Global Rating Prediction Mechanism for Trust-Aware Recommender System using K-Shell Decomposition

Linshan Shen[1], Weiwei Yuan[1*] and Donghai Guan[2]

[1]*College of Computer Science & Technology,*
Harbin Engineering Univ., Harbin, China
[2]*College of Automation, Harbin Engineering Univ., Harbin, China*
shenlinshan@hrbeu.edu.cn, yuanweiwei@hrbeu.edu.cn, dhguan@gmail.com

Abstract

The trust-aware recommender system (TARS) suggests the worthwhile information to the users on the basis of trust. Existing models of TARS use personalized rating prediction mechanisms, which can provide personalized services to each user, but they are computational very expensive. We therefore propose an efficient global rating prediction mechanism for TARS: we use the k-shell decomposition to find the most influential nodes in the trust network, and use the recommendations given by these nodes to predict global ratings on items. The experimental results verify that our proposed method can predict ratings accurately with low computational complexity.

***Keywords**: Global rating prediction, K-shell decomposition, Trust-aware recommender system, Trust network, Recommender system*

1. Introduction

The trust-aware recommender system (TARS) suggests the worthwhile information to the users on the basis of trust, in which trust is the measure of willingness to believe in a user based on its competence and behavior within a specific context at a given time [1]. Trust is transitive, so if there is no direct trust between users, it is able to build up some indirect trust relationships via the trust propagations. This contributes to the effectiveness of TARS in solving the well-known data sparseness problem and the cold start problem in the conventional recommender systems [1-5].

Existing models [1-11] of TARS use personalized rating prediction mechanisms. They predict ratings from each active user's point of view: propagating trust from each active user to build up its own trust network, estimating the active user's trust on each recommender, and using the active user's trust on the recommender to weight the recommendation given by this recommender. The advantage of the personalized rating prediction mechanism is obvious: it can provide personalized services to the users of TARS, and users can achieve better prediction accuracy. However, like each coin has two sides, the personalized rating prediction mechanism also has its limitation: it is computational expensive. The computational complexity on building the trust network for an active user is $O(k^d)$, where k is the average degree of the trust network and d is the trust propagation distance.

---

*Weiwei Yuan is the corresponding author of this paper.

To solve the above problem, we propose a novel global rating prediction mechanism for TARS using the k-shell decomposition. The k-shell decomposition, described in [12], has been shown to be effective at identifying influential spreaders in complex networks. Instead of building up an effort-consuming personalized trust network for each active user, we use the k-shell decomposition to find the most influential nodes in the trust network, and use the recommendations given by these nodes to predict global ratings on items. The experimental results verify that our proposed global rating prediction mechanism can predict accurate ratings for users of TARS very efficiently, *i.e.*, computational much less expensive than the personalized rating prediction mechanism.

The rest of the paper is organized as follows: Section 2 introduces the related works; Section 3 presents our proposed global rating prediction mechanism for TARS using the k-shell decomposition in details and gives the experimental results; Sections 4 concludes this paper and points out the future work.

2. Related Works

The architecture of TARS is shown in Figure 1. The inputs are the trust matrix and the rating matrix. The output is the predicted ratings on the items for different users. The trust matrix is the collection of the trust relations between the users of the recommender system. Each element of the trust matrix describes the trust between two users. The rating matrix records the users' ratings on the items. Each element of the rating matrix is the rating given by a user on a particular item. Table 1 gives an example of the inputs of TARS: a rating matrix records 4 users' ratings on 5 items and a trust matrix record 4 users' trust between each other, their relationships are shown in Figure 2.

