples. However, we may have to simulate the utilities for each type of the stakeholder due to the utility definitions. In this paper, we perform exploratory analysis on a speed-dating data, where the user expectations are clearly defined. We try to build a multi-dimensional utility framework by utilizing multi-criteria ratings. We further analyze the relationship between the utilities and recommendation performance, and achieve a tradeoff as the optimal solution. Even more, the proposed approach is able to outperform the exiting reciprocal recommendation algorithms in precision, recall and overall utilities. Finally, we derive a promising way to define and optimize utilities to be generalized in other applications or domains.

CCS CONCEPTS

• **Information systems** → **Recommender systems**;

KEYWORDS

recommender systems; reciprocal recommendation; multi-stakeholder; utility; optimization; multi-criteria

ACM Reference Format:
Yong Zheng, Tanaya Dave, Neha Mishra, and Harshit Kumar. 2018. Fairness In Reciprocal Recommendations: A Speed-Dating Study. In *UMAP'18 Adjunct: 26th Conference on User Modeling, Adaptation and Personalization Adjunct, July 8–11, 2018, Singapore, Singapore.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3213586.3226207

UMAP'18 Adjunct, July 8–11, 2018, Singapore, Singapore

ACM ISBN 978-1-4503-5784-5/18/07…$15.00
https://doi.org/10.1145/3213586.3226207

1 INTRODUCTION

Recommender systems is a well-known solution to assist decision making by suggesting items to the users. A traditional recommender system may produce a list of recommendations to match the user preferences in the history. For example, a learning-based recommendation algorithms may be developed to minimize the prediction errors or maximize the ranking of the top-N recommendations. These optimizations only takes the utility of the end users (i.e., users' explicit or implicit feedbacks) into account.

However, the receiver of the recommendations may not be the only stakeholder in the system, and other stakeholders may come into play. There are several examples:

- In the *dating* application [17], a young man may prefer a recommended woman who also wants to date with him, rather than only the pool of ladies that the young man likes.
- The similar thing may happen in the *job seeking* or *recruitment* case. We may need to recommend some job positions to a user while the corresponding employees may be willing to hire him or her.
- In the advertising area [18], not only the user preferences (e.g., interests on Ads, privacy concerns) but also the interests of the advertisers should be considered. For example, the advertiser would like to present the car advertisement to the appropriate customers, rather than any groups of the viewers who like cars. Teenagers may like cars too but they may not have the capability to make purchases, which decreases the utility of the advertisers in this example.
- The case of multi-stakeholders may also happen in the educational learning area, e.g., a successful recommendation of the projects may need to consider both the student preferences and the instructor expectations or requirements [28].

In this paper, we particularly focus on the reciprocal recommendations [16, 17, 20]. Reciprocal recommender systems have been applied in multiple domains, such as dating [16], recruitment [19, 22] and social connections [10]. However, most of the work may have to simulate the stakeholder values due to the difficulty of utility definitions or collections. For example, in the dating applications [16, 20], the peer utility may come from the match in terms of user demographic information, the number of dates, the frequency of communications, etc. In this paper, we perform exploratory analysis on a speed-dating data, where the user expectations and dating evaluations are clearly defined. We try to build a multi-dimensional utility framework by utilizing multi-criteria ratings, and demonstrate that we are able to obtain a successful tradeoff between the utility optimizations and the recommendation performance.

2 RELATED WORK

In this section, we introduce the related work in fairness and the multi-stakeholder recommender systems. And we particularly discuss the research on the reciprocal recommender systems, especially the dating applications.

The importance of "fairness" has been raised recently. Algorithmic fairness [6, 8, 23] could be defined as a type of constrained optimization which is designed to improve both the efficiency and equity of decisions. It is necessary because our machine learning algorithms with the standard utility-maximization objectives (e.g, maximizing prediction accuracy or minimizing prediction errors) sometimes resulted in algorithms that behaved in a way in which a human observer will deem unfair, often especially towards a certain minority, such as the cases with respect to user gender, race, disabilities, and so forth.

