Community Semantics Recommender Based on Association Link Network

Yang Liu and Xiangfeng Luo

Abstract Community semantics recommender offers the semantic communities for topically homogeneous terms with the user demand, which helps people to acquire personalized information in mass documents and saves users' cognitive energy significantly. However, most of current recommendation models merely focus on users' behavior (e.g., item purchase behavior or user browsing traces) while pay less attention on the knowledge itself, especially the community semantics. In this paper, we present a community semantics recommender based on Association Link Network. We organize and represent the knowledge by Association Link Network (ALN). Given a user demand, we first select topically relevant community based on ALN; then we conduct community semantics activation process on the activated community to produce the topically homogeneous terms. Experimental results demonstrate our model can offer the user topically homogeneous terms and help them to understand the implicit knowledge that hidden in the large scale of documents.

Keywords Community semantics · Recommender system · Association link network

1 Introduction

The knowledge recommendation offers users the knowledge service [1], such as refining the implicit fuzzy demand and extending the domain knowledge, which not only saves users' effort for complex information seeking but also improves users' experience significantly. Community semantics recommender provides the

Y. Liu · X. Luo (✉)
School of Computer Engineering and Science, Shanghai University, Shanghai, China
e-mail: luoxf@shu.edu.cn

Y. Liu
e-mail: nmgliuyang@shu.edu.cn

N.Y. Yen and J.C. Hung (eds.), *Frontier Computing*, Lecture Notes in Electrical Engineering 422, DOI 10.1007/978-981-10-3187-8_5

semantic communities, which conveys topically homogeneous knowledge to meet users' demands.

Current researches on knowledge recommender mainly focus on recommending documents by users' behavior. The two major recommender models are the collaborative-filtering model [2, 3] and the content-based model [4, 5]. The former explicits relevance feedback from users to update the user profiles [6], which is widely used in the e-commerce sites like Amazon and Taobao. The latter monitors document stream and pushes documents that match a user profile to the corresponding user [7, 8]. Above two kinds of recommender models have two commons: (1) recommend items/documents that are similar to what a user liked in the past; (2) recommend the items/documents which similar users group likes as target user's potential demand. Above two recommender models can hardly provide users their interested knowledge. Another widely used recommender for documents is the general search engines such as Google and Baidu, which recommend the documents that contain the terms in user's queries. The search engines offer different users same document list, if they have the same queries. Current recommenders seldom pay attention on the semantics of knowledge themselves, especially the community semantics, which leads the recommended results having abundant knowledge.

For these reason, this paper proposed an Association Link Network based community semantics recommender, which simulates the recommendation process of human memory based on the study of cognitive science. Specifically, we employ the Association Link Network (ALN) [9], an automatic built semantic link network [10], to organize and represent the knowledge of the documents. Based on ALN, we implement spreading activation [11, 12] model to simulate the knowledge recommendation process of human memory. Given a user demand, we first select topically relevant community based on ALN; then we activated community to produce the topically homogeneous terms. Therefore, our recommender not only can provide the user the community semantics for a better understanding of knowledge scenario, but also offer topically homogeneous terms.

The rest of the paper is organized as follows. The related work is discussed in Sect. 2. Section 3 describes community semantics recommender based on ALN. In Sect. 4, we demonstrate the effectiveness of our approach with comparative empirical results. Finally, we present the conclusions and point out future work in Sect. 5.

2 Related Work

Recommender is changing the way people interact with the Web and being more and more important. The recommender predicts the utility or relevance of the item to a particular user. Considering the different types of information used, the recommender can be divided into content-based recommender and collaborative filtering recommender.

The content-based recommender [13] recommends the documents which are similar to documents in the corresponding user's profile. Content-based recommender creates content-based user profiles using a weighted vector which denote the user feature. User profile can be computed by user's rating score using a variety of techniques such as Bayesian classifiers [14], decision trees [15], and neural networks [16] to estimate the probability that the user like the documents.

Collaborative filtering recommender [17] is based on the opinions of other users. The user ratings of documents are stored and can be used to predict ratings of documents which have not been rated by corresponding users. The user-based collaborative filtering model [18] finds the similar users according to users' rating history and then analyzes similar users' profile to predict the favor of target user. Different metrics have been used to calculate the similarity between users, such as mean squared difference [18], vector similarity [19] and weighted Pearson [20]. Item-based collaborative filtering model [21] finds the similar items by the items history and recommends the similar items to users. The item based model is better than the user based model because the similarity between items tends to be more static than the similarity between users. Taking the advantage of both item-based and user-based collaborative filtering models, a similarity fusion algorithm is proposed that has higher computational-complexity than both two algorithms [22].

