THE PRIVATE RECOMMENDATION BASED ON THE ANALYSIS OF USER DYNAMIC BEHAVIOR

Yang Hongyan[1], Liu Qun[1]

[1]Chongqing Key Laboratory of Computational Intelligence , Chongqing University of Posts and Telecommunications, Chongqing 400065,China

Keyword: private recommendation, hot topic, dynamic behavior network, interest similarity

Abstract . The traditional private recommendation system always ignores user dynamic behaviors. In consideration of the problem, the private recommendation based on the analysis of user dynamic behavior is provoked. This method recommends interested users in their virtual communities which are identified in the dynamic behavior network. The dynamic behavior network is built and made up of micro blog hot topics, users and participation behavior relationships. Meanwhile, this method not only considers the short-term dynamic interest, but also takes long-term stability interest into account. In order to get the weighted similarity of interest, establish the long-term interest model and short-term interest model, and trade off their contribution rate. Finally, experiment is done on a micro blog data set, The results show that this method has good effect.

1. Introduction

In recent years, social network site(SNS) has been increasing in both size and services. It has brought convenience to people's life, and with it the great deal of information make user choose hard. This is where personal recommendation system come in. In the domestic and oversea, the most popular SNS is Micro blog services, such as Sina,Twitter and so on. They allow users to post a short message for sharing viewpoints and acquire knowledge in real time.

Now so many methods have been researched for personal recommendation, but not all of them apply for friend recommendation. For example, the collaborative filtering for the personal recommendation is by far the most widely used and the most successful personal recommendation technology. However it cant be used friend recommendation. Traditional personalized friend recommendation always uses static attribute information or topological structure information[1,2]. Those kinds of information are static and they are not sufficient conditions for friend recommendation. As the short text micro-blog users post reflects their views and interests, so now the content of micro blog is considered in recommendation system[3,4], but it still ignores user' behavior information. Here, to improve the accuracy, we combine the user' behavior information, the topic of micro blog and topology information together. Meanwhile, we divide user's interest into short-term dynamic interest and long-term stability interest. The recommendation procedure is described as follows:

Step1: We analyze the user participation behavior for micro-blog hot topic, establish the dynamic behavior network model DBN = {U, C, E}.

Step2: Establish short-term dynamic interest model EI = (E1, E2, E3), the interest set E1 describes the frequency of interest, E2 represents participation vitality of interest, E3 description consistency of participation period. We use formu1a 9 to compute the similarity of short-term interest,and get the value sim1 (u, v).

Step3: We use formu1a 10 to compute the similarity of user's long-term interest, and get the value sim2 (u, v);

Step4: Trade-off the contribution rate of Short-term interest and long-term interest to get the weighted similarity of interest sim (u, v) = αsim1 (u, v) + (1-α) sim2 (u, v), where α is the contribution rate.

2. User dynamic behavior analysis

There are various social relationships in SNS. Except friend and following relationships, there are participation relationships that are built by users and topics. The main ways for   users   to   take participate in topics are: Post, Retweet, Comment, Collect and Praise. If there are more participation actions between users and topics, the higher the degree of interest is. The dynamic behavior network is shown as follows:
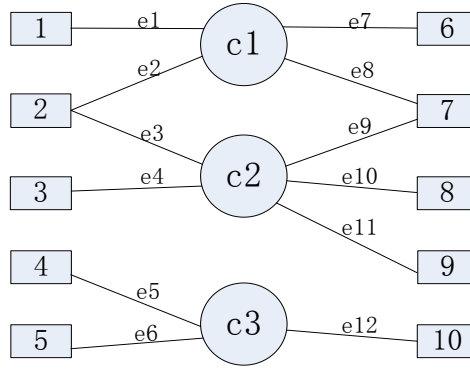

Fig. 1: An Example of Dynamic Behavior Network

Establish the dynamic behavior network DBN = {U, C, E}, where U = {u1, u2, u3 ...... un} is the set of user nodes, C = {c1, c2, c3 ...... cm} is the set of hot topic nodes, E = {e1, e2, e3 ...... en} is the set of edges, the edge e= (u, c, wuc) represents the participation relationship between a user node and a topic node, wuc is the weight of e, indicating that the user u' participation parameters to the topic c. For example, if u posts or retweets one micro blog which its content is about c, then wuc cumulative add degree of interest .

