

Multi Agent Collaborative Filtering Algorithm Based on Double Layer Behavior Driven

Yan Cheng

College of Electronics and Information Engineering, Tongji University;
College of Computer and Information Engineering, Jiangxi Normal University
Shanghai, China; Nan Chang, China
chyan88888@jxnu.edu.cn

Zhiming Yang

College of Computer and Information Engineering, Jiangxi Normal University
Nan Chang, China
709830862@qq.com

Weisheng Xu

College of Electronics and Information Engineering, Tongji University
Shanghai, China

Xiaoling Ao

College of Computer and Information Engineering, Jiangxi Normal University
Nan Chang, China

Abstract—In recent years, with the development of the Internet and network learning community rapid development, information technology is changing people's ways of e-learning with amazing speed. At the same time, learning trek, knowledge overload problems gradually emerged. Personalized learning recommendation has gradually become an important means for people to quickly learn and master the knowledge, but also reflects the people-oriented personalized learning mode. However, recommendation algorithm based on collaborative filtering still faces problems such as sparsity, scalability, cold start and precision. In this paper, a multi Agent collaborative filtering model based on double layer behavior is proposed for the problem of data sparsity and accuracy. Using social network relationship and dynamic agent perception and the ability to interact with learning, combined with individual behavior to measure the degree of interest, group behavior measure of trust and influence of learning resources for collaborative forecasting scale, reduce the score prediction error and the problem of learning trek.

Keywords—personalized learning recommendation; multi agent; interest; collaborative filtering

I. INTRODUCTION

With the popularity of the Internet and the rapid development of Web2.0[1], the traditional teaching mode has been unable to meet the people's demand for knowledge. The development of distance education[2] along with the development of information technology has made great progress, in this era that full information of the data, in the face of the massive information, searching engine (Google, Baidu, etc.) can help people quickly filter information. But it is really affect the learning efficiency in looking for information in the vast information[3]. So personalized learning recommendation [4]has gradually become the focus of research in the network teaching. To create a new teaching

mode which can provide the online learners with personalized learning suggestions and precise curriculum recommendation and meet the individual needs of the learners, which has become the inevitable direction of the development of distance education[5].

Scholar Wu et al. work on the design of intelligent search engine and they have put forward new opinions on the personalized, intelligent. But because it cannot fully understand students' query, search efficiency and result accuracy is not high[6]. Scholar Wang worked on apriority association rules algorithm, and improved k nearest neighbor algorithm and the rock clustering algorithm based on the similarity of user access patterns which using most similar neighbor user automatic recommended learning resources. And it also has accuracy and sparsity problem[7]. According to the characteristics of user's interest, the sense of the resources of interest or a list of information are recommended to the online learners which help users quickly find interesting content[8].At present, collaborative filtering[9]is one of the most widely used and mature algorithms in the personalized recommendation system. According to the characteristics of collaborative filtering, by reviewing the literature found filtering recommendation system based on collaborative has the following problems[10]: 1) the sparsity problem. 2) cold start problem. 3) the accuracy of the problem. 4) scalability issues.

The personalized recommendation technology is used in the field of electronic commerce which according to the user's preference to recommend the items of interest[11]. This paper design double driving behavior of multi agent collaborative scoring model considering the sparse problem in collaborative filtering algorithm and the problem of accuracy , indirect data filled and through the simulation experiment found that combines the double driven multi agent cooperation model

This work was supported by National Natural Science Foundation of China (NSFC, Grant No.61262080), Jiangxi Province Science and Technology Support Major Project (20151BBE50121, 20161BBE50086), Science and Technology Key Project of Jiangxi Province Education Department (GJJ150299)

score data sparsity problem is solved effectively, and improve the accuracy.

II. COLLABORATIVE SCORING MODEL BASED ON DOUBLE BEHAVIOR DRIVEN AGENT

Now the information technology is into the field of network education. The pursuit of information diversification and personalized learning trend develop quickly under the situation of education information overloading as well as learners and educators[12]. Through mobile phones, computers, television and other terminal devices to access the Internet, to achieve remote online education, online learning, in meeting the basic demand and began searching for the individual needs of higher level[13]. Namely learners do not take the initiative to search for learning resources situation. It is difficult to improve the personalization recommendation quality by single interest model and measurement between the learner's similarities to recommend.