Burke [2] extended this concept to the area of recommender systems. Basically, there could be two ways to incorporate fairness into recommender systems. One is a similar way as the original definition of algorithmic fairness, where we try to balance the equity in the algorithm components with respect to minority groups which may be defined by specific attributes, such as user gender or race. For example, by taking gender and race into consideration, Burke, et al. [3] proposed a way to balance the neighborhood selection in the recommendation algorithms. Another idea is recommending items by considering multi-stakeholders roles in the system.

The similar idea of "multi-stakeholder" has been discussed in the reciprocal recommendations, but the stakeholders are not limited to the "bilateral" interactions. For example, in a movie applications, the potential stakeholders involved in the process could be the receiver of movie recommendations, the producer of the movies, the cinema which plays the movies, and even the movie director. The major challenge in the multi-stakeholder recommender systems is the utility definitions and optimizations. Unfortunately, there are no unified ways to define the utility of stakeholders. We may simulate the utilities based on the interactions between the items and the stakeholders [4], or utilize the existing preference data, such as multi-criteria ratings [25]. In this paper, we will show case that how multi-criteria ratings can be helpful to define utilities. From the perspective of utility optimizations, it is highly correlated to the multi-objective learning [12] which is a popular solution to deal with such optimizations. The existing methods either incorporate the fairness into the recommendation models to re-rank the items [3], or directly setup a multi-objective learning process [15] to obtain the optimal solution. The ultimate goal is to produce a list of appropriate item recommendations which consider the values of multiple stakeholders with no or reasonable loss in the recommendation accuracy.

In this paper, we particularly focus on the reciprocal recommendations [16, 17, 20] in which the "two-sided" or the "bilateral" interactions between the peers are considered. Take the dating applications [16, 20] example, the recommended females to a male user cannot simply be the suggestions who match the male user's preferences. A successful dating recommendation may have to also consider whether the females' opinions. Therefore, the values of the peers (i.e., the two stakeholders) should be fair enough in the recommendation process. The corresponding utilities are usually derived from simulations. For example, the RECON [16] approach considers the match in user demographic information and the communications between the peers. Xia, et al. [20] additionally take the dating interactions by the user neighborhood into considerations. These information may indicate how a user likes others, but it may not be a good representation as user utility.

3 DATA AND PROBLEM STATEMENT

In this section, we discuss the speed-dating data and introduce the list of problems we would like to explore.

3.1 The Speed-Dating Data

The speed-dating data was collected and introduced by Fisman, et al. [9]. It is also available on Kaggle.com[1]. The data was gathered from participants in experimental speed dating events from the year of 2002 to 2004. During the events, the attendees would have a four minute "first date" with every other participant of the opposite gender. At the end of their four minutes, participants were asked if they would like to see their date again. They were also asked to rate their date on six attributes: attractiveness, sincerity, intelligence, fun, ambition, and shared interests.

The data contains rich information. We list the useful information that we may need in our work as follows:

- **Demographic Information**: Each subject was requested to fill in a questionnaire in which they need to provide a list of demographic information, such as age, gender, country, field of studies, degree, race, income level, and so forth.
- **Dating Goals or Expectations**: In the questionnaire, each subject defined their dating goals or expectations in advance. These information were stored in either textual options (e.g., the primary goal of looking for a partner, etc) or multi-criteria ratings. For simplicity, we utilize the multi-criteria ratings only. More specifically, the subjects need to describe their expectations in six criteria: attractiveness, sincerity, intelligence, fun, ambition and shared interests. The summation of the ratings should be 100 in total.
- **Dating Experience**: Each subject will represent their dating experience in multiple ways. First of all, each subject will leave the ratings on the same six criteria mentioned above to indicate how the partner meets their expectations. In addition, each subject also gave an overall rating (in scale 1 to 10) to the partner. Finally, each subject needs to present the willing which tells whether they want to have a further date with the same partner in the future or not. The binary variable, "match", indicates whether both of them would like to keep the relationship or continue the dating in the future.