The user's degree of interest is divided into four levels[19]: 0.0, 0.33, 0.67, 1.00. Rules are shown as follows:

If there are not any participation actions between u and c, the degree of interest equals 0.

If u post or retweet a micro blog which its content is about c ,the degree of interest equals 0.67.

If u comment, collect or praise a micro blog which its content is about topic c, the degree of interest equals 0.33.

If u has the actions in rule 2 and rule 3,the degree of interest equals 0.33.

In figure 1, we can see directly that users gathered for virtual communities which constitute the dynamic behavior network. Users in the same virtual community are attracted by the virtual community topic. Since dynamic behavior network is made up of virtual communities, we can only consider the virtual communities which target user in as recommendation space(RS). When we recommend interested users for target user, we just do it in his recommendation space.

3. User interest similarity framework

3.1 Short-term dynamic interest Model

1) Frequency similarity
The behavior characteristics between u and c reflect the usser u is interested in the topic c. The interest frequency distribution of topic set C is expressed as follows: $S_u=\{s_1，s_2，s_3……s_m\}$, where s_i equals w_{uci}. Then the user interest set E1=$\{e_1=(c_1, s_1)，e_2=(c_2, s_2)，e_3=(c_3, s_3)……e_m=(c_m, s_m)\}$，The similarity of E1 between u and v is expressed as follows,

$$sim_{E1}(u,v) = \frac{1}{1+dis(u,v)}$$

(1)

Where dis(u,v) is euclidean distance of S_u and $S_v.$,

$$dis(u,v) = \frac{1}{m} \sqrt[2]{\sum_{i=1}^{m} \overline{q_i} \times (sui - svi)^2} \qquad (2)$$

Where q_i is contribution rate,

$$q_i = \frac{s_i}{\sum_{j=1}^{m} s_j} \qquad (3)$$

$$\overline{q_i} = \frac{q_{ui} + q_{vi}}{2} \qquad (4)$$

2) Vitality similarity

The interest set E1 is not enough to decide and reflect the user's short-term interest. If the participation frequency is not high enough during the hot topic existing time, but it is high in a short time, we consider that during the short time the user's interest to the hot topic is high. Assume that the more times the user participates in the hot topics in a unit interval, the bigger vitality that the user has. Then we can see that user's participation vitality set Fu={f1, f2, f3……fn}, where fi is expressed as follows,

$$f_i = \frac{s_i}{t_{i2} - t_{i1}} \qquad (5)$$

the user interest set E2={e_1=(c_1,f_1)，e_2=(c_2,f_2)，e_3=(c_3,f_3)······e$_m$=(c_m, f_m)}. The similarity of E2 between u and v is expressed as follows,

$$sim_{E2}(u,v) = \frac{\sum_{i=1}^{m} (f_{ui} - \overline{f_u})(f_{vi} - \overline{f_v})}{\sqrt{\sum_{i=1}^{m} (f_{ui} - \overline{f_u})^2} \sqrt{\sum_{i=1}^{m} (f_{vi} - \overline{f_v})^2}} \qquad (6)$$

3) Time Consistency

User interest change over time. It is very important that weather users are interested in the same topic at the same time. Then we can see that user's participation period set Tu={<t11,t12>，<t21,t22>, <t31,t32>······ <tm1,tm3>}, where ti1 is the earliest time and ti2 is the latest time when u join

in the topic i.The user interest set E3={e1=(c1, <t11,t12>)，e2=(c2, <t21,t22>)，e3=(c3, <t31,t32>) ··· ··· em=(cm, <tm1,tm3>)}. Consistency of participation period reflects the characteristic weather u and v take part in the same topic at the same time. It is expressed as follows,

$$\begin{cases} \Delta t_i(u,v) = 0 & [ut_{i1}, ut_{i2}] \cap [vt_{i1}, vt_{i2}] = \varnothing \\ \Delta t_i(u,v) = \dfrac{t_2 - t_1}{ut_{i2} - ut_{i1}} & [ut_{i1}, ut_{i2}] \cap [vt_{i1}, vt_{i2}] = [t_1, t_2] \\ \Delta t_i(u,v) = 1 & [ut_{i1}, ut_{i2}] = [vt_{i1}, vt_{i2}] \end{cases} \qquad (7)$$

