A Novel Recommendation Algorithm Based on Multi-objective Evolution

Jiaqi Zhang[1]

[1]Anhui University, Anhui 230000, China

e01814009@stu.ahu.edu.cn

Abstract. During the past two decades, recommender systems have been playing an increasingly significant role in business decisions for providing personalized recommendations. However, the majority of the present algorithms attach excess importance to accuracy while overlooking users' demand for diverse items. We have proposed a novel multi-objective evolutionary recommendation algorithm, which has the capability to look for compromise among accuracy and diversity. Moreover, due to the fact that Pareto optimal solutions of recommendation tend to be sparse, the proposed algorithm implements a new strategy of population initialization with the consideration of the sparse nature of the Pareto solutions. The proposed algorithm will take effect in the recommender systems.

Keywords: Evolutionary algorithm, Multi-objective problems, Recommender systems

1. Introduction

With the quick improvement of science and innovation, the blast of information has become progressively escalated [1]. However, in this information explosion era, users often encountered difficulties in finding the desired content due to the great volume of data [2]. Recommender systems (RSs) have proved to be an efficient tool to cope with the problem [3]. The main idea of RSs is to establish a model to automatically recommend the proper items to users based on the historical behavior and preference information [4]. RSs can assist users in decision making including movies, books, web search, tourism, and so on [5]. The existing RSs can be arranged into a few classes: hybrid recommendation, knowledge-based recommendation, collaborative filtering recommendation, and content-based recommendation [6].

Normally, the major aim of RSs is to satisfy the user by the best choice of recommendations, which can be interpreted as the maximum of accuracy [7]. Nonetheless, being accurate is not enough to guarantee the desired recommendation for users [8]. The other metric of performance, diversity, ought to likewise be taken into consideration to satisfy the multiple demands of users [9]. In practical application, a diversity-accuracy dilemma has been found, where the recommendation of manifold items contributes to a reduction of exactness for users and vice versa [10]. To accomplish an appropriate harmony among accuracy and diversity, we adopt a multi-objective recommendation model to optimize the conflicting objectives of diversity and accuracy. Thus, naturally, the problem of personalized recommendation is modeled as a multi-objective problem (MOP) [11].

There have been a lot of researches on recommendation based on multi-objective optimization. Ribeiro et al. exploited the Pareto-efficiency concept and select Pareto-efficient items to optimize the

recommender systems by the means of weighted combination [12]. A model of personalized recommendation was proposed by Zuo et al., which optimizes the accuracy and coverage of the recommendation lists [13]. Agarwal et al. used a constrained optimization framework to naturally incorporate various application-driven requirements [14]. Rodriguez et al. framed an approach on optimizing non-smooth rank metrics for information retrieval.

To strike a better balance between accuracy and diversity, we furthered the research on the above recommendation techniques. In this paper, a novel multi-objective recommendation algorithm is proposed to generate diverse recommendations while reserving the competent accuracy of recommendation accuracy. The algorithm adopts bipartite network projection (Probs) as a recommendation technique to estimate accuracy, which is basic yet effective. Accordingly, the algorithm utilizes Coverage to measure the ability of diversity. Moreover, by encoding recommendations to different users in one individual, it is easy to obtain multiple recommendations for multiple users in one run. Via the technique of clustering, the numerous users can be separated into distinctive clusters to decrease the cost of computation. The main contributions can be given as follows: (1)A novel variation of Probs is implemented to make ratings. (2)A heuristic algorithm is adopted to initialize the population. (3)A new strategy of crossover based on the similar genes of parents is used to generate a child.

2. Technological background
In this section, the definition of recommendation problem and multi-objective optimization is given, respectively, and then the ProbS method is reviewed in detail.

2.1. Recommendation problem
In general, we can formulate the problem of recommendation as follows. Assuming that Users includes the whole users of a system, and Items includes all the items to be recommended. Thus, a rating matrix R can be utilized to estimate the preference of users to items. If user i $\in$ Users has priority to the item j $\in$ items, the value of R(i, j) can be equal to 1, and R(i, j) is regularly a real number or non-negative integer inside a specific reach [15]. It is obvious that only a handful of items are evaluated by a handful of users. Therefore, the rating matrix R turns into a sparse matrix. The process of recommendation can be presented briefly. Firstly, on the basis of matrix R, the suggested algorithm predicts the unknown ratings in R. Consequently, one or a set of items will be provided for the user, this can be expressed as:

$$\forall\, i \in \text{Users},\ j = \arg\max R(i, j),\ j \in \text{Items} \qquad (1)$$