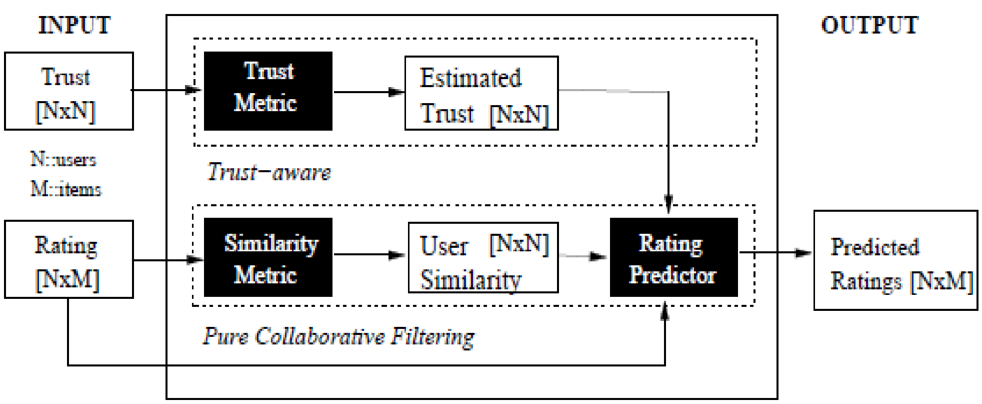


Figure 1. Trust-aware Recommender System Architecture [1]

Table 1. An Example of TARS Inputs: A Rating Matrix Records 4 Users' Ratings on 5 Items and a Trust Matrix Record 4 Users' Trust between Each Other

Rating	I_1	I_2	I_3	I_4	I_5	*Trust*	U_A	U_B	U_C	U_D
U_A	0	3	0	5	4	U_A	0	1	1	1
U_B	1	0	0	3	2	U_B	0	0	1	0
U_C	0	0	5	0	3	U_C	0	1	0	1
U_D	3	0	4	0	0	U_D	1	1	0	0

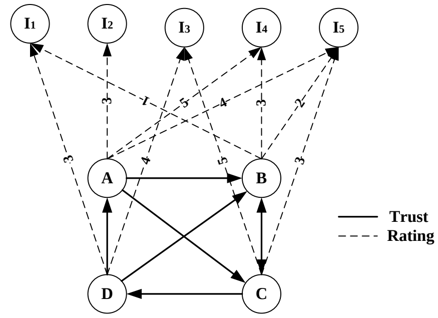


Figure 2. An Example of TARS Inputs with Matrices Given in Table 1

The personalized rating prediction mechanism of TARS is similar as the collaborative filtering (CF). The difference is that CF weights each recommendation based on the active user's similarity to the recommender, while TARS weights each recommendation based on the active user's trust on the recommender:

$$p_{a,i} = \overline{r}_a + \frac{\sum_{u=1}^{m} t_{a,u}(r_{u,i} - \overline{r}_u)}{\sum_{u=1}^{k} t_{a,u}}$$

where a is the active user, i is the item, u is the recommender, m is the number of recommenders, $r_{u,i}$ is u 's recommendation on i , $t_{a,u}$ is a 's trust on u , $\overline{r}$ is the user's average rating, and $p_{a,i}$ is the predicted rating on i for a . $t_{a,u}$ is calculated as:

$$t_{a,u} = \frac{d_{\max} - d_{a,u} + 1}{d_{\max}}$$

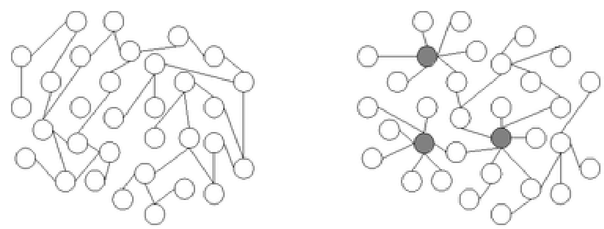
where $d_{\max}$ is the maximum allowable trust propagation distance between users of recommender systems, and $d_{a,u}$ is the trust propagation distance from a to u . In TARS, the trust propagation distance refers to the number of hops in the shortest trust propagation path from the trust or to the trustee.