According to the learner's individual behavior this paper measured learners on learning resources prediction score, and combined with the learning groups and other learners between interactive behavior and network effects of dynamic behavior, with multi agent collaborative score.

To construct the double behavior driven multi agent collaborative scoring model (see Fig.1.). It will score the interest resources collaborative for learners may interest. Generally use the interest degree to measure the learner's individual behavior, use the trust degree to measure learning and groups in other learners between interactive behaviors, and use the influence to measure network dynamic behavior in groups.

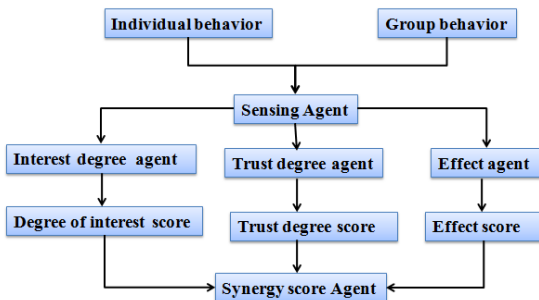


Fig. 1. Multi agent collaborative scoring model driven by double layer behavior

The behavior of individual interest degree model and group behavior trust and influence model is constructed according to the driving double behavior. According to the corresponding time period learners learning behavior discount factor, collaborative agent on the last score data for collaborative scoring which treat as a collaborative filtering recommendation based on the prediction score.

According to double behavior driven collaborative scoring model of multi Agent, this paper established the measurement model of interest, trust degree and influence:

A. Interest Measure Model of Individual Behavior

In this paper, we can measure the online learner's interest in learning content through the combination of the e-learning behavior of online learners. If in process of browsing page, the learner did some action like save, print, copy, keep, that means learners have higher interest for the things in this page. Then set up the learning interest 1; otherwise, the available secondary behavior to calculate the degree of interest. An important learning behavior formula is:

$$IR = \begin{cases} 1, & \text{if } action \in \{save, print, copy, keep\} \\ f(rt, rc, cc), & \text{otherwise} \end{cases} \quad (2-1)$$

Among them, IR represents to the learners' interest; $action$ represents actions in browsing the web, and save, print, copy, keep represent save, print, copy or collect the page. rt represents learners' reading time, rc represents the number of visitors in learning, cc represents click times when the learners browse page. $f(rt, rc, cc)$ represents the functions between rt, rc, cc .

According to the learner's individual learning behavior, if the degree of interest between the two learners is similar, it means that prediction score maybe more close. Hypothesis, the target learners have K closest neighbors, the score prediction that the learners made according to the degree of interest is as follow:

$$Inter(u, i) = \overline{IR}_u + \frac{\sum_{v \in K_u} sim(u, v)(R_{v,i} - \overline{R}_v)}{\sum_{v \in K_u} |sim(u, v)|} \quad (2-3)$$

Among them, \overline{IR}_u represents the average value of the interest degree measure of the learner u , $sim(u, v)$ represents the similarity value of the interest between the two learners. K_u represents K nearest neighbor of the learners u .

B. Construct the Trust Degree Measurement Model of Group Behavior.

In the reality of the social learning network, learning members are more willing to accept learning resource recommendation from the people who he trusts in the network[14]. According to the trust model, the combination between the online learners' interaction concerns, the degree of trust among learners can be quantified according to a certain scoring mechanism. The model can be defined as follows:

Definition 1: Trust relations is represented by a four tuple (E_a, E_b, c, t) . The entity E_a, E_b the two sides of the trust relationship; c represents the interaction between the entities; t represents the interaction time.

