

Framework for Social Relationship Recommendation

Jia Chen^{1,2,*}, Tongge Xu³ and Zhang Xiong^{1,2}

¹The State Key Laboratory of Software Development Environment, Beihang University, Beijing 100191, China

²Research Institute of Beihang University in Shenzhen, Shenzhen 518057, China

³Lab 673, School of Computer Science and Engineering, Beihang University, Beijing 100191, China

Abstract—Find out the potential favorite data is the essence of recommendation techniques, which facilitates the recommendation techniques become a vital issue in the big data researches and applications. The traditional recommendation research and application centers on the commercial items and resources. We apply the recommendation to social relationships in social networks. We design a framework which can analyze and recommend various types of relationships. The framework has four modules and three work modes. Furthermore, we design a hybrid recommendation approach in the framework for the balancing between accuracy and efficiency. The framework can recommend the social relationships with high accuracy and low cost.

Keywords—big data; framework; recommendation; social relationships

I. INTRODUCTION

The world has entered an era of big data and complex systems. Users are inundated by these big data and their relationships from various industrial applications [1]. Therefore, research on recommender systems has been proposed since the early 1990s, which can find out users' potential favorite items from big data by analyzing their historical data. By applying the theoretical research on recommender related techniques into complex data mining, recommendation have been used widely for internet commercial applications and improve the commercial profits significantly [2, 3].

The applications of recommendation have be categorized into eight main categories for the common commercial applications: e-government, e-business, e-commerce/e-shopping, e-library, e-learning, e-tourism, e-resource services and e-group activities [4]. The recommended commercial products include movies, songs, news, books, etc. The recommended resources include online lessons, favorited restaurants and favorite tour routes. With the personalized recommendation applications, the potential interested items are found out from the ocean of big data and recommended to the users.

Accompanied by the successful recommendations in big data mining, various fields have become to introduce and apply the recommendation, i.e., QoS analysis systems, users and user-user trust networks in social network-based services, and proteins and protein-protein interactomes in bioinformatics [5-7]. Among these new application fields, we also apply the recommendation in the social relationship data. The social relationship recommendation can find out the potential interesting social relationships in numerous social data, which

help users to find out various interesting relationships belongs to different persons.

Similar to other recommendation, our framework for social relationship recommendation should resolve the common issues, i.e., cold start problem, the balance between recommendation efficiency and accuracy, etc.. Besides these common considerations, we also need to consider some issues especially for social relationship recommendation. For example, social relationships belong to different persons. Also social relationships can be divided into many types, i.e., classmate relationships, colleague relationships, family relationships etc.

In this work, we aim to construct a framework to resolve the above issues and provide personalized recommendation on the social relationship with high accuracy. The main contributions of this paper include:

1) We have constructed a recommendation framework which has 4 modules and 3 work modes. This framework can adapt the characteristics of social relationships and recommend the social relationships per different requirements.

2) We have an overall analysis of the social relationships, which includes defining the related attributes like time and impacts. Furthermore, we give the categories of social relationships including but not limiting with the common friend relationships.

The rest of this paper is organized as follows. Section 2 gives the related works. Section 3 introduces the high-level framework design. Section 4 provides the detailed design of the framework. Finally, Section 5 concludes this paper.

II. RELATED WORKS

Various recommendation frameworks have been proposed and used for different applications successfully. However, to our best knowledge, no previous research work investigates on the analysis and recommendation for various types of social relationships. Fortunately, some of these works focus on recommending friends to users on social networks, which are so called friendship recommendation.

In [8], the framework for recommending potential friends to others in social networks has been constructed. The framework calculates the interpersonal similarity by analyzing the demographic, activity and some information about their relationships.

A friend recommendation framework has been proposed by [9] based on the user behaviors on particular social networks. The proposed framework measures the frequency of the

activities, classify the user behaviors and then apply multilayer thresholds for friend recommendation.

People recommendation application has been designed in [10], which works on an aggregated social network containing various resources. This work prove that the effective people recommendation can increase the number of social connections.