Current recommenders pay less attention on the knowledge, especially the community semantics. Content-based recommenders only consider the similarity of items while users' need for association cannot be meeting. The collaborative filtering doesn't consider the characteristic of knowledge itself. The semantic communities are topically homogeneous terms, which help users to obtain personalized knowledge from mass documents and save their cognitive energy. A user friendly knowledge recommender should have following features: (1) Recommender should provide the user the community semantics for a better understanding of knowledge scenario. Human will easily understand the topically relevant knowledge, extend the domain knowledge, and solve specific knowledge questions. (2) It should offer the user topically homogeneous terms, which are relevant terms related to user demand and provide extra choices to users for further interaction. In this paper, we present a community semantics recommender system, where we provide the users relevant semantic communities as well as topically homogeneous terms.

3 Community Semantics Recommendation

In this Section, we will propose the ALN based community semantics recommender system. As Fig. 1 depicted, the framework of our community semantics recommender has two phases: (1) community selection phase, which provides the user the community semantics for a better understanding of knowledge scenario; (2) community semantics activation phase, which offers the user topically homogeneous terms and provide extra choices for the next recommendation interaction. Above two phases are based on the ALN. We employ ALN to organize the

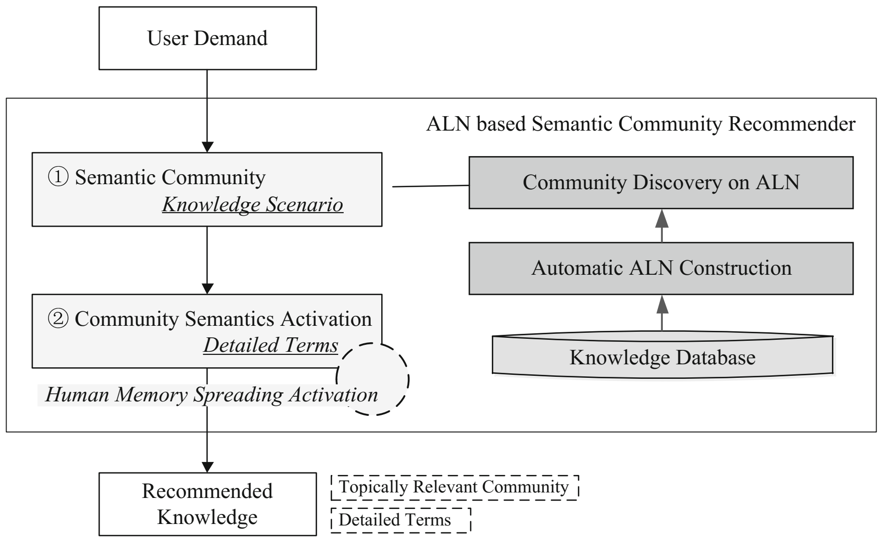

Fig. 1 The framework of ALN based community semantics recommender

semantics of knowledge database by association linked terms. Then, we discover communities on ALN, which can be used in community selection phase. We will detail each phase in the following.

3.1 Automatic ALN Construction

We employ ALN, which is an automatic built semantic link network [9], to organize the association linked nodes. Human knowledge retrieval is naturally directed. We construct ALN using significant terms, and the formula of computing association strength between term a and term b is as follows:

$$w^{a-b} = \frac{Co(a,b)}{\sqrt{DF(a)DF(b)}} \tag{1}$$

where $Co(a,b)$ represents the total co-occurrence times that term a and term b appear in one document. $DF(a)$ is occurrence times of word a in the document set.

The terms correspond to knowledge nodes in the human memory network. Term association has a value in the range [0, 1] to represent semantic related strength which corresponds to the strength between concepts in human memory. Thus, the generated ALN could be a knowledge network to simulating human domain knowledge, and based on it, we offer the user topically homogeneous terms.

ALN is able to organize the associated resources loosely distributed in the knowledge database for effectively supporting the recommender intelligent activities.

3.2 ALN-Based Community Discovery

After constructing ALN using the terms of knowledge database, we discover the community on ALN. The terms in a community are densely linked, and they have strong relationships between each other. We employ community detection algorithm to discover the community on ALN.

Enlightened by the thought of label propagation [23], each node is initialized with a unique label and at every step each node adopts the label that most of its neighbors currently have. In this iterative process, densely connected groups of nodes form a consensus on a unique label to form communities. Consequently, densely connected groups reach a common label quickly. The detailed process is as follows [23]:

(1) *Label Initialization*: we initialize the nodes' labels on ALN, and give each node on the ALN a unique node label.
(2) *Label propagation*: we do label propagation in iteration. For the iteration, each node adopts a label that a maximum number of its neighbours have, with ties broken uniformly randomly. As the labels propagate through the network in this manner, densely connected groups of nodes form a consensus on their labels.
(3) *Community Generation*: at the end of label propagation process, nodes with the same labels are grouped together as the communities discovered on ALN.

