E-commerce Service Equilibrium Prediction Simulation Based on Ontology

WU ZHONG
Physical Education School
Wuhan Business University
Wuhan, China
7849800@qq.com

Abstract— **Accurately predicting the needs of users in the process of e-commerce services can expand the scope of e-commerce services and improve the quality of e-commerce services. When forecasting demand, it is necessary to establish different forecasting indicators for different factors that affect e-commerce services. The traditional dual-neighbor selection strategy can only select recommended behaviors based on a single indicator, which reduces the demand for e-commerce services. Accuracy. This paper proposes an ontology-based method for predicting user demand balance in the process of e-commerce services. Firstly, construct an e-commerce service balanced recommendation user ontology, calculate the semantic similarity between the ontology, and obtain a set of similar users. On this basis, predict the neighbors of the target user as the prediction group, and calculate the trust subgroups with higher prediction probability in the group. Through the dynamic measurement of uncertain neighbors for balanced prediction, effectively completed the balanced prediction of user demand in e-commerce services. Simulation experiments prove that the proposed method can effectively improve the accuracy of user demand prediction in e-commerce services, and it is valuable.**

Keywords- E-commerce service, Ontology, Equilibrium Prediction, Simulation

I. INTRODUCTION

With the continuous development of mobile Internet technology, the amount of information on the Internet is increasing day by day, and the recommendation of business objects under the electricity supplier mode has attracted people's special attention [1]. As a new information service mode, demand forecast can effectively alleviate the problem of information overload [2]. However, in most e-commerce services, user demand forecasting system can not accurately calculate the similarity between users' ontology, accordingly reducing user demand prediction accuracy. Under such condition, improving the quality of e-commerce service demand forecast has been paid much attention by the relevant people, and some progress has been made.

However, there are many problems in existing research, such as large deviation [3], difficult to resist the interference of external factors [4], and can not accurately calculate the similarity between users [5].

In view of the problems arising above, it proposes a balanced recommendation method based on ontology for e-commerce services. The simulation experiment shows that the proposed method can effectively improve the accuracy of the recommended system.

II. FORECASTING PRINCIPLE OF USER DEMAND

In the e-commerce service demand forecasting process, the concept of business service context ontology is divided into two types the concept of characteristics and the concept of e-commerce service characteristics according to the semantics, establishing the demand forecast rules. Current user environment and user requirements details are matched by the actual-time demand forecast reasoning of e-commerce. And forecasting the characteristics of user preferences to e-commerce services under the real-time situation, selecting service projects in line with the feature which will be recommended to the user from the feature set, and completing business services in the forecast. The concrete realization of the process as described below.

It is assumed that E represents the set of users and forecast objects, R^l is the concept of situation state, p^l is the concept of e-commerce service characteristics, U_i is the potential feature vector of e-commerce services, then user demand forecast rules are established by equation (1).

$$P_E = E \times \frac{R^l \times p^l}{U_i \times a_{\Re}} \times \mathrm{B}^{\chi} \times S_A \tag{1}$$

In equation (1), $a_{\Re}$ is the relationship network between users, B^{χ} is any two forecast objects, S_A is the relationship between the recommended objects.

Suppose that e representing the user's current environment, using the equation (2) to match the user current environment and the rules of the user requirements forecast.

$$T_y = \frac{e}{P_E} \times p^l \times S_A \tag{2}$$

Suppose that $q_{\Im}$ represents the trusted neighbor user group, I_j is on behalf of the similarity of user j with all neighbor candidate users, by using the equation (3) to forecast the user preferences to e-commerce service features under real-time situations, establish the objective function of user demand forecast in e-commerce requirement service, and complete user demand forecast.

$$P_{\xi} = \frac{e}{T_y \cdot a_{\Re} \times P_E} \tag{3}$$

But by the traditional method the recommendation can only be selected on the basis of single index, which reduces the accuracy of the demand forecast of e-

2473-3636/21/$31.00 ©2021 IEEE

DOI 10.1109/DCABES52998.2021.00009

commerce service. A method of equilibrium forecast for e-commerce service demand based on ontology is proposed.

III. EQUILIBRIUM PREDICTION OF CUSTOMER DEMAND FOR E-COMMERCE SERVICE BASED ON ONTOLOGY

A. Calculation of similarity threshold between e-commerce users

When we do equilibrium prediction of customer demand for e-commerce service, we need to build user ontology for e-commerce service, getting the attribute system of user ontology knowledge, based on that system user attributes are divided into 3 different types. Firstly, when the users register, they provide the basic information of the document which is defined as user basic information. Secondly, it is the document of social information set, which we use to structure the users' knowledge. Thirdly, it is the information document of user's project set. If any user is interested in e-commerce service, he will evaluate the project that he is interested in. Accordingly, we define the document as the scoring information matrix of user interest items. According to the different types of the above documents, the basic attributes, main features and interest of a user can be described in details.