2.2. Multi-objective optimization
The multi-objective optimization can be mathematically defined as:

$$min\ F(\boldsymbol{x}) = \left(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_m(\boldsymbol{x})\right)^T \qquad (2)$$

where $\boldsymbol{x} = [x_1, x_2, ..., x_d] \in \Omega$ is a decision vector, and $F(\boldsymbol{x})$ is an objective function of m dimensions. In view of a pair decision vectors $\boldsymbol{x}_A \in \Omega$ and $\boldsymbol{x}_B \in \Omega$, if the equation is satisfied, it is said that $\boldsymbol{x}_B$ is dominated by $\boldsymbol{x}_A$, which can be given in the form of $\boldsymbol{x}_A \succ \boldsymbol{x}_B$. Consisting of decision variables, the pareto set refers to the solutions which are not dominated by any other solutions [16].

2.3. ProbS
By the means of ProbS, the basic element in rating matrix $R(i, j)$ is equal to 1 or 0, representing that user $i \in Users$ selected or not selected the item $j \in$ items. Although it may sacrifice some information, the adoption of explicit ratings can obtain this form with ease [17]. The overall procedure of Probs algorithm can be separated into three steps. The first one is the contribution of user-item bipartite network, which is determined by the relationship of users and items [18]. The second one is

distribution of the initial resource which are allocated to each item and then is distributed to adjacent users. The last step is that apportion users' resources back to the neighboring items. Ultimately, a matrix of nominalized transition can be formulated. The specific computation of allocation can be expressed as follows:

$$w_{mn} = \frac{1}{k_n} \sum_{i=1}^{M} \frac{r_{im}\gamma_{in}}{k_i}$$ （3）

where w_{mn} represents the fraction of the transition of initial resource from n to m, M is the summation of the users, and k_n represents the degree of node n. The ratings of items can be evaluated as follows:

$$f'_m = \sum_{n=1}^{N} w_{mn}f_n$$ （4）

where the number of items is amount to N.

3. Methodology
In order to keep a balance between accuracy and diversity of the recommendation problem, a state-of-the-art algorithm is proposed, which is named MOEA-NProbS. In this part, the MOEA-NProbS algorithm will be interpreted in detail, including a new strategy to initialize the population, a novel algorithm for recommendation and method to make crossover.

3.1. A novel probs (NProbS)
Although Probs algorithm is confirmed to be effective by Zou, there are two aspects to be improved. Firstly, the neglection of items below score 3 sacrifice some information. Secondly, the average allocation of resources narrows the gap between likes and dislikes. As a consequence, NProbS is proposed to handle the above problems. The specific steps of NProbS are as follows: (1) Construct the user-item bipartite network. If the user A evaluated the item, then set the resource of item as 1, other items as 0. (2) Distribute the initial resources of items to entire users depending on the similarity of priority. In the event of high similarity of preference between users, allocate more weight. Generally, the absolute value of the score difference for the specific items can be divided into 5 conditions, that is $\{0,1,2,3,4\}$. The lower score represents the higher similarity of preference. Therefore, the weights to be distributed are $\{5x, 4x, 3x, 2x, x\}$ respectively (x is an unknown number). (3) Distribute users' resources to items. Allocate different weights to items depending on the ratings that users made. The ratings have five possibilities $\{1,2,3,4,5\}$. As a result, the weights belong to them are $\{x, 2x, 3x, 4x, 5x\}$.

   The Fig. 1 elucidates the process of resource distribution in a simple bipartite network. The circles represent users and squares represent items. The target user is marked by red color.
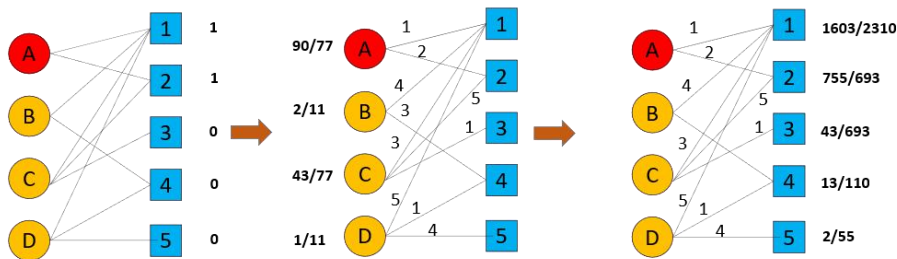