The advantage of this algorithm over the other methods is its simplicity and time efficiency. The algorithm uses the network structure to guide its progress and does not optimize any specific chosen measure of community strengths [23]. Furthermore, the community number and their sizes are not known a priori, and can be determined at the end of the algorithm. We conduct community discovery on ALN and provide the topically relevant community to users for a better understanding of knowledge scenario. And human will easily understand the topically relevant knowledge, which extends users' domain knowledge and solves specific knowledge questions.

3.3 Community Selection for Knowledge Scenario

For the user query, we select the topically relevant community to users, which gives the user comprehensive knowledge scenario for their understanding. A community can explain the topics of user demands, which offer the user knowledge context to extend their domain knowledge and solve the specific knowledge questions.

We select the topically relevant community for user demands based on ALN. The community is discovered based on the label propagation process on ALN. We view the terms with the largest degrees as well as association link weights as the most central terms for this community. We activate the most topically relevant community through computing the similarity between the user query and community:

$$I_c = \sum_{a \in (c \cap Q)} \frac{\sum w_c^{a-}}{\sum w_c} \tag{2}$$

where term a is the terms of user query. $\sum w_c^{a-}$ is the total association weight between term a and its adjacent terms on community c. $\sum w_c$ is the sum of all association weights of community c. We select the community with the max I_c value as the knowledge scenario for users.

3.4 Community Semantics Activation for Detailed Terms

After selecting the topically relevant community with user demands, we will offer the user topically homogeneous terms from the community based on the based on the spreading activation of human memory. The spreading activation of human memory tells the way that how human recommend the knowledge in human memory. We provide the relevant terms related and offer extra choices to users for further interaction.

In the human memory spreading activation process, user demand is the source of activation, and we conduct the spreading activation on the ALN-based topically relevant community.

(1) *Energy Initialization*: We give the terms of user queries an activation value represent their activation energy. And other terms of the community are zero.
(2) *Spreading Activation*: The activation energy is spreading in the community to get the topically homogeneous terms. After the spreading activation process, higher value represents user's higher focus.
(3) *Term Generation*: Terms with higher activation energy value will be generated as the topically homogeneous terms for the community semantics.

4 Experiments

In this section, we will make evaluations of the proposed community semantics recommender. We downloaded 1,222 documents from Tencent new channel as the knowledge database. Then, we construct ALN based on this dataset and perform community discovery on the ALN. We discovered 19 communities from the ALN.

Table 1 The Results of Community Semantics Recommendation

Query terms	Recommendation
Arizona, student, shooting	*Topically Relevant Community (Keywords of the Community)*: Shooting, U.S., Obama, Arizona, suspect, death, Tucson, condemn, college, American, head, police, blood, kill, federal, charge
	Recommended Detailed Terms: Obama, suspect, condemn, college, death
Iran, earthquake, Azerbaijan	*Topically Relevant Community (Keywords of the Community)*: Iran, earthquake, death, Azerbaijan, rescue, hospital, China, injure, Armenia, aftershock, shelter, damage, town, Ahar, emergency
	Recommended Detailed Terms: injure, damage, China, Armenia, rescue
Nigeria, flight, crash	*Topically Relevant Community (Keywords of the Community)*: Nigeria, China, death, Dana, passenger, flight, crash, disaster, people, aircraft, fire, damage, Lagos, reason, airframe, time
	Recommended Detailed Terms: China, Dana, passenger, disaster, aircraft
U.S, hurricane, Sandy	*Topically Relevant Community (Keywords of the Community)*: Sandy, storm, U.S, hurricane, flood, Obama, victim, weather, power, people, kill, wind, electric, speed, hospital, Cuba, town
	Recommended Detailed Terms: Victim, Obama, electric, power, power

Table 1 shows parts of recommendation results of our recommender system. Given a user query, our recommender first selects topically relevant community based on ALN, where we show the keywords of the relevant community. The recommended community provides users the knowledge scenario for a better understanding of relevant semantics. For example, the topically relevant community of the query “Arizona, student, shooting” can well explain the context how the Arizona shooting happens. Secondly, we recommend the detailed terms that are related to user demand, and provide extra choices to users for further interaction. For example, the detailed terms of the query “Arizona, student, shooting” describes some detailed information about the “Arizona shooting”.