where [$ut_{i1,}ut_{i2}$] is the period when u takes part in the topic i, $ut_{i1}<ut_{i2}$. The degree of E3 similarity between u and v is expressed as follows,

$$sim_{E3}(u,v) = \frac{1}{m} \sum_{i=1}^{m} \Delta t_i(u,v) \qquad (8)$$

4) Weighted degree of short-term interest

The traditional model of users' interest may ignore user's participation vitality and the consistency of time. Concerning to the feature of users' interest changing fast, we weight on the similarity of interest frequency, the similarity of interest vitality and the consistency of interest time.

Then we can get the similarity of short-term interest, shown as follows,

$$sim_1(u,v) = \sum_{i=1}^{3} \lambda_i sim_{Ei}(u,v) \tag{9}$$

where λ_i in the formula reflects the contribution rate of the similarity of interest set Ei to weighted degree of short-term interest, $\lambda_1+\lambda_2+\lambda_3=1$.

3.2 Long-term interest similarity

The user's behavior characteristics reflect his interest. And the relationship of attention among the users can also reflect the user's characteristics. The relation of attention will not easily break down as time flows once established in general, and it is a relatively stable relations. The relation between the target user u and the interest users can reflect u' long-term stable interest. The more common-focus users they have, the more similarity of long-term interest they have.

$$sim_2(u,v) = \frac{I_u \cap I_u}{max(I_u, I_u)} \tag{10}$$

Where I is the interested users set.

3.3 Weighted similarity of interest

User's interest depends on short-term dynamic user interest and long-term stable interest. We add linear weight to short-term interest similarity and long-term interest similarity to calculate the weighted interest similarity of users,

$$sim(u,v) = \alpha * sim_1(u,v) + (1-\alpha) * sim_2(u,v) \tag{11}$$

where $sim_1(u, v)$ in the formula is the short-term similarity of user u and user v, and $sim_2(u, v)$ stands for the long-term stability similarity. α $\in [0, 1]$, It Here, we get it according to the experimental test. Suppose u is the target user, and v is a recommended user, then v must be the node which is in the recommend space.

4. Experiment and analysis

4.1 Experimental data

In this paper, Sina micro-blog data is used, and it includes the following:

(1) 63641 Sina Micro-blog users information.

(2) 84,168 tweets with 12 topics are collected from 2014-04 to 2014-05. The 12 topics include Meizu, Xiaomi, Rockets, Jeremy Lin, Hengda, Korean, Haze, House prices, My Old Classmate, Civil Servants, Corrupt Officials, Transgenosis.

(3) 1,391,718 users' relationships between friends, each record consists of two parts: suid and tuid. The users relationship represents concern from suid to tuid.

4.2 Experimental evaluation criteria

In this paper, we adopt Precision, Recall, F-measure, Ranking Precision as Recommended evaluation.

(1) Precision: Pr=p/r. in which p represents the target users' interested users in recommended list. r represents all interested users to target users.

(2) Recall: Re=p/n, where p represents the target users' interested users in recommended list. n represents the number of interested users recommended.

(3) F-measure: F=2Pr*Re/(Pr +Re).

(4) Ranking Precision: RP=(1/n)*Σ(1/posi), where posi means that the ranking position in the user's interest recommended list. n represents the number of interested users in recommendation list.

4.3 Analysis of experimental results

In this experiment, the micro-blog data set is divided into five equal parts, of which a quarter of the training set, a test set. For each data set, we obtain Top-10 interested users. Meanwhile, we using the following recommend algorithm to compare.

(1) Common Neighbor (M1) is adopted to obtain a recommendation list.

(2) Character Similarity Calculation of Community Site Based on User's Interest Change[3] (M2)

is adopted to obtain a recommendation list.