In general, the trust degree of the network is divided into direct trust and indirect trust. The former indicates that there is a direct relationship between the network learners; the latter represents the indirect trust that is formed by the third party. According to the definition 1, $D(E_a, E_b, c, t)$ the direct trust

between the learners. In the same way, indirect trust can be recorded as $I(E_a, E_b, c, t)$. λ represents the weight of the direct trust, then the overall trust can be defined as:

$$O(E_a, E_b, c, t) = \lambda D(E_a, E_b, c, t) + (1 - \lambda) I(E_a, E_b, c, t), 0 \leq \lambda \leq 1 \quad (2-4)$$

As a result of the quantitative trust, the specific evaluation results and the participation of learners have a relationship. Through a learner E_a needs comprehensive evaluation of the trust of another learner E_b . Generally, $D(E_a, E_b, c, t)$ can be directly obtained through the connection of the network nodes, but the indirect trust degree of the learners E_b needs to be evaluated comprehensively according to all the entities in the domain.

According to the formula (3-4), the learner can be quantified as:

$$O_a^b(E_a, E_b, c, t) = \lambda D_a^b(E_a, E_b, c, t) + (1 - \lambda) \frac{1}{n} \sum_{i \in Q} I(E_i, E_b, c, t) \quad (2-5)$$

$D_a^b(E_a, E_b, c, t)$ refers to the direct trust degree evaluation of learners from E_a to E_b ; Q represents all the attention refers to a collection of entities of E_b ; n represents a collection of learners Q , $\frac{1}{n} \sum_{i \in Q} I(E_i, E_b, c, t)$ said other learners E_b trust the average value of evaluation.

According to the network relationship between social network learners, after calculated for learners u of trust, based on collaborative filtering theory, assumptions, learners u exist as K nearest neighbor, according to the trust degree of the resource object of rating prediction formula is:

$$Trust(u, i) = \bar{O}_u + \frac{\sum_{v \in K_u} O(u, v)(O_{v,i} - \bar{O}_v)}{\sum_{v \in K_u} |O(u, v)|} \quad (2-6)$$

Among them, \bar{O}_u the represents the average value of trust in a period of time for the learners u , $O(u, v)$ represents the value of trust between learners u and v .

C. The Measure Model of The Influence of Group Behavior.

Common influence evaluation model, based on the user's forwarding behavior modeling. Due to the learning website, network learners learning time is limited and the learners learning reading time is also related to the, so in models need to consider learners' reading behavior and forwarding delay, according to a learner's article is forwarded and reading, to estimate the study of influence.

The calculation formula $I_u(t_s, t_e)$ for the overall expected value of the articles published by the learner u at the time t_s of the article before t_e being read:

$$I_u(t_s, t_e) = \sum_{v, v \rightarrow u} \int_{t_s}^{t_e} P_v(t_s, t) [1 + q_v I_v(t, t_e)] dt \quad (2-7)$$

Among them, $P_v(t_s, t_e)$ refers the probability that the learner v at the moment t reading released at the moment t_s ; learner v read this article after forwarding the conditional probability representation, reflect the forward preference learner v ; $v \rightarrow u$ learner v focus on learner u .

The expected value calculation considers not only the learner brings to your friends directly influence, which $\int_{t_s}^{t_e} P_v(t_s, t) dt$ reflects the learner v directly reading the number of expected value, but also takes into account the learning to bring indirect effects, which $\int_{t_s}^{t_e} P_v(t_s, t) q_v I_v(t, t_e) dt$ reflects the forwarding brings global reading times expected value.

Due to the forwarding a learner can again be forwarded, the Global Forwarding behavior will not be recorded directly, if from a global study to measure the disseminators of knowledge capacity. This model in reading times expected on the basis of the values defines the study forward frequency measurement model:

$$RP_u(t_s, t_e) = \sum_{v, v \rightarrow u} \int_{t_s}^{t_e} P_v(t_s, t) q_v [1 + RP_v(t, t_e)] dt \quad (2-8)$$

$RP_u(t_s, t_e)$ refers to the learner overall expected value of the number of times released by the learner u at the time t_s . The probability of the model q_v to determine whether the learner v forward this article, as long as the forward, will be the learner's contribution into account.