In [11], a novel visual system has been proposed to support users to explore and find friends interactively. The recommendation framework use both semantic structure of activity data and topological structure of social networks to find out the potential friends.

A unified framework can recommend items, groups and friends also attract our interests, which integrate three different types of recommendation in a single framework. In [12], each type of recommendation is promoted by the other two information resources, which improve the recommendation accuracy effectively.

All these frameworks provide simple friend relationship recommendation in the social network. In this paper, we focus on the framework to recommend various social relationships, i.e., classmate relationships, colleague relationships, family relationships etc. The users not only focus on different persons but also different relationships.

Similar to these works, we also emphasize on persons and their relationships. Furthermore, our framework needs to investigate the various types of social relationships, rather than the friendship relationship only. Therefore, the relationship analysis and recommendation approach become more complex, which leads specific recommendation framework requirements.

III. THE FRAMEWORK DESIGN

We design the social relationship recommendation framework by 1) designing the functionality and interfaces of modules, 2) analyzing the input and output data, 3) designing the work modes for different requirements and input data, 4) designing the hybrid recommendation approach and choosing the appropriate algorithms, 5) resolving the cold start problem

and other considerations when designing the models. Based on these 5 steps, recommendation framework for social relationships can be constructed.

The designed framework consists of 4 modules: 1) User data collection and profile construction; 2) Social relationship analysis; 3) Hybrid social relationship recommendation approach; 4) Recommendation filtering and representation. The functions and work flows in these models are depicted in. FIG. I.

As depicted in FIG. I., user attributes and behaviors in module 1, social relationships in module 2 are two input data of the framework. And the 1st and 2nd modules focus on the data collection and transformation. Recommendation representation data collection in module 4 is the output data of the framework. The module 3 is the main analysis and recommendation module, which uses different recommendation algorithms in different work modes.

The recommendation framework has three work modes: initialization mode, offline work mode and online work mode.

The initialization mode runs in recommendation initialization phase.

The offline mode runs periodically. All the new records and updates are checked and analyzed. All the users are re-analyzed and the whole rating matrices are recalculated. The online work mode gives the recalculated recommendation lists to all the users.

The online mode runs when some users need the real time recommendation. For the registered user, online mode only updates the relevant matrix, and recommend algorithm only calculate the updated section. This scheme satisfies the real time recommendation requirements and with low cost.

For new user, KNN similarity algorithm will be applied for recommendation. The KNN algorithm also does not consume much cost.