3. Our Proposed Global Rating Prediction Mechanism using k-shell Decomposition

3.1 Degree: A Heuristic Measure on Building the Global Rating Prediction Mechanism

In the trust network, the indegree of a node represents the number of nodes trust this node, and the outdegree of a nodes represent the number of nodes this nodes trusts. The higher indegree a node has, the more it is trusted by other users in the trust network, and the higher reputation this node has. Therefore, a heuristic measure on building the global rating prediction mechanism of TARS is to use the indegree of the nodes in the trust network (for the undirected graph, we consider the degree of the nodes).

Furthermore, we have verified that trust network is the scale-free network [1-3]. The *scale-free network* is the network whose degree distribution follows a power law, *i.e.*, $P(k) \sim k^{-\gamma}$, where $P(k)$ is the probability that a randomly selected node has k connections, and γ is the power of the degree distribution. The most notable characteristic in the scale-free network is the existence of nodes with degrees greatly exceeds the average. These highest-degree nodes are often called "hubs". Though the number of hubs is limited, they dominate the connectivity of the scale-free network. The comparison between the structure of the random network and the structure of the scale-free network is given in Figure 3. The scale-freeness of the trust network indicates that: in the trust network, there exist some users (hub users) whose reputations are much higher than others. Heuristically, it would be effective to build the global rating prediction mechanism for TARS by finding the hubs users and using the recommendations given by these hub users.



(a) Random network (b) Scale-free network

Figure 3. Comparison between the structure of (a) random network and (b) scale-free network. In the scale-free network, the hubs are highlighted with black nodes

3.2 K-shell: A Measure to Detect the Most Influential Users in Trust Networks

The k-shell (also called k-core) decomposition is a well-established method for detecting the core of a network [13-14]. It assigns an integer index or corness, k_S, to each node representing its location according to successive layers (k-shell) in the network. The higher k_S a node has, the closer this node is to the core of the network. Furthermore, the k_S index is a quite robust measure and the nodes ranking is not influenced significantly in the case of incomplete information. In [12], the k-shell decomposition is used to identify the most influential spreaders in a complex network: the users in the higher k-shell levels are more influential in the network than users in the lower k-shell levels.

Nodes are assigned to k-shell level according to their remaining degree, which is obtained by successive pruning of nodes with degree smaller than the k_S value of the current layer. Once a node has been identified, it is marked (and removed from the network) and the search continues until all nodes in shell k have been found. The process then moves to the next larger k-shell value (and continues until all nodes have been marked).

An example of k-shell decomposition, from [12], is given in Figure 4. Firstly, nodes with degree $k = 1$ are removed. Nodes are keeping on pruned until all nodes with degree $k = 1$ are removed. The removed nodes along with the links connecting them form the $k_S = 1$ k-shell. Next, the pruning process is repeated in a similar way for the nodes of degree $k = 2$ to extract the $k_S = 2$ k-shell and subsequently for higher value of k until all nodes are removed. As a result, the network can be viewed as a set of adjacent k-shells.

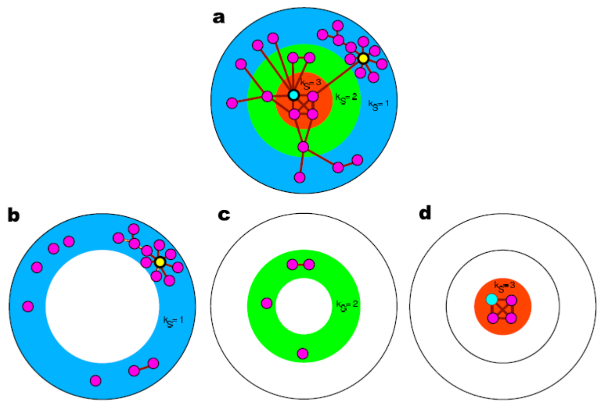


Figure 4. The illustration of the k-shell extraction method [12]. a, A schematic network is represented as a set of 3 successively enclosed k-shells labeled accordingly. b, Nodes with edges forming $k_S=1$ shell of the network. c, Nodes with edges forming $k_S=2$ shell of the network. d, Nodes with edges forming $k_S=3$ shell of the network.