There are 392 subjects in total. The number of dates by each subject is at least 9 and at most 21. The whole data contains 8,378 records which describe the information for each date. Note that

[1] https://www.kaggle.com/annavictoria/speed-dating-experiment

a date between user a and user b will be described by two rows or records in the data. Only 16.5% of the records present matched dating experience.

3.2 Problem Statement

Take suggesting user b to user a for example, we define user a as the *target user*, while b will be the recommended *partner*. Our major task is to recommend a list of appropriate partner users to a target user. Particularly, the appropriateness is defined by the *matched* dating experience which can be indicated by the binary variable "match". However, there are also other variables related to the dating experience, which leaves a challenge to define the utilities of each stakeholder. In this paper, we specifically focus on the following research problems:

- how the multi-criteria ratings can be utilized to define the utility framework.
- whether the proposed multi-stakeholder approach can outperform the single-stakeholder approaches.
- whether the proposed multi-stakeholder approach can outperform the existing reciprocal recommendation methods.
- whether the utility optimizations may hurt the recommendation performance.
- how to find the trade-off between utilities and recommendation accuracies.

4 MULTI-STAKEHOLDER APPROACHES

As mentioned previously, one of the major challenges in the multi-stakeholder recommender systems is the utility definition. In this speed-dating example, we have different kinds of the user preferences on the dating experience, such as the overall rating and multi-criteria ratings on their partner, as well as the match result. Our goal is to recommend partners that the two users would like to have further dates, i.e., the match value will be one.

Simply, we believe the multi-criteria ratings on the six criteria (i.e., attractiveness, sincerity, intelligence, fun, Ambition and shared interests) can be used to represent user satisfactions. More specifically, assume the target user a is going to date with user b. User a's expectations can be represented by such a rating vector based on the multi-criteria ratings, $\overrightarrow{E_a}$. And the dating experience can also be described by a similar rating vector, $\overrightarrow{R_{a,b}}$. The utility for the target user a to date with user b can be indicated by the similarity based on these two rating vectors as depicted by Equation 1. Accordingly, the utility for the partner b to date with a can be described by Equation 2. Note that $\overrightarrow{R_{a,b}}$ and $\overrightarrow{R_{b,a}}$ denote the rating vectors based on the multi-criteria given by a to b, and b to a respectively.

$$Utility(a) = similarity(\overrightarrow{E_a}, \overrightarrow{R_{a,b}}) \tag{1}$$

$$Utility(b) = similarity(\overrightarrow{E_b}, \overrightarrow{R_{b,a}}) \tag{2}$$

The similarity between two vectors can be calculated by the popular similarity or distance measures, such as cosine similarity, Pearson correlation or Euclidean distance. We tried these three measures and found that Pearson correlation can help us obtain the best results. We will report the results based on the Pearson correlation in the following sections.

By taking multi-stakeholders into consideration, we expect our approach can optimize the utilities for both user a and b. It enables them to date together after the four-minute speed dating, which results in a 1 value in the variable "match". The process can be viewed as a multi-objective learning. There are usually three methods [21] to address these multi-objective learning problems: 1). trade-off weight method which transforms the multi-objective problem to a single-objective one by assigning a set of trade-off weights; 2). constraint-based method which optimizes the most important objective and utilize other objectives as constraints by setting different thresholds; 3). learning based optimizations which use specific optimizations (such as evolution or genetic algorithms) to find the optimal solutions.