(3) Personal recommendation based on the analysis of user behavior (M3) is adopted to obtain a recommendation list.
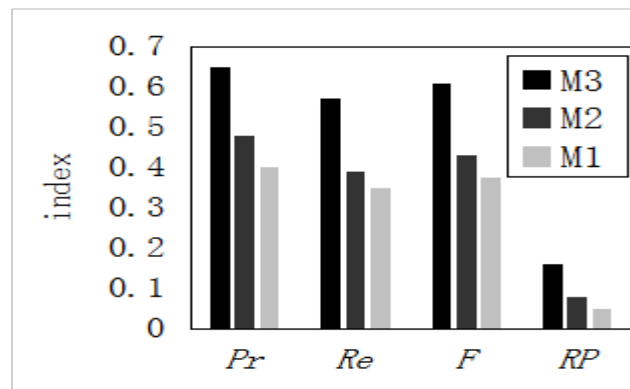

Fig. 2: Recommendation Indexes of *PR*,*Re* ,*F* And *RP*

As can be seen from Figure 2,M1 is relatively low. This is because that the selected recommended attributes are relatively simple, using only the friend relationship of micro-blog users (namely, mutual concern between users). Moreover, the recommendation indexes of M2 is higher than those of M1. This is because M2 constructs user interest model, but it does not take time consistency into account and ignores user' long-term exception. Its recommendation results should be improved. The paper presents a personal recommendation based on analysis of user behavior(M3) to make up for the shortage of M2 method, whose recommended results are the best. Considering a large number of micro-blog users and micro-blog user increment, the recommended range would be narrowed to the target user's virtual social group, which can improve recommended efficiency to some extent. After several tests verification, recommending effect would be optimal when α, representing the weight of weighted interest similarity, equals to 0.6.

5. Conclusion

In this paper, the dynamic behavior of the micro blog users was analyzed. Regarding the current hot topic of micro blog, we establish the users' dynamic behavior network and found the network's virtual community so as to build user's short-term dynamic interest model. While, we take advantage of long-term concern information to establish long-term stability interest model. Finally, the interested users in the recommendation space are recommended to target users. In this paper, user behavior information, the content of micro blog and topology information are used to improve the accuracy and the experimental results show that recommendation results are better than tradition methods. In the next study, more characteristics could be used and the weight α can be computed by the recommendation system rather than can repeat test experiment.

Acknowledgments

References

[1] Wu Zhipeng, Jiang Shuqiang, Huang Qingming. Friend Recommendation According to Appearances on Photos[C].the seventeen ACM international conference on Multimedia, China, 2009:987-988.
[2] Akbari F, Tajfar A H, Nejad A F. Graph-Based Friend Recommendation in Social Networks Using Artificial Bee Colony[C]//Dependable, Autonomic and Secure Computing (DASC), 2013 IEEE 11th International Conference on. IEEE, 2013: 464-468.

[3] ZHANG XiaoBin, PANG Haiyan. Character Similarity Calculation of Community Site Based on User's Interest Change[J]Computer Systems & Applications,2013, 22(2): 112-115.

[4] KHATER S, ELMONGUI H G, GRACANIN D, Personalized microblogs corpus recommendation based on dynamic users interests[C]// Proceeding of the 2013 International Conference on Social Computing. Piscataway:IEEE, 2013: 979-982.

[5] ZHU S.Personal recommendation based on micro-blogging references[D]. Beijing:Beijing University of Post and Telecommunications, 2013: 39-40.

[6] YAN Q, WU L, ZHENG L. Social network based microblog user behavior analysis[J]. Physica A: Statistical Mechanics and its Application, 2013, 392(7): 1712-1723.

[7] WANG Weiwei, XIA Xiufeng, LI Xiaoming. Personal interest degree model base on consumer behavior. Computer Engineering and Application, 2012,48(8):148-151.

[8] Sharma, A., Gemici, M., Cosley, D.: Friends, strangers, and the value of ego networks for recommendation. In: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, pp. 721–724 (2013)

[9] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann. Time-aware point-of-interest recommendation. In SIGIR, pages 363–372, 2013.

[10] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann. Who, where, when and what: discover spatio-temporal topics for twitter users. In KDD, pages 605–613, 2013.

[11] W. Zhang, J. Wang, and W. Feng. Combining latent factor model with location features for event-based group recommendation. In KDD, pages 910–918, 2013.