3.3 Comparing the Effectiveness of Degree and k-Shell on TARS

We first compare the effectiveness of degree and k-shell on the rating prediction performances for TARS.

3.3.1 Experimental Data: The experiments are held on the public released TARS dataset Epinions, which is available at trustlet.org†. Epinions consists of 40163 users' 664824 ratings on 139738 items. In this work, we consider the undirected trust network, *i.e.*, if user A trusts users B, user B also trusts user A. The user trust is extracted using the method given in [3], where $Thres_s$=0.8 and $Thres_I$=10. We get 1426 undirected trust statements between 903 users; these users give 125740 ratings on 46194 items. The degree distribution and the k-shell distribution (using the method mention in Section 3.2) of our selected trust network are given in Figure 5 and Figure 6 respectively. The maximum degree of the selected trust network is 54 and the maximum k-shell level of the trust network is 6. The degree distribution of the trust network clearly shows that the trust network is the scale-free network, which means there exists some nodes with superior reputations in the trust network.

To compare the effectiveness of degree and k-shell on TARS, we randomly select 94 users from the total 903 users. These 94 users rated 5752 items and got 250694 recommendations from 872 recommenders.

† http://www.trustlet.org/wiki/Datasets

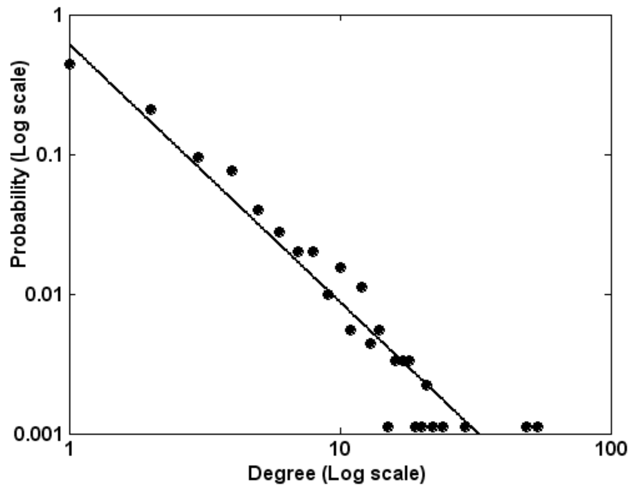


Figure 5. The Degree Distribution of Select Trust Network, where γ =1.8

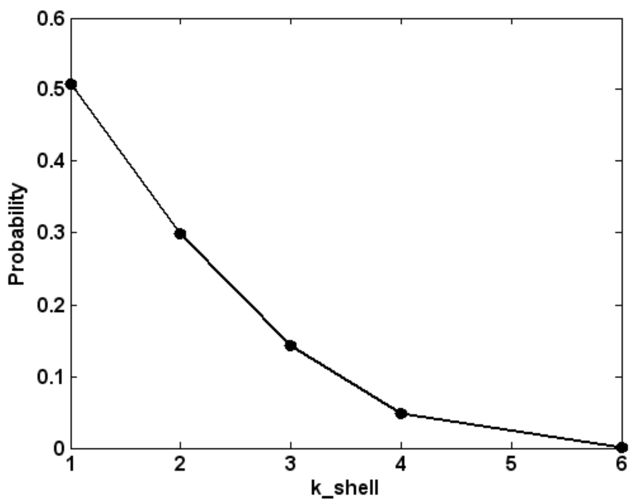


Figure 6. The k-shell Distribution of Select Trust Network