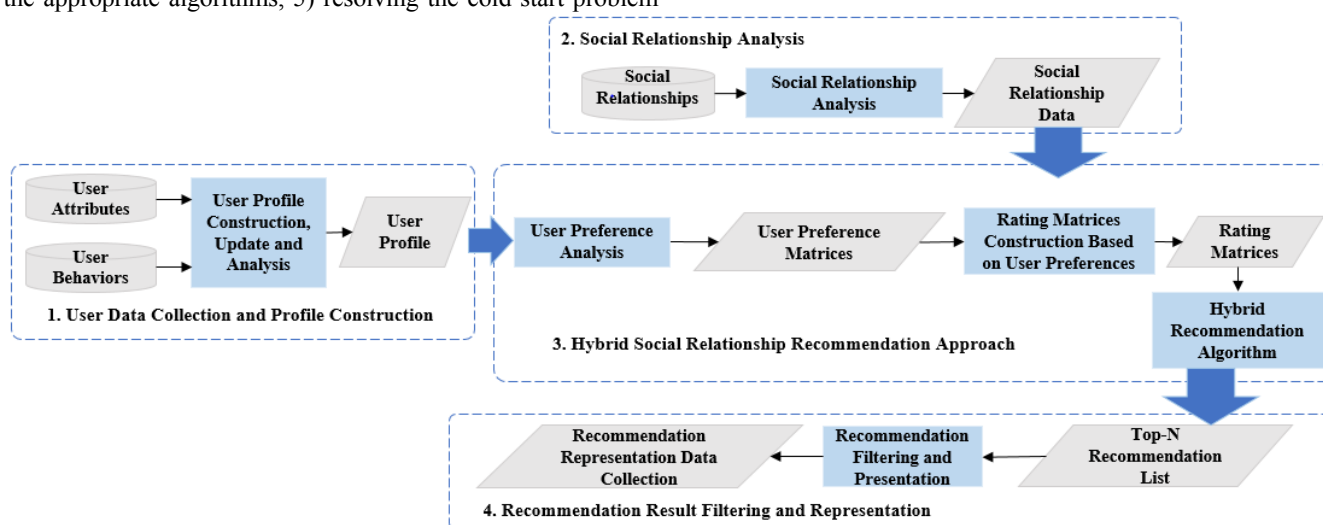


FIGURE I. THE FRAMEWORK OF SOCIAL RELATIONSHIP RECOMMENDATION

IV. DATA ANALYSIS AND HYBRID RECOMMENDATION ALGORITHM

The input data consists of user data and social relationship data. The output is the recommendation representation data collection. The hybrid recommendation algorithm analyzes the user data to determine the user preference type and then employs different algorithms. In this paper, we focus on the input data analysis and hybrid approach introduction, which improve the accuracy of data mining and recommendation.

A. User Data Analysis

User data consists of user attributes and user behaviors. User attributes are filled by the new users when they log in the application in the first time. Unless users modify their data, their attributes will be unchanged. Meanwhile, users' operations are monitored in the application and recorded in user profile as user behaviors.

We give symbols \mathbf{PR} to denote the user profile set, \mathbf{PR} consists of all the user profiles,

$$\mathbf{PR} = \{\mathbf{PR}_1, \mathbf{PR}_2, \dots, \mathbf{PR}_M\}. \quad (1)$$

where M is the number of users. M increases continuously with new users added in the application.

\mathbf{PR}_m is the user profile of u_m , $m \in M$, which consists of attributes and behaviors, formulated by

$$\mathbf{PR}_m = \{\mathbf{Atr}_m, \mathbf{Bh}_m\} \\ = \{\{Atr_1^m, Atr_2^m, \dots, Atr_{NA}^m\}, \{bh_1^m, bh_2^m, \dots, bh_W^m\}\} \quad (2)$$

where \mathbf{Atr}_m is the attribute vector of user u_m , $\mathbf{Atr}_m = \{Atr_1^m, Atr_2^m, \dots, Atr_{NA}^m\}$, NA is the number of attributes predefined in application. Each user has NA attributes. \mathbf{Bh}_m is the behavior vector of user u_m , $\mathbf{Bh}_m = \{bh_1^m, bh_2^m, \dots, bh_W^m\}$. Each user has W behaviors, which is varying for different users.

Each behavior recorded in the profile falls into a certain behavior type. For a specific behavior bh_w^m , we explain it as follow,

$$bh_w^m = [bh_id., ob_p(u_m, j), kb(u_m, j), \\ t(u_m, j), ob_s(u_m, j), im(u_m, j)], \quad (3)$$

where j is the behavior number for u_m , $bh_id.$ is the automatic generated number for user behaviors, $ob_p(u_m, j)$ is the person pointed by the behavior, $kb(u_m, j)$ is the behavior type, $t(u_m, j)$ is the happen time of behavior, $ob_s(u_m, j)$ is the social relationship pointed by the behavior, $im(u_m, j)$ is the behavior impact.

B. Social Relationship Analysis

We give the symbol \mathbf{SR} to denote the social relationship set

and list all the symbols involved in Table I.