In this paper, we would like to examine how the recommendation performance (such as precision, recall) change by different utilities scores. We prefer to employ the trade-off weight method. More specifically, we calculate a utility score for the dating experience between user a and b, as described by Equation 3:

$$Score(a \leftarrow b) = \beta \times Utility(a) + (1 - \beta) \times Utility(b) \tag{3}$$

The utility score will be produced for each candidate. Equation 3 shows that the system estimates the utility score if it is going to recommend b to the user a. We assign a weight β ($0 < \beta < 1$) to the utility from the perspective of user a, and leave weight 1-β to the utility from the perspective of user b. We are able to learn the best option for β by a learning-based optimization. In this paper, we simply try an exhaustive search to vary the value of β, in order to track how the recommendation performance change if we assign different weights to the utility components.

The workflow of our approach can be described as follows:

- To recommend a list of partners to user a, we figure out a list of candidates in the opposite gender. Assume user b is one of the candidates. To obtain the utility score in Equation 3, we extract the expectation rating vector $\overrightarrow{E_a}$ and $\overrightarrow{E_b}$ from the knowledge base. Furthermore, we need to predict multi-criteria ratings to estimate how a will rate b and how b will rate a in the six criteria. We apply the biased matrix factorization [13] to each rating matrix based on the ratings from each criterion. The rating on each criterion is predicted individually and independently. Criteria chains [24] is one multi-criteria prediction method which considers correlations among multiple criteria, and we will explore it in our future work.
- We rank the candidates by the utility score defined in Equation 3 to obtain the top-N recommendations. These recommendations will be evaluated based on the "match" information in the test set. Namely, only the candidates who like each other with the target user will be considered as appropriate recommendations.
- We vary the value of β from 0 to 1 with 0.1 increment in each step to define different scoring functions, and track how the recommendation performance change accordingly.

Note that, we also have user a's overall rating on the user b. We can add this rating as additional dimension in the multi-criteria