3.3.2 Comparison Results: We compare the effectiveness via calculating the error of the recommendations with the active users' real ratings on the items, given specific degree and k-shell level. Particularly, we use the mean absolute error (MAE), since it is most appropriate and useful for evaluating prediction accuracy in offline tests [1]. Figure 7 and Figure 8 give the overall gap between the recommendations and the real ratings in according to degree and k-shell respectively. Figure 9 and Figure 10 give the further analysis on the gaps between the recommendations and the real ratings when the real ratings varying from 1 to 5 respectively. The simulation results illustrate that: (1) the recommendations given by users with higher k-shell levels are more reliable; (2) the errors of recommendations decrease when the active users' real ratings increase; (3) the recommendations given by users with higher degrees are not always more reliable, *e.g.*, the recommendations given by users with degrees 18 are least reliable in our simulation, even worse than the recommendations given by users whose degrees equal to 1; the recommendations given by users with degrees 29 are most reliable in our simulation data, even better than the recommendations given by users whose degrees equal to

54; (4) the overall ranges of the recommendation errors by using the degree and k-shell are similar, while the recommendation errors by using k-shell tends to be smaller: the errors are in the range of (0.8, 1.2) by using degree, while the errors are in the range of (0.8, 1.1) by using k-shell.

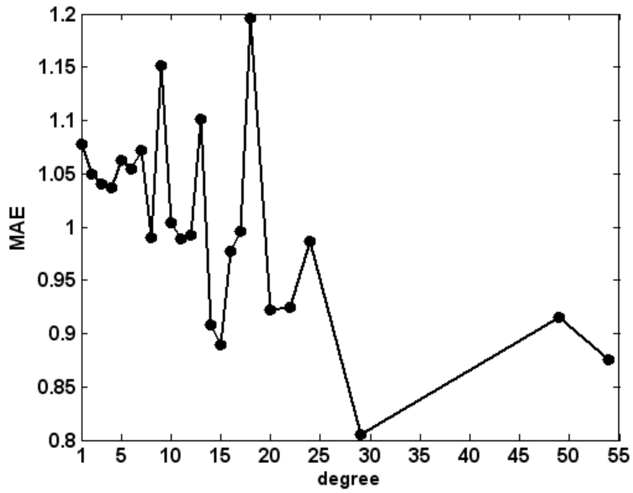


Figure 7. Comparison on the Effectiveness of Degree for All Ratings

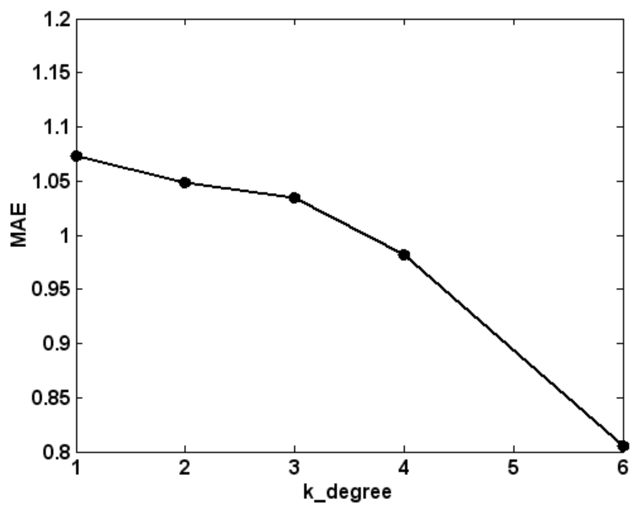


Figure 8. Comparison on the Effectiveness of k-shell on TARS for All Ratings

The above comparisons show that k-shell is a better measure than degree on building the global rating prediction mechanism: the recommendations given by users with higher k-shell levels always tend to be more reliable, and the overall ranges of the recommendation errors by using k-shell are no worse than that by using degree.