Each specific social relationship belongs to P_n can be defined as,

$$\mathbf{SR}_s^n = \{ob(P_n, ns), tp(P_n, ns), im(P_n, ns), tp(P_n, ns), dt(P_n, ns)\} \quad (4)$$

TABLE I. THE SYMBOLS INVOLVED IN SOCIAL RELATIONSHIP ANALYSIS

Parameter	Description
TS	Number of social relationship categories
\mathbf{SR}	Set of social relationships, $\mathbf{SR} = \{SR_{ts}, ts=1,2,3,\dots, TS\} = \{SR_1, SR_2, SR_3, \dots, SR_{ts}\}$
\mathbf{P}	Dataset of persons, each relationship belong to one person, $\mathbf{P} = \{P_n, n=1,2,3,\dots,N\} = \{P_1, P_2, P_3, \dots, P_N\}$, N is the number of persons.
\mathbf{SR}_n	Dataset of social relationships belonged to person P_n , $\mathbf{SR}_n = \{SR_1^n, SR_2^n, \dots, SR_S^n\}$, where S is the number of each person's social relationships, S is different for different person, and increases with new relationship added.
\mathbf{SR}_s^n	A sth social relationship record for person P_n .

where $ob(P_n, ns)$ is the objective of \mathbf{SR}_s^n . $tp(P_n, ns)$ is the social relationship type of \mathbf{SR}_s^n , $tp(P_n, ns) \in \mathbf{SR}$. $im(P_n, ns)$ is the strength, $t(P_n, ns)$ is the start time, $dt(P_n, ns)$ is the duration time.

All the parameters of the social relationship except $im(P_n, ns)$ can be extracted from the known personal information. When the recommendation framework runs, the latest relationship data will be added into dataset by grabbing from Internet.

Each relationship has its own lifetime. Some relationships happened in the past and does not exist now. Although these relationships disappeared, we still need to record them and represent them to the users. All the relationships in the dataset can't be deleted.

C. Hybrid Recommendation Approach

When design the recommendation approach in the framework, we need to consider the following aspects: 1) User attributes and behaviors can be considered as the hidden preference to the social relationships, so the user preference matrix to social relationships can be constructed based on behavior data; 2) the preference matrix is the high dimension sparse matrix, ELBLF based CF recommendation algorithm can be used for its high accuracy and efficiency [13]; 3) for new relationships and new users, the cold start problem should be considered.

Based on the above considerations, we design a hybrid recommendation approach based on ELBLF and KNN algorithm in the framework. The KNN algorithm based on user attributes or relationships is used for resolving cold start problem, while ELBLF algorithm is the main recommendation algorithm.

The input data of ELBLF algorithm is rating matrix, which

is constructed by calculating the preference matrices for all the users based on their attributes and behaviors.

We define β_m as the influence parameter to combine the attributes and behaviors, which mainly considers the behavior weights and frequency in a period. In offline work mode, all the behaviors happened after T_{snt} need to be calculated. While, in online mode, only the behaviors happened after $T_{nt} (< T_{snt})$ need to be calculated. The reason is that only the nearest happened behaviors should be considered for the online real time recommendation.

$$\beta_m = \frac{\sum_{\forall T(u_m, j) \leq T} wkb(kb(u_m, j))}{N \cdot TS}, T = \begin{cases} T_{nt} & \text{online} \\ T_{snt} & \text{offline} \end{cases} \quad (5)$$

where $wkb()$ is the weight factor of behavior type. Then, we construct u_m 's preference matrix by combining the attribute and behavior preference matrices together. The attribute preference matrix of user u_m to all the social relationships belonged to all the persons can be defined as,

$$RA_m = \begin{bmatrix} ra_{1,1}^m & ra_{1,2}^m & \cdots & ra_{1,TS}^m \\ ra_{2,1}^m & ra_{2,2}^m & \cdots & ra_{2,TS}^m \\ \vdots & \vdots & \ddots & \vdots \\ ra_{N,1}^m & ra_{N,2}^m & \cdots & ra_{N,TS}^m \end{bmatrix} \quad (6)$$

In (6), each row in the matrix represents the u_m attribute preference to all the TS types relationships belonged to one person.

The behavior preference matrix of user u_m to all the social relationships belong to all the persons can be defined as,

$$RB_m = \begin{bmatrix} rb_{1,1}^m & rb_{1,2}^m & \cdots & rb_{1,TS}^m \\ rb_{2,1}^m & rb_{2,2}^m & \cdots & rb_{2,TS}^m \\ \vdots & \vdots & \ddots & \vdots \\ rb_{N,1}^m & rb_{N,2}^m & \cdots & rb_{N,TS}^m \end{bmatrix} \quad (7)$$

In (7), each row in the matrix represents the u_m behavior preference to all the TS types relationships belong to one person.

