

Research on Group Recommendation Mode based on Item Features and Time Factor

Zhufeng Qiao^{1, 2, a}, Lei Chen^{1, 2, b}, JianXin Guo^{1, 2, c}

¹Institute of Agricultural Information and Economics of BAAFS

²The Research Center of Beijing Engineering Technology for Rural Remote Information Services, Beijing, China

^aqiaozf@agri.ac.cn, ^bchenl@agri.ac.cn, ^cguojx@agri.ac.cn

Keywords: Item Features, Preference Fusion, Group Preference.

Abstract. This paper integrated item domain features, time factors and user ratings, using TF-IDF method constructing single member user preference model in each feature. Then, analyze the interaction between groups of users and the influence of user preference on group model by calculating the similarity of user and group attributes. Finally, the group preference features are obtained.

1. Introduction

In recent years, research on the development of recommendation system is very rapid, appeared many types of recommender systems. However, many of the daily activities are involved in the form of a group of multiple users. According to the user behavior similarity, divide them into different groups. Group recommendation can achieve the minimum cost to meet the needs of more users, and alleviate computation recommendation system complexity. Compared to the personalized recommendation, the group recommendation has some specific problems, such as how to get the group preference, how to solve the members of the preference conflict, how to ensure the whole group and personal satisfaction.

This paper uses the TF-IDF method to build user group preference feature model. Then, analyze the interaction between groups of users and the influence of user preference on group model by calculating the similarity of user and group attributes. Finally, the group preference features are obtained.

2. Item Features Model

In this paper, the main item features considers five aspects: type, program source, publish time, content, label. The feature model of courseware item m_k is :

$$\overline{I}_{m_k} = (C_{m_k p_1}, C_{m_k p_2}, \dots, C_{m_k p_i}, \dots, C_{m_k p_s}) \quad (1)$$

Among, $C_{m_k p_i} = 1$ indicates that the courseware contains the attributes g_i , $C_{m_k p_i} = 0$ means that the courseware does not contains attributes g_i .

3. Personal Preference Model Algorithm

In this paper using TF-IDF way calculate the degree of preference of the statistical user for each feature attribute in all items. TF-IDF is a classical method for calculating the weight of feature items based on statistical analysis. The IF-IDF formula is as follows:

$$TF-IDF = TF(t_i, d_j) \times IDF(t_i) = \frac{f_{t_i, d_j}}{\sum_{m=1}^k f_{t_m, d_j}} \times \log\left(\frac{n}{n_{t_i}}\right) \quad (2)$$

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (3)$$

Among, $TF(t_i, d_j)$ is word frequency, the word t_i in the document d_j appears in the frequency. $f(t_i, d_j)$ is the number of times the word t_i appears in the document d_j . $IDF(t_i)$ is a reverse file frequency.

1) Build single user preference model based on item feature attributes by TF-IDF.

User u_i preference model based on item feature attributes:

$$\overline{L}_{u_i} = (L_{u_i p_1}, L_{u_i p_2}, \dots, L_{u_i p_j}, \dots, L_{u_i p_t}) \quad (4)$$

The formula of the preference degree of user u_i to attributes p_j :

$$L_{u_i p_j} = \frac{n_{u_i p_j}}{n_{u_i p}} \log \frac{m_p}{m_{p_j}} \quad (5)$$

Among, $n_{u_i p_j}$ represents the number of a courseware that has watched by user u_i and containing properties p_j , m_p represents the total number of all properties in distance education platform's resources, $n_{u_i p}$ represents the number of attributes that the user u_i has watched, m_{p_j} represents the total number of courseware for all distance education platform resources that contain the attribute p_j .

2) User preference model optimization based on user score

In order to better represent the user preferences. When calculating the number $n_{u_i p_j}$ of attributes of the courseware that was watched by user u_i . In this paper, the [1, 2, 3, 4, 5] score corresponds to the weight of [0.6, 0.8, 1.0, 1.2, 1.4]. So:

$$n_{u_i p_j} = \sum_{k=1}^n w_{u_i m_k} \quad (6)$$

Among, $k \in [1, n]$, n is the number of user u_i to watch the courseware, $w_{u_i m_k}$ is the user u_i to view each courseware attributes based on the weight of the score.