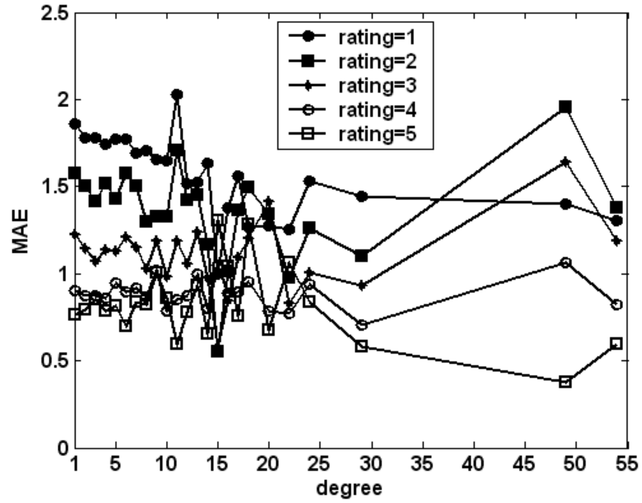


Figure 9. Comparison on the Effectiveness of Degree on TARS when the Active Users' Real Ratings Varying from 1 to 5

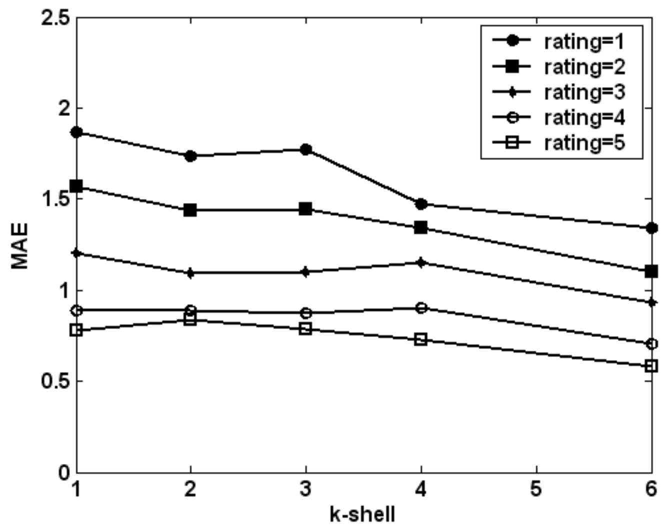


Figure 10. Comparison on the Effectiveness of k-shell on TARS when the Active Users' Real Ratings Varying from 1 to 5.

3.3.3 Global rating prediction mechanism using k-shell decomposition: Based on the analysis in Section 3.3, we propose a global rating prediction mechanism using k-shell decomposition for TARS. The key idea of our rating prediction mechanism is: firstly, based on the degrees extracted from the trust network, the k-shell level of each user is calculated; secondly, we set a threshold value for the k-shell level, and the recommendations given by recommenders whose k-shell levels are greater than this threshold value are valid for the rating prediction; finally, we calculate the mean of the valid recommendations as the predicted rating on the target item. The details of our proposed global rating prediction mechanism are given in Table 2.

The performance of our proposed global rating prediction mechanism is verified on the following experimental data: based on the data given in Section 3.3.1, we randomly select 92 users from the total 903 users, these 92 users rated on 5086 items, and they got 225843 recommendations from 872 recommenders. We verify the rating prediction coverage and rating prediction accuracy respectively for TARS. The rating prediction coverage of TARS is the portion of items that TARS is able to predict, *i.e.*, the portion of items that the active user can get at least one recommendation. The rating prediction accuracy is measured by the mean absolute error (MAE): the predicted rating is compared with the real rating and the difference (in absolute value) is the prediction error, this error is then averaged over all predictions to obtain the overall MAE.

