Recommending with Limited Number of Trusted Users in Social Networks

Weiwei Yuan, Donghai Guan
Dept. of Computer Science and Technology
Nanjing University of Aeronautics and Astronautics
Nanjing, City
{yuanweiwei, dhguan}@nuaa.edu.cn

Asad Masood Khattak
College of Technological Innovation
Zayed University
Abu Dhabi, UAE
Asad.Khattak@zu.ac.ae

***Abstract*— To estimate the reliability of an unknown node in social networks, existing works involve as many opinions from other nodes as possible. Though this makes it possible to approximate the real property of the unknown nodes, the computational complexity increases as the scale of social networks getting bigger and bigger. We therefore propose a novel method which involve only limited number of social relations to predict the trustworthiness of the unknown nodes. The proposed method involves four rating prediction mechanisms: FM use the recommendation given by the most reliable recommender with the shortest trust propagation distance from the active user as the predicted rating, FMW weights the recommendation in FM, FA uses the mean value of recommendations with the shortest trust propagation distance from the active user as the predicted rating, and FAW weights recommendations in FA. The simulation results show that the proposed method can greatly reduce the rating prediction calculation, while the rating prediction losses are reasonable.**

Keywords—node reliability, trusted user, social networks, social computing

I. INTRODUCTION

To estimate the reliability of an unknown node in the social network, a common sense is to involve the opinions of other nodes. Inferring other nodes' social trust on the unknown node, it is possible to predict the trustworthiness of this unknown nodes.

Existing works [1, 2, 3] involve as many opinions as possible to estimate the reliability of the unknown nodes. With plenty of the opinions, it is possible to use different kinds of learning methods to approximate the real property of the unknown nodes. However, as the scale of the social network is getting bigger and bigger in the real applications, the computational complexity of finding all possible opinions is getting higher [4, 5]. This is because, the whole social network needs to be searched to find the possible opinions. In addition, since the social relations can be propagated, the propagation always need longer time when many nodes need to be found in the social network.

To solve the problems of existing work, this work proposes an alternative method, instead of using as many social relations as possible, the proposed method only involves limited number of social relations to predict the trustworthiness of the unknown nodes. In the proposed model, opinions from very limited number of users are used, while opinions from other users are ignored. The proposed method is inspired by the famous economics law: the Watch Law. It is mentioned that "a person watches have a chart, you can know what time it is, and when he has two not being skillful to determine. Two tables and cannot differentiate a person extra precise time, so they will watch the people lose trust in precise time." This work proposes 4 rating prediction mechanisms using recommendations given by few recommender: FM use the recommendation given by the most reliable recommender with the shortest trust propagation distance from the active user as the predicted rating, FMW weights the recommendation given by the most reliable recommender with the shortest trust propagation distance from the active user as predicted rating, FA uses the mean value of recommendations with the shortest trust propagation distance from the active user as the predicted rating, and FAW weights recommendations with the shortest trust propagation distance from the active user as the predicted rating. The simulation results show that the proposed 4 rating prediction mechanisms can greatly reduce the rating prediction calculation, while the rating prediction losses of the proposed 4 rating prediction mechanisms are reasonable. Among the proposed 4 rating prediction mechanisms, FAW tends to have the best rating prediction performances. It can greatly reduce the rating prediction calculation with only 5%-10% of rating prediction accuracy loss.

The rest of this paper is organized as follows: Chapter 2 introduces the related works, Chapter 3 presents the proposed methods, Chapter 4 gives the simulation results held on the real application data, and Chapter 5 concludes this paper.

II. RELATED WORKS

Trust-aware recommender system (TARS) [9, 10, 11, 12, 13] predicts ratings based on the trust relationships between users. It is the improved collaborative filter (CF) algorithm [6, 7, 8]: CF predicts ratings based on user similarity. Since the rating matrix is very sparse in the real applications, it is always difficult for CF to calculate the user similarity between users. Trust is the measure of willingness to believe in a user based on its competence and behavior within a specific time.