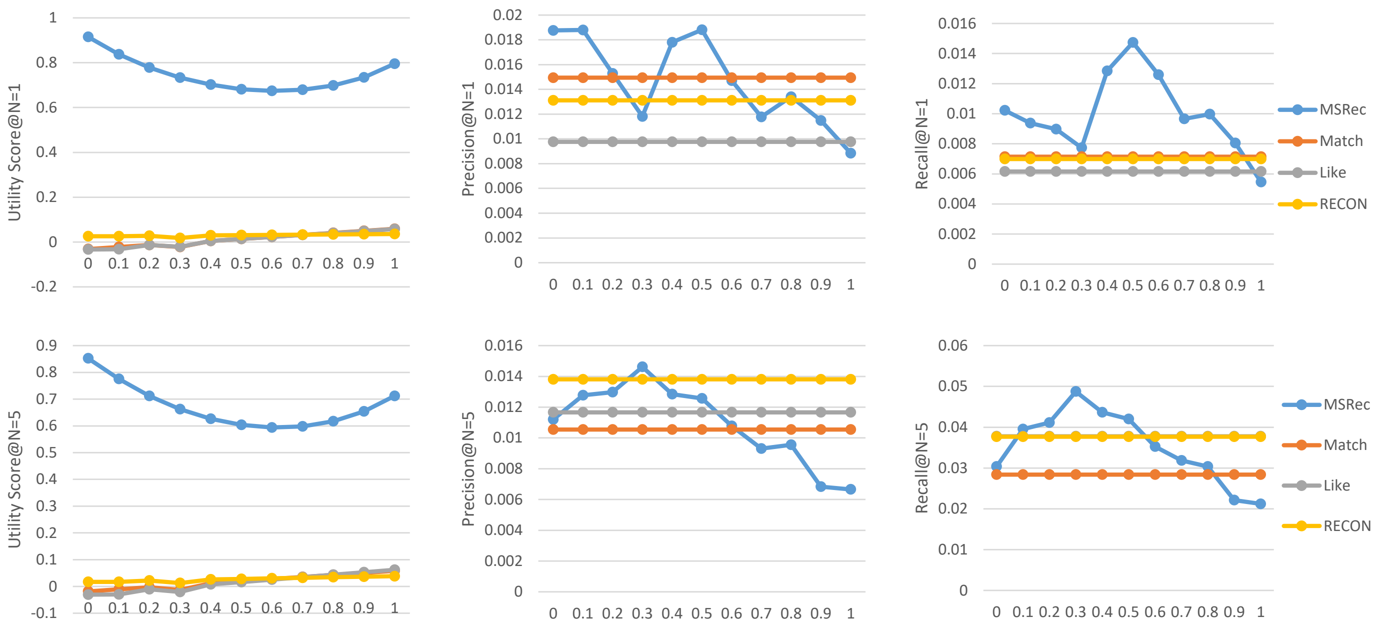

Figure 1: Experimental Results

rating vector. We ignore this operation in this work and we may explore how to better utilize this overall rating in the future.

5 EXPERIMENTAL ANALYSIS

5.1 Baseline Approaches

We design three baseline approaches in our work:

- The model *Match* is the biased matrix factorization model by using the "match" variable as user preferences, where we only consider user, item and match results to build the recommendation model.
- The model *Like* is another biased matrix factorization model while we utilize the overall rating as user preferences.
- The model *RECON* [16] is one reciprocal recommendation algorithm for online dating. It utilizes the user demographic information and communication history. We use user age, gender, race and dating purpose (i.e., a fun, a date, a serious relationship or just to meet new people) to build user attributes. The user preferences were build based on the existing dating experience in the training set, such as how many males or females a user dated, how many partners in different age groups, the distribution of races and different dating purposes, etc. Unlike the dating application in [16], users cannot send messages to others. Alternatively, we simply use a true or false value to indicated whether a user has dated with another one in the training set.

Note that the *Match* and *Like* approaches just rank the items by minimizing the sum of squared errors in the overall rating or the "match" variable, without considering multi-stakeholder utilities. The *RECON* method relies on a compatibility score which is similar to multi-stakeholder utilities. But this compatibility score is calculated based on the user attributes and preferences. Our approach particularly utilizes the multi-criteria rating vectors to represent the user expectations and examinations on the dating experience.

5.2 Evaluation Protocols

We have three goals in our work: 1). optimize utility scores to produce recommendations; 2). find the optimal β, and examine how the recommendation performance change by different utility scores; 3). compare proposed method with baseline approaches to find the optimal model. We use the 5-fold cross validation and evaluate the models based on top-1 and top-5 recommendations. The values in the variable "match" are considered as the ground truth in the test set. We use precision and recall as the corresponding evaluation metrics to measure the recommendation performance.

5.3 Results and Findings

Figure 1 presents our experimental results. The x-axis indicates the values of β which is the weight in Equation 3. The figures on the first row shown in Figure 1 present the results in utility scores, precision and recall of top-1 recommendations, while the 2nd row shows the outcomes for top-5 recommendations. We use "MSRec" to represent the multi-stakeholder approach proposed in this paper. Note that the three baseline models do not rank items based on the utility score defined in Equation 3. Therefore, the precision and recall associated with these three approaches do not change by varying the values of β. For each recommended partner, we calculate the utility score based on the Equation 3. The utility score in the Figure 1 is the average utility score based on the top-N recommendations.