Table 2. Our Proposed Global Rating Prediction Mechanism of TARS

Input: $R_{n\times m}$ (n users' ratings on m items), $T_{n\times n}$ (n users' trusts on each other)

Output: $P_{n\times m}$ (predicted ratings for n users on m items)

Parameters: $Thers_{k_S}$ (the threshold value of the k-shell level), de_k (degree of user k), de_{k_S} (k-shell level of user k), $p_{a,i}$ (the value in the a^{th} row and i^{th} column of $P_{n\times m}$), $r_{u,i}$ (recommender u 's recommendation on item i)

1: using $T_{n\times n}$, calculate the degrees of n users respectively
2: set $de_{\max}$ = the maximum degree of the n users
3: for i =1 to $de_{\max}$
4:   for k =1 to n
5:     if $de_k = i$
6:       $de_{k_S} = i$
7:       remove user k
8:     end if
9:   end for
10:   for k =1 to n
11:     if $de_k = i$
12:       go to step 3
13:     end if
14:   end for
15: end for
16: for a =1 to n
17:   for i =1 to m
18:     $p_{a,i} = \frac{1}{num}\sum_{u-1}^{num} r_{u,i}$ where $de_{u_S} > Thers_{k_S}$
19:   end for
20: end for

We use the personalized rating prediction mechanism in [1] to compare the performance of our proposed global rating prediction mechanism. The coverage and MAE by using the personalized rating prediction mechanism are 80.17% and 0.8241 respectively. The coverage and MAE by using our proposed global rating prediction mechanism with different threshold values of k-shell level are given in Table 3.

The computational complexity of the personalized rating prediction mechanism in [1] is $O(k^d)$, where k is the average degree of the trust network and d is the trust propagation distance, and the computational complexity of our proposed global rating prediction mechanism is $O(c)$, where c is a constant. The experimental results show that: in our experiments if we set $Thers_{k_S} = 3$, our proposed global rating prediction mechanism has similar rating prediction coverage and slightly worse rating prediction accuracy as the personalized rating prediction mechanism, while our proposed method is computational much less expensive.

Table 3. The Rating Prediction Coverage and Rating Prediction Accuracy of TARS by using Our Proposed Global Rating Prediction Mechanism

Our proposed global rating prediction mechanism	Coverage	MAE
$Thers_{k_S}=1$	100.00%	0.9392
$Thers_{k_S}=2$	92.08%	0.8934
$Thers_{k_S}=3$	78.61%	0.8654
$Thers_{k_S}=4$	57.88%	0.8473
$Thers_{k_S}=6$	7.22%	0.8201

4. Conclusions and Future Works

Existing works of TARS use the personalized rating prediction mechanism to provide personalized services to the users. However, the personalized rating prediction mechanism is always effort consuming since it need to build up the trust network for each active user. We therefore propose a novel global rating prediction mechanism for TARS, which is able to efficiently predict ratings. Our proposed rating prediction mechanism is based on the k-shell decomposition: we find via experiments that the recommendations given by users with higher k-shell level always tend to be more reliable, so our proposed rating prediction mechanism uses the recommendations given by the recommenders with highest k-shell levels to predict ratings. The simulation results show that our proposed global rating prediction mechanism can achieve reasonable rating prediction performances with much less expensive computational complexity than the personalized rating prediction mechanism.

In the future, we plan to focus on more details of the global rating prediction mechanism: the possible attributes which would affect the prediction performances and the relationships between these attributes. Though the research on the global rating prediction mechanism of TARS is still at the beginning stage, we do believe that it presents a promising path for the future research.

Acknowledgements

This research was supported by National Natural Science Foundation of China (Grant No. 61100007, 61100081, 61100006, 61272184), and Fundamental Research Funds for the Central Universities (HEUCF100605, HEUCF041204).