978-1-5386-4366-2/18/$31.00 ©2018 IEEE

One basic property of trust is that it is transitive. Even if there is no direct trust between users, some indirect trust can be built up based on trust propagations. So TARS can achieve higher prediction coverage than CF in the real applications.

To achieve high rating prediction accuracy, existing TARS models involve as many recommenders as possible. Recommendations given by these recommenders are weighted based on the trust propagation distance from the active user to the recommenders. The shorter the trust propagation distance from the active user to the recommender, the more important the recommendation given by the recommender. The rating prediction mechanism of classical TARS model is similar to that of CF, while the difference is that CF use user similarity to weight the recommendations and TARS use user trust to weight the recommendations. The details of rating prediction mechanism of classical TARS is given Table 1.

TABLE I. RATING PREDICTION MECHANISM OF EXISTING TARS MODELS

Input: **T** : the trust relations between users, **R** : ratings given by users on different items

Output: **P** : predictions on unrated items

Parameters: a : active user, i :item, r :user, R_{ri} : r 's rating on i , d_{ar} : shortest trust propagation distance from a to r , $d_{\max}$: maximum allowable trust propagation distance, $\overline{R}$:average rating of the user

For each active user a
  For each item i
    For each user r
      If $R_{ri} > 0$
      If $0 < d_{ar} \le d_{\max}$

$$w_{ar} = \frac{d_{\max} + 1 - d_{ar}}{d_{\max}}$$

$$P_{ai} = \overline{R}_a + \frac{\sum w_{ar}(R_{ri} - \overline{R}_r)}{\sum w_{ar}}$$

    End for
  End for
End for

As shown in Table 1, the trust propagation distance from the active user to the recommenders should be no longer than the maximum allowable trust propagation distance. If the trust propagation distance is too long, the recommendations given by the recommender may have very limited effects on the rating prediction, while this greatly increases the recommender searching cost. It has been verified in the related works that when setting a suitable value for the maximum allowable trust propagation distance, TARS can achieve high rating prediction accuracy with almost similar rating prediction coverage as searching the whole trust network.

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. THE PROPOSED METHODS

This work proposes four algorithms involving only limited number of recommenders in rating prediction. The four algorithms are named as FM, FMW, FA and FAW respectively. The existing method is called FE in this paper. FE predicts ratings with all possible recommenders which could be found by the trust propagation from the active user. The proposed four algorithms are defined as follows:

FM predicts ratings with recommendation given by the most reliable user with the shortest trust propagation distance from the active user. The user’s indegree is used to measure user reliability. A user’s indegree means the number of users trust this user. The higher a user’s indegree is, the more users trust this user in the trust network, so this user tends to be more reliable. The trust propagation distance from the active user to the recommender measures the strength of their relationship. The shorter the trust propagation distance is, the stronger the relationship between the active user to the recommender, so it is more likely the active user and the recommender will give similar rating on items. So the most reliable user with the shortest trust propagation distance means this user is the most reliable one among those who have the closest relationship with the active user. FM uses the ratings given by this selected recommender as the prediction for the active user on the target item. The details of FM are given in Table 2.

TABLE II. RATING PREDICTION MECHANISM OF FM

For each active user a
  For each item i
    $d_t = d_{\max}$
    $in_t = 0$
    For each user r
      If $R_{ri} > 0$
      If $0 < d_{ar} \le d_t$
        $d_t = d_{ar}$
        If $in_r > in_t$
          $in_t = in_r$
          $P_{ai} = R_{ri}$
    End for
  End for
End for

FMW predicts ratings with similar rating prediction mechanism of FM. They both predict ratings with recommendation given by the most reliable recommender with the shortest trust propagation distance from the active user. The difference between FMW and FM is that FM uses the recommendations given by the selected recommenders as the predicted ratings directly, while FMW weighted the recommendations based on the propagation distance from active user to the recommender. The weights are calculated by the mechanism of FE:

$$w_{ar} = \frac{d_{\max} + 1 - d_{ar}}{d_{\max}} \tag{1}$$

When predicting ratings, we refer to the rating prediction of FE:

$$P_{ai} = \overline{R_a} + \frac{\sum w_{ar}(R_{ri} - \overline{R_r})}{\sum w_{ar}} \tag{2}$$

To predict the active user's rating on one item, only one recommender is chosen, so (2) can be transformed to:

$$P_{ai} = \overline{R_a} + R_{ri} - \overline{R_r} \tag{3}$$

The details of FMW are given in Table 3.