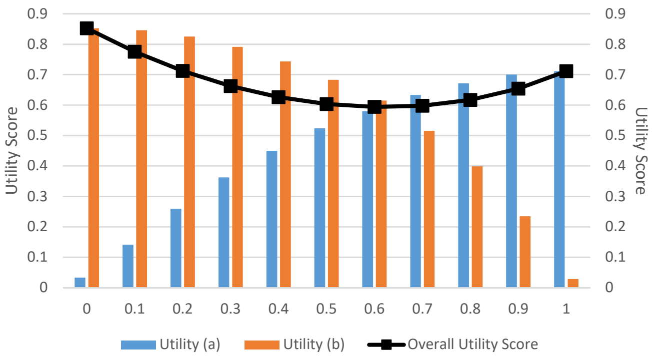

Figure 2: Individual Utility vs. Overall Utility

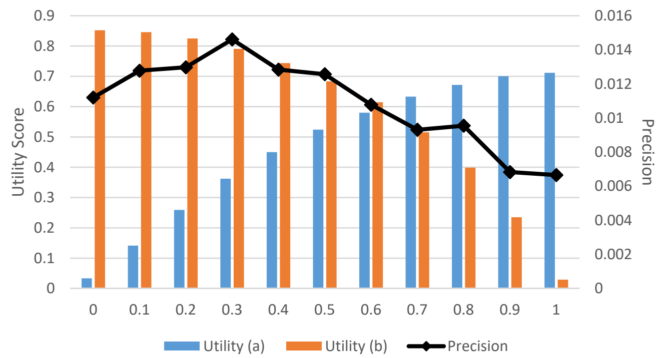

Figure 3: Utility Score vs. Precision

5.3.1 *General Patterns.*

Assume that we may recommender user b to date with the target user a, we only consider the utility from the perspective of a if β is set as 1. Accordingly, we only focus on the desire of user b, if we set β as zero. In terms of the precision and recall results, we can observe that the top-N recommendation performance may generally go down if the β value is larger than 0.5. Note that a *matched* dating experience may require the user satisfaction from the perspective of both user a and b. It is also confirmed from our results. The precision and recall results may not be the best ones if we simply set β as zero or one in the top-5 recommendations.

The figures which plot the utility scores demonstrate that MSRec is able to offer significant higher utility scores, no matter which β values we choose. The RECON method relies on a compatibility score which is calculated based on the user attributes and previous dating experience. The result in our work may infer that the compatibility score in the RECON approach may not be significant correlated with the utility score defined in our work.

5.3.2 *Solution by Optimizing Recommendations.*

By seeking the optimal solution based on the precision and recall results only, we find out that 0.5 and 0.3 are the best β value for top-1 and top-5 recommendations respectively. The corresponding precision and recall can beat all the baseline approaches. Particularly, the best setting in our proposed approach is able to significantly outperform the RECON approach in both top-1 and top-5 recommendations in terms of precision and recall. Note that the precision and recall values are relatively small in our experiment, due to that fact that only 16.5% of the records in the data have *matched* dating experience.

However, this optimal solution may decrease the utilities. Figure 1 shows that the overall utility score in the top-1 recommendation is almost the smallest one if we set β as 0.5. The situation in the top-5 recommendations is better, but the overall utility is decreased by 20% from the top if we set β as 0.3. We can also observe that the precision and recall results are generally higher when the β value is smaller. It infers that the "match" result may rely more on the perspective of the parter instead of the target user.

5.3.3 *Solution by Optimizing Utilities.*

Figure 2 presents the comparison between individual utilities and the overall utility in the top-5 recommendations by using different β values shown as the x-axis. Setting β as zero can help obtain the best overall utility, but the utility of the target user a is minimized. The individual utilities for both user a and b could be balanced by setting β as 0.5 or 0.6, but the overall utility becomes the minimal one in this case.

Therefore, the best setting to optimize the utilities could be more complicated. A trade-off value is required to balance the overall utility and the individual utilities. One possible solution is that we can setup a threshold of the minimal utility. Namely, we may need to find the optimal solutions that enables a descent overall utility score but also achieves the minimal utility for each stakeholder. If this threshold is set as 0.3, it tells that the results based on corresponding settings of β ($0.3 \leq \beta \leq 0.8$) are qualified. Particularly, we can achieve the best overall utility score by setting β as 0.3, though there is not big difference if we set β as 0.4 or 0.8.

5.3.4 *Optimal Trade-Off Settings.*

Out ultimate goals is to recommend partners while we consider the utility of the peers at the cost of recommendation performance. Figure 3 presents the trend of precision at top-5 recommendations associated with different individual utilities. By setting β as 0.3, we can achieve the best precision but the overall utility, and specially the utility of a are decreased. 0.4 or 0.5 could be a better choice, where precision is dropped by 12%.