Figure 1. Elucidation of NProbS

3.2. A heuristic method to initialize population based on selection
As a variation of optimization algorithms, heuristic method can be combined with genetic algorithm to reduce time for searching. Moreover, when it comes to select items according to ratings, users are

more likely to choose the item with high score. As a result, we proposed a heuristic method to of initialization. It can be illustrated as follows: (1) Assuming that the individual size is $|U| \times L$ ($|U|$ represent the total number of users and L represent the length of recommendation list). Use NProbs algorithm to obtain the rating matrix R_1 for prediction, which is as large as $|U| \times L$. (2) For each user, select the top-L items as the recommendation list according to the row predictive vector in R_1 . Then combine the top-L items to generate an individual of the initial population to guide the evolution of the population. (3) The rest initial individuals can be obtained by the following mean: Firstly, normalize R_1 to get R_2 . Then for the user i, select the item j at random and generate a random number between 0 and 1. If number $< R_2 (i, j)$, then add the item into the recommendation list of user i, vice versa. (4) Repeat the above process L times and get the recommendation list of user i . Subsequently, repeat $|U|$ times to get the initial individuals. Finally, repeat the times of population size to get the initial population.

3.3. A crossover strategy based on the conservation of parent similar gens

To further the diversity of recommendation results, a novel crossover strategy is proposed. For the beginning, utilize the binary tournament to select an individual based on dominated sorting and crowing degree and repeat this procedure 3 times to get Patent1, Parent2 and Parent3. For the user i, recommendation lists rem_list1, rem_list2, rem_list3 can be obtained to compose rem_list. Record the appearance of different items. If the appearance time is greater than 1, mark the item lists as list1. If the appearance time is equal to 1, mark the item lists as list2. If the items quantity M in list1 is larger than L, preserve L items to the child by random selection. If M $\leq$ L, retain all the items in list1 to the child and randomly select the rest L-M items from list2. Repeat the process to gain the Child and new child population.

3.4. Framework of the proposed MOEA-NProbS

As shown in the following Fig. 2, the proposed recommendation algorithm has a similar framework to NSGA-II. Firstly, classify users U into k clusters. Then for each cluster, execute the following steps: utilize NProbS to obtain predictive rating matrix; initialize the population and calculate the target function; execute fast non-dominated sorting and calculate crowding distance. Afterward, iterate the population by following steps until it meets the maximum. Select F2, F2, F3 from parent P by the binary tournament. Generate offspring based on the proposed crossover strategy. Execute fast non-dominated sorting, calculate crowding distance and obtain new parent generation by elitist strategy. At last, select the first frontier individuals.

Algorithm 1 Framework of the Proposed MOEA-NProbS

Input: N (population size); R (rating matrix); $gmax$ (maximum iterations)
Output: FS (Pareto set)
1: $[U_1, U_2, ..., U_k] \leftarrow U$ //classify users into k clusters
2: **for** $i = 1$ to k **do**
3:      $R \leftarrow NProbS(R)$ // obtain predictive rating matrix for unknown items
4:      $P \leftarrow Initialization(U_i, R, N)$ // initilize population
5:      $P \leftarrow$ Do Fast non-dominated sorting on P
6:      $P \leftarrow CrowdingDistance(P)$
7:      **for** $gen = 1$ to $gmax$ **do**
8:          $F_1, F_2, F_3 \leftarrow BinaryTournament(P)$ //select three parents
9:          $Child \leftarrow GeneticOperator(F_1, F_2, F_3)$ //generate new generation by crossover strategy
10:          $P \leftarrow$ Do Fast non-dominated sorting on $P \cup Child$
11:          $P \leftarrow CrowdingDistance(P)$
12:          $P \leftarrow Elitism(P)$ //generate new parent generation
13:      **end for**
14: **end for**
15: $FS \leftarrow Selection(P)$
16: return FS

Figure 2. Framework of the Proposed MOEA-NProbs

4. Conclusion

In this paper, a novel multi-objective model is proposed to optimize accuracy and diversity at the same time. NProbS is implemented to make ratings, which is a variation of the Probs algorithm. Moreover, a heuristic algorithm is adopted to accelerate the initialization of the population. The proposed MOEA-NProbS utilize a new strategy of crossover based on the similar genes of parents. From theoretical analysis, MOEA-NProbS is effective to improve the diversity and accuracy of the recommendation. In addition, for multiple users, the algorithm can generate multiple recommendations in only one run.