TABLE III. RATING PREDICTION MECHANISM OF FMW

```
For each active user a
  For each item i
    d_t = d_max
    in_t = 0
    For each user r
      If R_ri > 0
      If 0 < d_ar ≤ d_t
        d_t = d_ar
        If in_r > in_t
          in_t = in_r
          P_ai = R̄_a + R_ri − R̄_r
    End for
  End for
End for
```

FA predicts ratings with recommendations given by all recommenders with the shortest trust propagation distance from the active user. FM and FMW consider both the reliability of the recommender and the similarity between the recommender and the active user, while FA only considers the similarity between the recommender and the active user. Since the similarity of users are measures by their trust propagation distance in TARS, so FA chooses recommenders with the shortest trust propagation distance to find the most similar recommenders for the active user. The prediction is the average rating given by all selected recommenders. The details of FA are given in Table 4.

TABLE IV. RATING PREDICTION MECHANISM OF FA

```
For each active user a
  For each item i
    d_t = d_max
    n = 0
    P_ai = 0
    For each user r
      If R_ri > 0
      If 0 < d_ar ≤ d_t
        d_t = d_ar
        P_ai = P_ai + R_ri
        n = n + 1
    End for
    P_ai = P_ai / n
  End for
End for
```

FAW predict ratings with similar prediction mechanism of FA. They both predict ratings with recommendations given by all recommenders with the shortest trust propagation distance from the active user. The difference between FAW and FA is that: FA uses the average value of all recommendations as the rating prediction value, while FAW weights the recommendations before merging. Recommendations are weighted by the mechanism of FE：

$$w_{ar} = \frac{d_{\max} + 1 - d_{ar}}{d_{\max}} \tag{4}$$

When predicting ratings, we refer to the rating prediction of FE:

$$P_{ai} = \overline{R_a} + \frac{\sum w_{ar}(R_{ri} - \overline{R_r})}{\sum w_{ar}} \tag{5}$$

To predict the active user's rating on one item, all recommenders with the same trust propagation distance from the active user are chosen, so all recommenders have similar weight based on the calculation of (4). Formula (5) can be transformed to:

$$P_{ai} = \overline{R_a} + \frac{\sum (R_{ri} - \overline{R_r})}{n} \tag{6}$$

where n is the number of recommenders which have the shortest trust propagation distance from the active user in the trust network.

The details of FAW are given in Table 5.

TABLE V. RATING PREDICTION MECHANISM OF FAW

```
For each active user a
  For each item i
    d_t = d_max
    n = 0
    P_ai = 0
    For each user r
      If R_ri > 0
      If 0 < d_ar ≤ d_t
        d_t = d_ar
        P_ai = P_ai + R_ri − R̄_r
        n = n + 1
    End for
    P_ai = R̄_a + P_ai / n
  End for
End for
```

IV. SIMULATION RESULTS

The experiments are held on Epinions dataset [14]. We randomly choose 5000 users from Epinions dataset. There are totally 205935 trust relations between these users. These 5000 users rated 291255 ratings. The simulations are repeated 50 times. For each simulation, we randomly choose 100 users from the 5000 users as the active users, and use their rated items as the target of rating prediction. The performances of the proposed 4 algorithms FM, FMW, FA and FAW are compared with the performances of the existing rating prediction mechanism of TARS, i.e., FE, which is given in Section 2.