Based on the results above, we can observe that not only the balance between utilities and the recommendation accuracy, but also the balance between overall utility and individual utilities should be considered to find the optimal solution. Our suggestion is that we need to set up the degree of tolerance for these elements, including recommendation performance (e.g., precision, recall), individual utilities and the overall utility. For example, the precision or the utility cannot be hurt by a specific pre-defined percentage as the constraints in the optimization process.

6 CONCLUSIONS, DISCUSSIONS AND FUTURE WORK

In this paper, we conduct our initial work on the topic of multi-stakeholder recommender systems by using a speed-dating data as the case study. Particularly, we define the utility as the similarity between user expectation and examination on the dating experience. And we propose to recommend the partners to the target user by ranking a utility score which combines the preferences from the perspective of the peers. Our experimental results reveal that our approach is able to outperform the baseline approaches in both the utility and recommendation optimizations. We also discover that not only the balance between utilities and the recommendation accuracy, but also the balance between overall utility and individual utilities should be considered to find the optimal solution.

There are plenty of lessons learned from our work and there could be potentially extensions we can product to generalize the idea to other applications or domains. We discuss these promising topics and future work in the following sections.

6.1 Dating Applications

We use a speed-dating case study in our work, but the general dating applications could be more complicated in the real-world. For example, there could be VIP users or the users with the paid membership [11] in the dating applications. The priority of these target users should be given, while we may optimize the utility from the perspective of these VIP users to find the optimal solution.

6.2 Optimization Methods

As mentioned in Section 4, there could be three methods to address the multi-objective learning problem. Our initial work in this paper utilizes the exhaustive search to iterative the value of weights in the utility function. To speed up the learning process and reduce complexity of human efforts, we could employ other more efficient optimization approaches in the future, such as the genetic or evolution algorithms [1, 7].

6.3 Conflicting Interests

The optimizations could be more complicated if there are conflicting interests [5, 14]. Take an educational learning case [26–28] for example, students may prefer to select an easy topic for their final project in the class, but they may not be able to get a high score since it does not meet the instructor's expectations. There are plenty of solutions by using the genetic algorithms, but there will be more constraints or trade-off processes in the learning process.

6.4 Generalized Applications

The major advantage in this speed-dating data is that the survey collects the user expectations and evaluations on the dating experience by ratings on six criteria – attractiveness, sincerity, intelligence, fun, ambition and shared interests. Though we may not always have user expectations, the idea in this paper can be easily extended to other applications, especially the ones with multi-criteria ratings. More specifically, in the applications with multi-criteria ratings, such as TripAdvisor[2] and OpenTable[3], we can learn user expectations in shape of the multi-criteria ratings, and finally use the similarity between the expectations and the examinations as the utility function in the optimization process.

[2]https://www.tripadvisor.com/

[3]https://www.opentable.com/

REFERENCES

[1] Ajith Abraham and Lakhmi Jain. 2005. Evolutionary multiobjective optimization. In *Evolutionary Multiobjective Optimization*. Springer, 1–6.

[2] Robin Burke. 2017. Multisided Fairness for Recommendation. *arXiv preprint arXiv:1707.00093* (2017).

[3] Robin Burke, Nasim Sonboli, and Aldo Ordonez-Gauger. 2018. Balanced Neighborhoods for Multi-sided Fairness in Recommendation. In *Conference on Fairness, Accountability and Transparency*. 202–214.

[4] Robin D Burke, Himan Abdollahpouri, Bamshad Mobasher, and Trinadh Gupta. 2016. Towards Multi-Stakeholder Utility Evaluation of Recommender Systems. In *UMAP (Extended Proceedings)*.

[5] Cheng Chuntian and Kwok-Wing Chau. 2002. Three-person multi-objective conflict decision in reservoir flood control. *European Journal of Operational Research* 142, 3 (2002), 625–631.