3) Single user preference optimization model based on time factor

User preferences are not static. So use the forgetting function to reduce the score. In this paper uses the forgetting function based on exponential function to simulate the effect of time on user preference. The formula is as follows:

$$f(t) = \exp\left(\frac{t - t_{\min}}{t_{\max} - t_{\min}} - 1\right) \quad (7)$$

Among, t_{\min} is the difference between the user's scoring time and the reference time, t_{\max} is the difference between the user's latest rating time and the reference time, t_{\min} , t_{\max} are in days, t is access time for users, forgetting function $f(t)$ is monotone increasing function, Range is (0,1). The user scoring formula based on time factor:

$$n'_{u_i p_j} = \sum_{k=1}^n [w_{u_i m_k} f(t_k)] \quad (8)$$

Among, t_k users watch the time of courseware m_k .

The preference degree of user u_i to attribute p_j is expressed as:

$$L'_{u_i p_j} = \frac{\sum_{k=1}^n [w_{u_i m_k} f(t_k)] \times f_{m_k p_j}}{\sum_{k=1}^n [w_{u_i m_k} f(t_k)]} \log \frac{m_p}{m_{p_j}} \quad (9)$$

Among, $f_{m_k p_j}$ is to determine whether the attribute p_j in the courseware m_k .

4. Group Preference Model

Preference fusion need to meet the overall satisfaction, fairness, understandability and other requirements. In this paper, analyze the user's preference features and the similarity of the group average preference features based on the mean value strategy, the higher the similarity between the two, the greater the impact on the group features. Specific steps are as follows:

(1)The average degree of preference of the group to the courseware attributes is as follows:

$$L_{G_i p_j} = \frac{\sum_{u_x \in G_i} \dot{L}_{u_x p_j}}{|G_i|} \quad (10)$$

Among, $|G_i|$ represents the number of members of the user group, $\dot{L}_{u_x p_j}$ represents the degree of preference of the group member u_x to the property p_j .

(2)User feature attribute similarity weight

If the similarity between the user u_x preference and the group average preference is higher, the greater the role of the user u_x in the formation of group preference, Indicates that the user u_x preferences are more representative:

$$\lambda_u = \frac{\text{sim}(u_x \times G_i)}{\sum_{u_j \in G_i} \text{sim}(u_j \times G_i)} \quad (11)$$

Among, $\text{sim}(u_x \times G_i)$ is cosine similarity formula.

Therefore, the preference degree of the group G_i to the item features p_j is:

$$\dot{L}_{G_i p_j} = \frac{\sum_{u_x \in G_i} \lambda_{u_x} \dot{L}_{u_x p_j}}{\sum_{u_x \in G_i} \lambda_{u_x}} \quad (12)$$

5. Summary

This paper presents a method of user group preference modeling based on project features and time factor, Constructed user and group preference model. The method makes full use of user rating, time factors and project domain features. Due to the particularity of group recommendation, this paper takes into account the interaction between the members of the group preference model, which is based on the similarity weight of the user item feature attributes. The experiment is applied to the data set of the distance education platform.

Acknowledgments

Research on Distance Education Group recommendation system based on large data (Youth Research Fund of BAAFS 2017)

References

- [1] Zhang Yujie, Du Yulu, Meng Xiangwu. Group recommendation system and its application [J]. Journal of Computer Science, 2016, 39(4):745-764.
- [2] Meng Xiangwu, Ji Hong. Recommendation system in big data environment [J]. Journal of Beijing University of Posts and Telecommunications, 2015, 38(2): 1-5.
- [3] Guo Junpeng, Zhao Mengnan. Research on community recommendation algorithm for online community users [J]. Computer application research, 2014, 31(3): 696-699.
- [4] Zhu Yuxiao, Lv Linyuan. Review on evaluation index of recommendation system [J]. Journal of University of Electronic Science and technology of China, 2012, 41(2): 163-167.

- [5] Wang Li, Zhang Shulian. Review of group recommendation algorithm based on user preference fusion [J]. Computer knowledge and technology, 2014(7):1511-1515.