5 Conclusion

In this paper, we present a community semantics recommender, which focuses on the knowledge itself and provides the users knowledge including topically relevant community and related terms. The former provides the comprehensive knowledge scenario for user understanding. The latter offers relevant terms representing detailed community semantics. We employ Association Link Network (ALN) to organize and represent the knowledge. Based on the spreading activation of human memory, we activate topically relevant community and related semantics, and recommend them to users, which can better serve users and will be widely used in the field of e-learning and web service.

Acknowledgements The research work reported in this paper is supported by the National Science Foundation of China (grant no. 61471232), and the Key Innovation Program of Shanghai Municipal Education Commission (grant no. 13ZZ064).

References

1. P. Resnick, and H. R. Varian, "Recommender Systems," Communications of the ACM, vol. 40, no. 3, pp. 56–58, 1997.
2. J. Bobadilla, F. Serradilla and J. Bernal, "A new collaborative filtering metric that improves the behavior of recommender systems," Knowledge-Based Systems, vol. 23, no. 6, pp. 520–528, 2010.
3. Y. Koren, and R. Bell, "Advances in collaborative filtering", Recommender Systems Handbook, Springer, pp. 145–186, 2011.
4. P. Lops, M. de Gemmis and G. Semeraro, "Content-based recommender systems: State of the art and trends," Recommender Systems Handbook, Springer, pp. 3–105, 2011.
5. R. J. Mooney, and L. Roy, "Content-based book recommending using learning for text categorization," Proceedings of the fifth ACM conference on Digital libraries, pp. 195–204, 2000.
6. Y. Zhang, and J. Koren, "Efficient Bayesian and Hierarchical user Modeling for Recommendation Systems," Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 47–54, 2007.
7. J. Bian, A. Dong, X. He, S. Reddy, and Y. Chang, "User Action Interpretation for Online Content Optimization," IEEE Transactions on Knowledge and Data Engineering, vol. 25, no. 9, pp. 2161–2174, 2013.
8. M. J. Pazzani, and D. Billsus, "Content-based recommendation systems," The adaptive web, Springer, Berlin Heidelberg, pp. 325–341, 2007.
9. X. F. Luo, Z. Xu, and J. Yu, "Building association link network for semantic link on web resources," IEEE Transactions on Automation Science and Engineering, vol. 8, no. 3, pp 482–494, 2011.
10. Z. Xu, X. F. Luo, and W. Lu, "Association link network: An incremental semantic data model on organizing web resources," IEEE 15th International Conference on Parallel and Distributed Systems, pp 793–798, 2009.
11. A. M. Collins, and E. F. Loftus, "A spreading-activation theory of semantic processing," Psychological review, vol. 82, no. 6, pp. 407–428, 1975.
12. J. R. Anderson, "A spreading activation theory of memory," Journal of verbal learning and verbal behavior, vol. 22, no. 3, pp. 261–295, 1983.
13. P. W. Foltz and S. T. Dumais, "Personalized information delivery: An analysis of information filtering methods," Communications of the ACM, vol. 35, no. 12, pp 51–60, 2000.
14. C. Chen, L. F. Pau, and P. S. Wang, "Handbook of pattern recognition and computer vision," Imperial College Press, 2010.
15. N. Golbandi, Y. Koren, and R. Lempel, "Adaptive bootstrapping of recommender systems using decision trees," Proceedings of the fourth ACM international conference on Web search and data mining, pp. 595–604, 2011.
16. S. S. Park, K. K. Seo, and D. S. Jang, "Expert system based on artificial neural networks for content-based image retrieval," Expert Systems with Applications, vol. 29, no. 3, pp. 589–597, 2005.

17. U. Shardanand, and P. Maes, "Social information filtering: algorithms for automating 'word of mouth'," Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, pp. 210–217, 1995.
18. P. Resnick, N. Iacovou, and M. Suchak, "GroupLens: an open architecture for collaborative filtering of netnews," Proceedings of the 2001 ACM conference on Computer supported cooperative work. ACM, pp. 175–186, 2001.
19. J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., pp. 43–52, 1998.
20. J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. Riedl, "Evaluating collaborative filtering recommender systems," ACM Transactions on Information Systems (TOIS), vol. 22, no. 1, pp 5–53, 2004.
21. B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," Proceedings of the 10th international conference on World Wide Web. ACM, 285–295, 2001.
22. J. Wang, A. P. De Vries, M. J. T. Reinders, "Unifying user-based and item-based collaborative filtering approaches by similarity fusion," Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, pp. 501–508, 2006.
23. U. N. Raghavan, R. Albert and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," Physical Review E, Vol. 76, Issue. 3, 2007.