According to the learner's situation, the establishment of the impact assessment model for reading and forwarding:

$$R_{u,i} = I_u(t_s, t_e) + RP_u(t_s, t_e) \quad (2-9)$$

After calculating the influence of the learner, the target learners u need to be calculated to the resource i according to the influence:

$$Influ(u, i) = \bar{R}_u + \frac{\sum_{v \in Q} w_v (R_{v,i} - \bar{R}_v)}{\sum_{v \in Q} w_v} \quad (2-10)$$

The above three models are used to predict scores. Double action driven multi agent collaborative scoring model, dynamic perceptual learning of social network information to assist for the score of collaborative user interest degree, trust degree and influence, reduce a single score data sparsity problem. The improved collaborative filtering can be extended to improve the accuracy of recommendation algorithm.

III. MULTI AGENT COLLABORATIVE SCORING MODEL DRIVEN BY DOUBLE LAYER BEHAVIOR

For a social learning network platform, multi Agent interaction (see Fig.2.) through the perception of the environment and space environment. In the interactive process, the agent perception accept user U environment state of the input, and mapped to the perceived as the agent behavior A and the marked behavioral output to measure agent (including the interest of agent and trust degree of agent and influence agent), and metric agent accept the discount factor η for the environment.

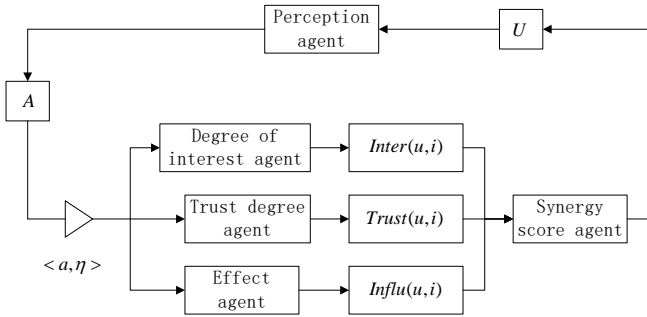


Fig. 2. Agent and environmental interaction process

According to the interaction process between Agent and the environment, the model can be formalized as a nine tuple $\langle S_{gt}, U, A, a, \eta, Inter_{gt}, Trust_{gt}, Influ_{gt}, C_{gt} \rangle$.

Where S_{gt} said agent perception; U represents the users of the activities of the network learner; A represents of learning behavior collection; representation of a set of control behavior, $\langle a, \eta \rangle$ represents the current behavior a , discount factor η for a state; and, respectively, $Inter_{gt}$, $Trust_{gt}$, $Influ_{gt}$ indicating interest agent, degree of agent and influence agent trust based on the corresponding model prediction score, C_{gt} represents output value of cooperative agent score.

Assuming that when mapping from the Agent learning environment state space to the Agent behavior space, the Agent action will be instantaneous to get the environmental return value without affecting the subsequent state and action. The behavior state perception function is defined as:

$$T: U \times A \rightarrow T(U) \quad (3-1)$$

Among them, $T(U)$ represents the probability distribution behavior of a set of discrete on the user set U , the behavior of the state aware function defines the current state of the behavior of the state of perception of Agent may reach the next state of the behavior of probability.

When Agent aware of the learner u , the corresponding behavior a is occurring, Agent on the behavior of the tag, and output to the corresponding metrics model, to get the current moment of synergy Agent score:

$$C_{gt}(u, a) = \alpha Inter_{gt}(u, a) + \beta Trust(u, a) + \gamma Influ(u, a) \quad (3-2)$$

For this model, if the value of the three factors

$\alpha = 1, \beta = 0, \gamma = 0$ is that the model considers only the learner's interest, the two layer behavior driven model will degenerate into a general individual behavior driven model.

IV. EXPERIMENT AND ANALYSIS

This experiment is aimed at the IT of CSDN in Beijing city of China and the Blog member network relationship and the historical score of download resources collected. The original data sets are 52168 members, 4695 of the amount of learning resources, the relationship between the learners are 265298, the history of the record and learning behavior are 596580. For the research of recommendation system, the common data sets are Netflix and Lens Movie, and the corresponding data sparsity is 1.2% and 4.5%. In contrast, this study uses a more sparse data.

In addition, statistics of one learning resources, the number of scores, followed the basic law in line with the distribution, as shown in Fig. 3.