However, the MOEA-NProbS need to be verified by experiments and compared with prevalent MOEA algorithms. Our future work will confirm the validity of MOEA-NProbs based on the data set MovieLens. Moreover, to validate the advantage of MOEA-NProbs, it will be compared with well-known techniques, including CF and MF.

References

[1]   R. Mu, "A Survey of Recommender Systems Based on Deep Learning," in IEEE Access, vol. 6, pp. 69009-69022, 2018.

[2]   M. Lerato, O. A. Esan, A. Ebunoluwa, S. M. Ngwira and T. Zuva, "A survey of recommender system feedback techniques, comparison and evaluation metrics," 2015 International Conference on Computing, Communication and Security (ICCCS), 2015, pp. 1-4.

[3]   M. Liphoto, C. Du and S. Ngwira, "A survey on recommender systems," 2016 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2016, pp. 276-280.

[4]   Q. Ren, Y. Zheng, G. Guo and Y. Hu, "Resource Recommendation Algorithm Based on Text Semantics and Sentiment Analysis," 2019 Third IEEE International Conference on Robotic Computing (IRC), 2019, pp. 363-368.

[5]   S. Cheng, G. Horng and C. Chou, "The Adaptive Recommendation Mechanism for Distributed Group in Mobile Environments," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 6, pp. 1081-1092, Nov. 2012.

[6]   Nair, C. Paralkar, J. Pandya, Y. Chopra and D. Krishnan, "Comparative Review on Sentiment analysis-based Recommendation system," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-6.

[7]   Rohit and A. K. Singh, "Comparison of measures of collaborative filtering recommender systems: rating prediction accuracy versus usage prediction accuracy," 2017 International Conference on Innovations in Control, Communication and Information Systems (ICICCI), 2017, pp. 1-4.

[8]   Saúl Vargas and Pablo Castells. 2011. Rank and relevance in novelty and diversity metrics for recommender systems. In Proceedings of the fifth ACM conference on Recommender systems. Association for Computing Machinery, New York, NY, USA, 109–116.

[9]   Mi Zhang and Neil Hurley. 2008. Avoiding monotony: improving the diversity of recommendation lists. In Proceedings of the 2008 ACM conference on Recommender systems. Association for Computing Machinery, New York, NY, USA, 123–130.

[10]   J. Li, H. Wu, C. Hu and F. Liu, "Case-based Reasoning for Resolving the Diversity-accuracy Dilemma of Recommendation System through Subdivision," 2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI), 2020, pp. 357-360.

[11]   Andrii Maksai, Florent Garcin, and Boi Faltings. 2015. Predicting Online Performance of News Recommender Systems Through Richer Evaluation Metrics. In Proceedings of the 9th ACM Conference on Recommender Systems. Association for Computing Machinery, New York, NY, USA, 179–186.

[12]   Marco Tulio Ribeiro, Nivio Ziviani, Edleno Silva De Moura, Itamar Hata, Anisio Lacerda, and Adriano Veloso. 2015. Multiobjective Pareto-Efficient Approaches for Recommender Systems. ACM Trans. Intell. Syst. Technol. 5, 4, Article 53 (January 2015), 20 pages.

[13] Y. Zuo, M. Gong, J. Zeng, L. Ma and L. Jiao, "Personalized Recommendation Based on Evolutionary Multi-Objective Optimization [Research Frontier]," in IEEE Computational Intelligence Magazine, vol. 10, no. 1, pp. 52-62, Feb. 2015.

[14] Deepak Agarwal, Bee-Chung Chen, Pradheep Elango, and Xuanhui Wang. 2011. Click shaping to optimize multiple objectives. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. Association for Computing Machinery, New York, NY, USA, 132–140.

[15] M. A. Ghazanfar and A. Prugel-Bennett, "A Scalable, Accurate Hybrid Recommender System," 2010 Third International Conference on Knowledge Discovery and Data Mining, 2010, pp. 94-98.

[16] L. Huettenberger, C. Heine and C. Garth, "Decomposition and Simplification of Multivariate Data using Pareto Sets," in IEEE Transactions on Visualization and Computer Graphics, vol. 20, no. 12, pp. 2684-2693, 31 Dec. 2014.

[17] G. Wang, H. Liu and L. Qing, "Bipartite network projection and its application in recommendation systems," 2013 25th Chinese Control and Decision Conference (CCDC), 2013, pp. 5052-5057.

[18] LÜ, L., LIU, W.. Information filtering via preferential diffusion[J]. Physical review, E. Statistical, nonlinear, and soft matter physics,2011,83(6 Pt.2):066119-1-066119-12.