The preference matrix of u_m is represented as,

$$R_m = (1 - \beta_m) RA_m + \beta_m RB_m \quad (8)$$

For all the users, the preference matrix R as the input for EBELF recommendation algorithm, can be represented as,

$$R = [vec(R_1) \quad vec(R_2) \quad \cdots \quad vec(R_M)]^T \quad (9)$$

Then, we use R as the input matrix for EBELF algorithm. The detailed training process and theoretical formula of

EBELF algorithm have been represented in [13].

The result of the EBELF algorithm is the estimated rating matrix of all the users. All the missed elements in R are filled in the estimated rating matrix, which represents the potential user preference to the social relationships.

V. CONCLUSIONS

In this paper, we design a framework for the social relationship recommendation. The framework consists of 4 modules and 3 work modes. The hybrid recommendation approach based on ELBLF and KNN algorithms are applied for different recommendation requirements. We plan to introduce the influence of time-variance, user preference types and relationship strength into the framework in our future research.

ACKNOWLEDGMENT

This work was partially supported by the National Natural Science Foundation of China (No. 61472021, 61502320), SKLSDE project, and the Science Foundation of Shenzhen City in China (JCYJ20140509150917445).

REFERENCES

- [1] V. Mayer-Schonberger and K. Cukier, "Big Data: A Revolution That Will Transform How We Live, Work and Think," Houghton Mifflin Harcourt Publishing Company, 2013.
- [2] P. Resnick and H. R. Varian, "Recommender systems," Communications of the ACM, vol. 40, no. 3, pp. 56-58, 1997.
- [3] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," IEEE Trans. Knowledge and Data Engineering, vol. 17, pp. 734-749, Jun. 2005.
- [4] Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G. "Recommender system application developments". Decision Support Systems, vol. 74, pp.12-32, 2015.
- [5] P. Kazienko, K. Musiał, and T. Kajdanowicz, "Multidimensional social network in the social recommender system," IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 41, no. 4, pp. 746-759, 2011.
- [6] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle," IEEE Trans. on Knowledge and Data Engineering, vol. 26, no. 7, pp. 1763-1777, 2014.
- [7] X. Luo, Z. You, M. Zhou, S. Li, H. Leung, Y. Xia, and Q. Zhu, "A highly efficient approach to protein interactome mapping based on collaborative filtering framework," Scientific Reports, vol. 5, pp. 7702, 2015.
- [8] Przemysław Kazienko, Katarzyna Musiał, "Recommendation Framework for Online Social Networks", Advances in Web Intelligence and Data Mining, vol.23, pp. 111-120, 2006.
- [9] Hasan, M. M., Shaon, N. H., Marouf, A. A., Hasan, M. K, "Friend recommendation framework for social networking sites using user's online behavior," international Conference on Computer and Information Technology, pp.539-543, 2015.
- [10] Guy, I., Ronen, I., and Wilcox, E, "Do you know? Recommending people to invite into your social network," in Proceedings of the 14th international conference on intelligent user interfaces, pp. 77-86, 2009.
- [11] Gou, L., You, F., Guo, J., Wu, L., and Zhang, X. L, "SfViz: Interest-based friends exploration and recommendation in social networks," in Proceedings of the 2011 visual information communication – international symposium, pp. 15:1-15:10, 2011.
- [12] Chen, L., Zeng, W., Yuan, Q, "A unified framework for recommending items, groups and friends in social media environment via mutual resource fusion," Expert Systems with Applications, vol.40, no. 8, pp. 2889-2903, 2013.

- [13] Jia Chen, Xin Luo, Ye Yuan, Mingsheng Shang, Zhong Ming, and Zhang Xiong, "Performance of latent factor models with extended linear biases," *Knowledge-Based Systems*, vol. 123, pp. 128-136, 2017.