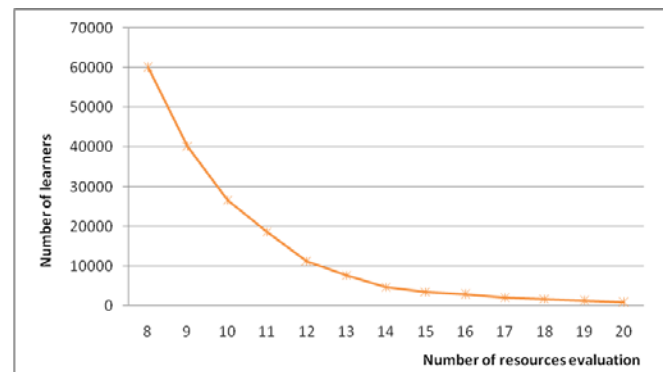


Fig. 3. The statistical distribution map about evaluation number corresponding to resources

At present, the common recommendation algorithm accuracy evaluation standard has the decision support precision measurement and statistical precision measurement method. This experiment chose statistical accuracy which calculating the results recommended by MAE (Mean Absolute Error)[15] which is a measure of the actual value and the predicted value between the average absolute deviation, the verification algorithm of resource prediction score and actual learning resources between the scores of the deviation, to measure the recommendation algorithm results accuracy. It is concluded that the results of the proposed algorithm I_r are focused on the learning resources prediction score $\{p_1, p_2, \dots, p_k\}$, and the actual score of the corresponding target learners $\{q_1, q_2, \dots, q_k\}$ can be obtained according to the calculation formula:

$$MAE = \frac{\sum_{i=1}^k |p_i - q_i|}{k}, i = 1, 2, \dots, k \quad (4-1)$$

Cross validation method is used in this experiment, and the original sample set is divided into 5 parts. Then, each sample data is recommended and evaluated, and the values of MAE

in each experiment are recorded. According to algorithm in chapter three double behavior driven by multi agent collaborative scoring model, selecting the reasonable interest degree, trust degree and influence weighting factor values α, β, γ , the recommended algorithm are better recommendation results, avoid the interference factor of α, β, γ to the algorithm.

In order to verify the three factors, to get a group of the best value of α, β, γ , and make the value of MAE minimum. By controlling variable method, firstly, the parameter value of the interest degree α is fixed, and the parameter values of trust β and influence γ are adjusted and calculated the value of MAE . The Fig.4. $\alpha = 0.4$, calculated MAE in some different groups of β, γ .

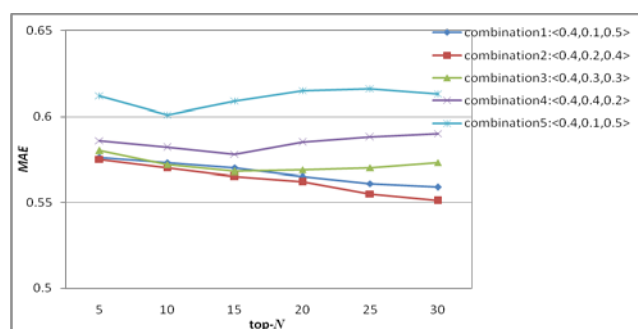


Fig. 4. Value of MAE for different parameters

The Fig. 4-2 shows that the value of interest α is fixed in score model, when learners influence factor is bigger than the trust; the accuracy of the recommendation algorithm is gradually increased. Experimental data show that when $\alpha = 0.4$, $\beta = 0.2$, $\gamma = 0.4$ when selected, and recommended, the accuracy of the results reached the highest.

Similarly, by the traversal of all the value of α, β, γ , the calculation shows that when α, β, γ are 0.4, 0.2, 0.4, the minimum MAE that fusion factors recommended value algorithm is 0.551 which get better recommendation results.

On the basis of the above experiments, in full consideration of the three factors, the recommendation algorithm can obtain better accuracy. The relationship of social network structure can effectively reduce the interference problem of data sparsity and improve the accuracy of recommendation algorithm.