According to the principles above, the similarity of the two users' ontology is calculated by using the equation (4).

$$Sim_{property}(o_1,o_2)=\frac{|p_1\cap p_2|}{|p_1\cap p_2|+\theta\|p_1-p_2\|+|p_2-p_1|} \quad (4)$$

In equation (4), p_1 and p_2 mean the attribute set of O_1 and O_2 respectively. $|p_1\cap p_2|$ is the elements number of p_1 and p_2 intersection. $\|p_1-p_2\|$ is the elements number of set p_1 not belonging to set p_2. While $\|p_2-p_1\|$ represents the elements number of set p_2 not belonging to set p_1. θ represents the similarity threshold calculated by equation (5).

$$\theta=\frac{\{depth(O_1),depth(O_2)\}}{depth(O_1)+depth(O_2)\times Sim_{property}(o_1,o_2)} \quad (5)$$

In equation (5), $depth(O_i)$ is the maximum value of the entire conceptual path of ontology.

B. The accurate calculation of similarity between users

In the two users' ontology, there is the common concept of sequence. In the common sequence between ontologies the concept point is consistent, and the order is the same, which can reflect the similarity more accurately between two users' ontology. These 3 factors: the number of common sub sequence, levels, program are calculated respectively, then in a comprehensive summary, calculating the accurate similarity between the user ontology. Specific steps as follows.

1) The calculation of the number of common sub sequences of user ontology

It is assumed that $CS(O_1,O_2)$ represents the common sub sequence of two users' ontology. And the similarity of $CSnum(O_1,O_2)$ represents the common sub sequence is calculated by the equation (6).

$$CSnum(O_1,O_2)=\frac{2\times NSC(O_1,O_2)\times SC(O_1,O_2)}{S(O_i)+S(O_i)} \quad (6)$$

In equation (6), O_i is the number of its sub sequences, $NCS(O_1,O_2)$ represents the common sub sequences of the two ontology.

2) The calculation of the hierarchy of the common sub sequence of the user's ontology

Suppose that DC represents the depth of the common sub sequence, the similarity between the levels of the common sub sequence is obtained by the equation (7).

$$CS_{Pec}(O_1,O_2)=\frac{D_c-\theta\cdot depth(LCS(O_1,O_2))}{CSnum(O_1,O_2)} \quad (7)$$

In formula (7), depth represents the distance from the concept to the root node of the ontology. $LCS(O_1,O_2)$ represents the minimum common sub sequence.

3) The length calculation of the common sub sequence of the user's ontology

In the two ontology, the longer the length of the common sub sequence is, the higher the similarity is between the corresponding two users. The length factor of the common sub sequence is obtained by using the equation (8).

$$CSlen(O_1,O_2)=\frac{Max(lenth(CS(O_1,O_2)))}{CS_{Pec}(O_1,O_2)} \quad (8)$$

Combining the equations (6), (7) and (8), the similarity function of the user ontology is obtained by using the equation (9).

$$\mathrm{si}\,m(O_1,O_2)=\frac{f_1\times CSnum(O_1,O_2)CSlen(O_1,O_2)\bullet f_2}{CS_{Pec}(O_1,O_2)\otimes CSnum(O_1,O_2)} \quad (9)$$

In the equation (9), f_1 and f_2 represent the weights of the similarity between the two types of user ontology.

The initial weights of user ontology similarity function are set up in sigmoid. The similarity threshold above is taken into consideration, and the similarity function is introduced into the equation (9). The similarity between users is accurately calculated by using equation (10).

$$f(x)=\frac{1-e^{\varphi x}}{1+e^{\varphi x}}\times\frac{2}{1+e^{\varphi x}}\times\frac{\mathrm{si}\,m(O_1,O_2)}{\theta} \quad (10)$$

In equation (10), φ represents the fixed constant on the basis of specific application. x is the similarity value calculated by various factors and $f(x)$ is the initial weight of different factors.

C. Realization of equilibrium prediction of e-commerce service

On the basis of the user' ontology similarity, it collects the neighbor objects of the forecast targeted users, establish groups of users' forecast demands, calculates the trusty subgroup with higher prediction probability, predicts the dynamic measurement of uncertain neighbors.