References

[1] W. Yuan, D. Guan, Y. K. Lee, S. Lee and S. J. Hur, "Improved trust-aware recommender system using small-worldness of trust networks", Knowledge-Based Systems, vol. 23, **(2010)**, pp. 232-238.
[2] W. Yuan, D. Guan, Y. K. Lee and S. Lee, "The Small-World Trust Network", Applied Intelligence, **(2010)** April 27, pp. 1-12.
[3] W. Yuan, L. Shu, D. Guan, Y. K. Lee and S. Lee, "iTARS: Trust-Aware Recommender System using Implicit Trust Networks", IET Communications, vol. 14, **(2010)**, pp. 1709 – 1721.
[4] P. Massa and P. Avesani, "Trust-aware Collaborative Filtering for Recommender Systems", Proceeding of Federated Int'l. Conf. on the Move to Meaningful Internet, **(2004)**, pp. 492-508.
[5] P. Massa and P. Avesani, "Trust Metrics in Recommender Systems", Proceeding of Computing With Social Trust, **(2009)**, pp. 259-285.
[6] Y. Li and C. Kao, "TREPPS: A Trust-based Recommender System for Peer Production Services", Expert Systems with Applications, vol. 36, **(2009)**, pp. 3263-3277.
[7] F. Walter, S. Battistion and F. Schweitzer, "A model of a trust-based recommendation on a social network", Autonomous Agents and Multi-Agent System, vol. 16, **(2008)**, pp. 57-74.
[8] P. Massa and P. Avesani, "Trust-aware recommender systems", Proceeding of the 2007 ACM Conference on Recommender Systems, **(2007)**, pp. 121-126.
[9] M. Kim, Y.S. Jeong, J. H. Park and S. O. Park, "Group-Aware Social Trust Management for a Movie Recommender System", Proceeding of 5th International Conference on Information Technology, vol. 107, Part 5, **(2011)**, pp. 495-501.
[10] J. Sun and X. Yu, "Research on Trust Recommender Model based on experts' social networks", Proceeding of International Conference on Fuzzy Systems and Knowledge Discovery, **(2010)**, pp. 2503-2506.
[11] N. Sahli and G. Lenzini, "Trust-Aware Recommender Systems for Open and Mobile Virtual Communities", Collaborative Computer Security and Trust Management, DOI:10.4018/978-1-60566-414-9.ch007, **(2010)**.
[12] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley and H. A. Makse, "Identifying Influential Spreaders in Complex Networks", Nature Physics, vol. 6, **(2010)**, pp. 888-893.
[13] S. B. Seidman, "Network Structure and Minimum Degree", Social Networks, vol. 5, no. 3, **(1983)**, pp. 269–287.
[14] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt and E. Shir, "A Model of Internet Topology using K-Shell Decomposition", Proceeding of the National Academy of Sciences of the United States of America (PNAS), vol. 104, no. 27, **(2007)**, pp. 11150-11154.

Authors

Linshan Shen is a Ph.D. candidate of College of Computer Science and Technology, Harbin Engineering University (HEU), Harbin, China. He received hisB.S. degree in computer science and technology in 2003, and M.S. s in computer engineering from Kumoh National Institute of Technology, Korea, in 2006. He is currently a lecturer in the College of Computer Science and Technology, Harbin Engineering University. His research interest focuses on distributed computing, autonomic computing, task scheduling etc

E-mail: shenlinshan@hrbeu.edu.cn

Weiwei Yuan received Bachelor's degree and Master's degree in Automation and Computer Engineering in 2002 and 2005 respectively from Harbin Engineering University, China. In 2010, she got her Ph.D. degree in the Department of Computer Engineering, Kyung Hee University, South Korea. Since September 2010, she has been an assistant professor in Harbin Engineering University, China. Her research interests are social computing, information security and machine learning.

E-mail: yuanweiwei@hrbeu.edu.cn

Donghai Guan received his B.S. from Harbin Engineering University, Harbin, China. He got his M.S. degree in Computer Science from Kumoh National Institute of Technology (KIT), Gumi, South Korea in 2004. He got his Ph.D. degree in Computer Science from Kyung Hee University, South Korea in 2009. From 2009, he was a Post Doctoral Fellow at Computer Science Department, Kyung Hee University. Since February 2011, he has been an assistant professor in Harbin Engineering University, China. Since March 2012, he has been an assistant professor in Kyung Hee University, South Korea. His research interests are Machine Learning, Data Mining, Activity Recognition, and Social Computing.

E-mail: dhguan@gmail.com