V. CONCLUSION

In this paper, a multi agent collaborative scoring model is designed. Using agent dynamic sensing ability and learning space environment interact and it will mapping behavior state space and giving the collaborative score according to the corresponding interest agent, trust degree of agent and influence agent. The model makes full use of individual behavior metrics of interest, and group behavior measure of trust and influence between learners and resources for prediction score, which reduced the traditional collaborative

filtering algorithms only consider user interest degree score to bring the data sparseness problem, also improve the algorithm recommendation accuracy. The experiments were carried out and algorithm analysis of establish convergence degree of interest and trust degree and influence on the comprehensive evaluation of the similarity model, and three factor of the algorithm accuracy of the recommendation effect, select the optimal factor combination α, β, γ . The experimental results show that the proposed algorithm has higher accuracy. By simulating the update of the score, it is proved that the algorithm has been greatly improved. After three sets of experimental results, the algorithm has a better performance.

ACKNOWLEDGMENT

This research was supported by National Natural Science Foundation of China (NSFC Grant No.61262080), Jiangxi Province Science and Technology Support Major Project (20151BBE50121, 20161BBE50086), Science and Technology Key Project of Jiangxi Province Education Department (GJJ150299). All support is gratefully acknowledged.

REFERENCES

- [1] Russell Newman, Victor Chang, Robert John Walters, Gary Brian Wills. Web 2.0—The past and the future[J], International Journal of Information Management, 2016,04(4):591-598.
- [2] Ruben Heradio, Luis de la Torre, Daniel Galan, Francisco Javier Cabrerizo, Enrique Herrera-Viedma, Sebastian Dormido, Virtual and remote labs in education: A bibliometric analysis[J], Computers & Education, 2016,06 (98):14-38.
- [3] Lin Shuangmei, Wang Dickinson, Chen Yiqiu. User modeling and feature selection in personalized recommendation system [J]. Computer Engineering ,2007,No.28517:196-198+230.
- [4] Sasithorn Chookaew, Patcharin Panjaburee, Dechawut Wanichsan, Parames Laosinchai, A Personalized E-Learning Environment to Promote Student's Conceptual Learning on Basic Computer Programming[J], Procedia - Social and Behavioral Sciences, 2014,02(116): 815-819
- [5] Wang Jing. Research on personalized learning based on collaborative filtering technology in E-learning system [D]. Shanghai Normal University,2015.
- [6] Wu Lihua, Luo Yunfeng, Feng Jianping. Personalized intelligent agent search engine for web based education [J]. Computer aided engineering,2007,No.6302:65-68.
- [7] Wang Rong. Research on Personalized Recommendation System of Education Website Resources [D]. North China University of Technology,2012.
- [8] Xia Peiyong. Research on collaborative filtering algorithm in personalized recommendation technology [D]. Ocean University of China,2011.
- [9] Nitin Pradeep Kumar, Fan Zhenzhen, Hybrid User-Item Based Collaborative Filtering[J], Procedia Computer Science, 2015(60):1453-1461
- [10] Leng Yajun, Lu Qing, Liang Changyong. Review on Collaborative Filtering Recommendation Technology [J]. Pattern recognition and artificial intelligence,2014,v.27;No.13408:720-734.
- [11] Kong Weiliang. Research on key problems of collaborative filtering recommendation system [D]. Huazhong Normal University,2013.
- [12] Jiang Zhihui. Model construction and strategy optimization of personalized learning under Network Environment [J]. China Distance Education,2013,02:48-51+95.

- [13] Liu Dehua, Zhang Peng, Zheng Qinghua, An efficient online active learning algorithm for binary classification[J], Pattern Recognition Letters, 2015,12(68):22-26, ISSN 0167-8655
- [14] Ding Yonggang, Zhang Xin, Sang Qiuxia, Zhang Hongbo. Fusion of collaborative filtering learning resources for learners' social networks [J]. Modern educational technology,2016,02:108-114.
- [15] Wang Yuxiang, Qiao Xiuquan, Li Xiaofeng, Meng Luoming. Research on context aware mobile social network service selection mechanism [J]. Journal of Computer Science,2010,11:2126-2135