It effectively completes the prediction of users' demands. The specific implementation process, as shown below.

Assuming I_j representing user a of a service without scoring, $R_{a,j}$ is the forecast user scoring to the service. In the service without user's score, the forecast user provides the maximum score to recommended selectively to the user, by using the equation (10) to get $f(x)$ and search k neighbors for the unknown I_j, using the equation (11) to obtain the scores with user ontology forecast.

$$R_{a,j} = \bar{R}_x + \frac{sim'(U_a, U_b)^a \times (R_{b,j} - \bar{R}_b)^\delta}{\sum sim'(U_a, U_b) \times f(x) \cdot I_j} \tag{11}$$

In equation (10), $\bar{R}_a$ represents the scoring for other service types by user U_a in e-commerce, $sim'(U_a, U_b)$ representing the scoring intersection of user U_a and user U_b. $\bar{R}_b$ represents the scoring for other service types by user U_b.

It is assumed that $S(U_a)$ the prediction set selected by user ontology and $S(I_j)$ the prediction set selected by e-commerce, according to the equation (12), through them we calculate the common scoring number between user ontology more than user demand forecast group of the set threshold ε.

$$S'(U_a) = \varepsilon \cdot \frac{\{U_x | sim'(U_a, U_x) \ge \mu\}}{S(I_j) \times R_{a,j} \times S(U_a)} \tag{12}$$

In equation (12), U_x represents x users set. μ represents the threshold of similarity calculation between the service types of the e-commerce system.

By the equation (13), the trusty subgroup with higher prediction accuracy is obtained from the user's demand prediction group.

$$S'(I_j) = \varepsilon \cdot \frac{\{U_x | S'(U_a) \ge \mu\}}{S(I_j) \times R_{a,j} \times S(\gamma)} \tag{13}$$

In equation (13), γ is the set trusty degree threshold. By equation (14), the user demand equilibrium prediction model in the process of e-commerce service is established.

$$R'_{a,j} = \lambda \left[\bar{R}_a + \frac{(R_{x,j} - \bar{R}_x)}{sim'(U_a, U_b)} \times \frac{U_x \times (I_j, I_y)}{(\bar{R}_y, \bar{R}_j)} \right] \tag{14}$$

In equation (14), $\bar{R}_a$ and $\bar{R}_x$ represents the average score to other services of user U_a and U_x respectively. $\bar{R}_j$ and $\bar{R}_y$ represents the average score value of the total number of known users I_j and I_y. λ represents the neighbor of the user.

IV. Simulation Experiment Results and Analysis

In order to prove the effectiveness of the proposed method of user demand equilibrium prediction in the process of ontology e-commerce service, simulation experiments are needed. The data set used in the experiment is derived from the data of a university mobile catering service platform. During the experiment we totally recruited 200 student volunteers to participate in the use of the prototype system. The experimental time was 100 days. During the experiment, the experimenter had to order every meal through the interactive interface of mobile client inputting that time situation and and selecting restaurants. The system generated one user demand forecast table based on the situation.

A. Ratio of customer demand forecast efficiency in e-commerce service

By using the improved algorithm and literature 3, 4 and 5 algorithm to predict user demand in e-commerce service, the prediction efficiency of different methods is compared at different time, and the results are shown in Figure 1 as below.

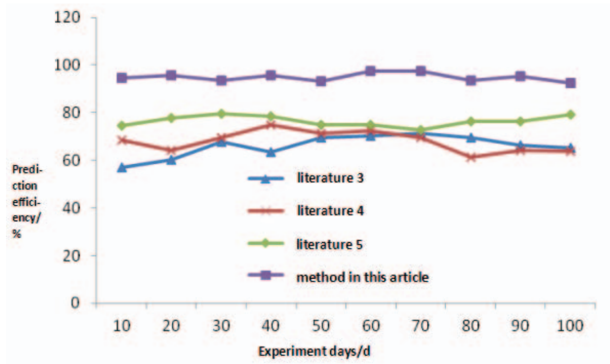

Figure 1. Comparison of the prediction efficiency of different algorithms.

Figure 1 analysis shows that with the same experimental days in the literature of 3, 4, and 5 the prediction efficiency all was lower than that of the improved algorithm. Through the analysis of the whole experimental process, we could know that the prediction average efficiency of literature 3, 4, and 5 is about 66.03%, 67.98%, 76.48% respectively, while that of the improved algorithm is about 94.92%, being increased by 43.73%, 39.62%, 24.10% to compare the three method above, which shows that the improved algorithm has high prediction efficiency.