[6] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 797–806.

[7] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation* 6, 2 (2002), 182–197.

[8] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*. ACM, 214–226.

[9] Raymond Fisman, Sheena S Iyengar, Emir Kamenica, and Itamar Simonson. 2006. Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics* 121, 2 (2006), 673–697.

[10] Ido Guy. 2015. Social recommender systems. In *Recommender Systems Handbook*. Springer, 511–543.

[11] Conor Hayes and Erik Aumayr. 2017. The path to success: A study of user behaviour and success criteria in online communities. In *Proceedings of the 2017 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*. ACM.

[12] Yaochu Jin. 2006. *Multi-objective machine learning*. Vol. 16. Springer Science & Business Media.

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

[14] Ida Kristoffersson, Leonid Engelson, and Maria Börjesson. 2017. Efficiency vs equity: Conflicting objectives of congestion charges. *Transport Policy* 60 (2017), 99–107.

[15] Phong Nguyen, John Dines, and Jan Krasnodebski. 2017. A Multi-Objective Learning to re-Rank Approach to Optimize Online Marketplaces for Multiple Stakeholders. *arXiv preprint arXiv:1708.00651* (2017).

[16] Luiz Pizzato, Tomek Rej, Thomas Chung, Irena Koprinska, and Judy Kay. 2010. RECON: a reciprocal recommender for online dating. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 207–214.

[17] Luiz Pizzato, Tomek Rej, Thomas Chung, Kalina Yacef, Irena Koprinska, and Judy Kay. 2010. Reciprocal recommenders. In *8th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems, UMAP*.

[18] Michael J Polonsky and Michael R Hyman. 2007. A multiple stakeholder perspective on responsibility in advertising. *Journal of Advertising* 36, 2 (2007), 5–13.

[19] Hong Wenxing, Chen Yiwei, Qi Jianwei, and Hui Yin. 2015. iHR+: A mobile reciprocal job recommender system. In *Computer Science & Education (ICCSE), 2015 10th International Conference on*. IEEE, 492–495.

[20] Peng Xia, Benyuan Liu, Yizhou Sun, and Cindy Chen. 2015. Reciprocal recommendation system for online dating. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. ACM, 234–241.

[21] Lihong Xu, Qingsong Hu, Haigen Hu, and Erik Goodman. 2010. Conflicting multi-objective compatible optimization control. In *New Achievements in Evolutionary Computation*. InTech.

[22] Murat Yagci and Fikret Gurgen. 2017. A Ranker Ensemble for Multi-objective Job Recommendation in an Item Cold Start Setting. In *Proceedings of the Recommender Systems Challenge 2017*. ACM, 2.

[23] Rich Zemel, Yu Wu, Kevin Swersky, Toni Pitassi, and Cynthia Dwork. 2013. Learning fair representations. In *International Conference on Machine Learning*. 325–333.

[24] Yong Zheng. 2017. Criteria Chains: A Novel Multi-Criteria Recommendation Approach. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. ACM, 29–33.

[25] Yong Zheng. 2017. Multi-Stakeholder Recommendation: Applications and Challenges. *arXiv preprint arXiv:1707.08913* (2017).

[26] Yong Zheng. 2018. Exploring User Roles In Group Recommendations: A Learning Approach. In *Adjunct Proceedings of the ACM Conference on User Modelling, Adaptation and Personalization: 3rd Workshop on Human Aspects in Adaptive and Personalized Interactive Environments*.

[27] Yong Zheng. 2018. Identifying Dominators and Followers In Group Decision Making Based on The Personality Traits. In *Companion Proceedings of the 23rd International on Intelligent User Interfaces: 2nd Workshop on Theory-Informed User Modeling for Tailoring and Personalizing Interfaces*.

[28] Yong Zheng. 2018. Personality-Aware Decision Making In Educational Learning. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. ACM, 58.