B. The average absolute deviation ratio of user demand forecast in e-commerce service

Mean Absolute Error(MEA) which represents the average absolute deviation is defined as the recommended accuracy measurement standard for different algorithms. By MEA we calculate the deviation between the predicted user scoring and the actual user scoring. The smaller the MAE value is, the higher the accuracy of the recommendation results is. The equation (15) is used to calculate MAE.

$$MAE = \frac{1}{n}\left[r_{u,i} - \hat{r}_{u,i} \right] \quad (15)$$

In equation (15), $\hat{r}_{u,i}$ represents the forecast results, $r_{u,i}$ representing the collection of the types of e-commerce services, and n representing the size.

By the improved algorithm and literature 3, 4, and 5 algorithms to predict user demand in e-commerce service, the average absolute deviation prediction of different methods is compared at different time, and the results are shown in Figure 2 as below.

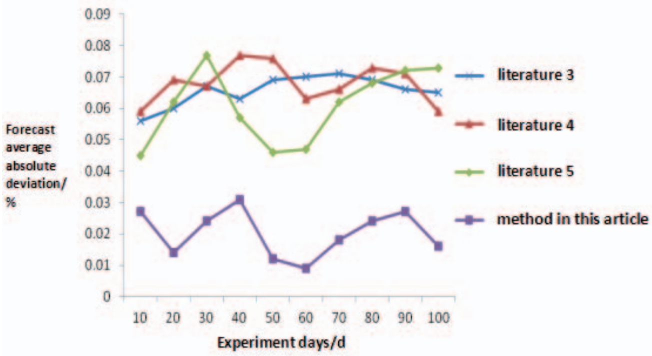

Figure 2. Comparison of predicted average absolute deviation of different algorithms.

Figure 2 shows that with the same experimental days under the condition that the average absolute deviation of literature 3, 4, and 5 is much higher than that of the improved algorithm. From the analysis of the whole experimental process, the prediction average efficiency of literature 3, 4, and 5 is about 0.0656%, 0.068% and 0.0609% respectively, while that of improved algorithm about 0.0202%, which shows that the improved algorithm has small prediction deviation.

C. *Comprehensive and effective comparison of user demand forecast*

By the improved algorithm and the algorithms of literature 3, 4, and 5 to forecast users' demands, within a given time period, comparing the accuracy of several algorithms (T)%, reliability E (%) and memory, with the results we could measure the comprehensive effectiveness of different algorithms to e-commerce service equilibrium recommendation. The results are shown in table 1.

TABLE I. THE COMPREHENSIVE EFFECTIVENESS WITH DIFFERENT ALGORITHMS TO USER DEMAND FORECAST OF E-COMMERCE SERVICE

Algorithms	(T)%	E（%）
algorithm of literature 3	69.17	62.91
algorithm of literature 4	71.13	65.28
algorithm of literature 5	76.31	73.12
Improved algorithms	95.32	94.38

Analysis of Table 1 shows that in terms of accuracy, the prediction accuracy of literature 3, 4, and 5 is about 69.17%, 71.13% and 76.31% respectively, while the improved method is about 95.32%. In terms of the reliability of prediction, the prediction accuracy of 3, 4, and 5 is about 62.91%, 65.28% and 73.12% respectively, while the improved method is about 94.38% . It is improved that this method is more effective than other three, and it can meet the actual demands of users.

Acknowledgment

This research was supported by the Application Foundation Frontier Project (2020010601012294), the Wuhan University of Business Doctoral Fund projects (2021KB002),the Teaching and Research Program at City-owned Universities in Wuhan (2019102), the Provincial University Teaching Research Project of Hubei Province (2018466), and the Construction of Specialty Features of Sports Economics and Management Major in Wuhan Business University (2019N013).

References

[1] L. Guo, J. Ma, Z. M. Chen and H. R. Jiang, Incorporating Item Relations for Social Recommendation, Chinese Journal of Computers. 37 (2014) 219-228.

[2] A. T. Yang, Y. Tang, J. B. Wang and J. G. Li, Personalized Friends Recommendation System Based on Game Theory in Social Network, Computer Science. 42 (2015) 191-194.

[3] Z. H. Huang, Top-n Recommendation Algorithms for Cloud Data, Acta Electronica Sinica. 43 (2015) 54-61.

[4] Y. Jiang, Research and Simulation on Information Trust Calculation Model of P2P Network Nodes, Computer Simulation. 31 (2014) 298-301.

[5] J. Y. Tang, S. C. Huang, Application of optimized similarity calculation in recommender system, Electronic Design Engineering. 23 (2015) 46-48.