The comparison of prediction calculation reduction is given in Fig. 1. The prediction calculation of FE is regarded as 1. FM and FMW both select the most reliable recommender with the shortest trust propagation distance from the active user, so FM and FMW have the same complexity in prediction calculation. FA and FAW both select all recommenders with the shortest trust propagation distance from the active user, so FA and FAW have the same complexity in prediction calculation. Since FA and FAW use recommendations from more recommenders, the calculation of the prediction reduction is less than that of FM and FMW. As shown in Fig. 1, in the 50 times repeated experiments, FA and FAW can reduce around 80% to 88% of the calculation in rating prediction compared with existing work. FM and FMW can reduce around 96% to 98% of the calculation in rating prediction compared with existing work. All of the four proposed algorithm can greatly reduce the rating prediction calculation in rating prediction compared with existing work.

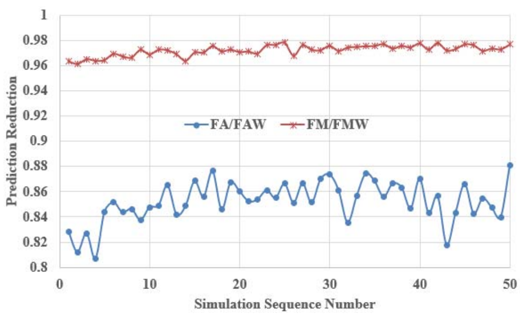

Fig. 1. The prediction reduction of the proposed methods compared with the existing method.

The cost of achieving high prediction calculation reduction is the loss of the rating prediction accuracy. This is because very few number of recommenders are used. However, it has been verified by experiments that compared with the gain for prediction calculation reduction, the loss of the rating prediction accuracy is very limited. Fig. 2 shows the rating prediction accuracy loss of FM, FMW, FA and FAW compared with that of FE. This is done for all the ratings, not differentiate the values of the real ratings, i.e., in this experiments, we do not consider whether the real rating is 1, 2, 3, 4 or 5. Intuitively, the rating prediction accuracy of existing methods is always better when the real rating is 4 or 5. This is because in this case, the prediction algorithm gets more data to be trained. It is shown in Fig. 2 that the accuracy loss of FMW is the highest, around 13% to 23%. The accuracy loss of FM are around 0 to 10%. The accuracy loss of FA is around -5% to 10%. The minus value of accuracy loss means the proposed method has better rating prediction accuracy compared with the existing work. The accuracy loss of FAM is around 5% to 10%. The variance of FAW's accuracy loss is the smallest among the four proposed methods.

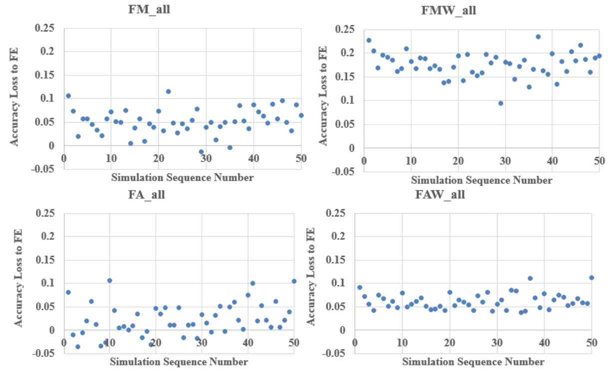

Fig. 2. The comparison of accuracy between existing work and the proposed methods.

Fig. 3 shows the accuracy loss of FM, FMW, FA and FAW compared with the prediction accuracy of the existing method FE, given the real ratings equal to 1. The weighted recommendations showed better performance than the unweighted recommendations: the accuracy loss of FM is higher than the accuracy loss of FMW, and the accuracy loss of FA is higher than the accuracy loss of FAW. The performance of FM is the worst among the proposed 4 methods given the rating real equals to 1. Its accuracy loss is around 10% to 60%. Weighting the recommendation given by the most reliable recommender can significantly reduce the accuracy loss. The accuracy loss of FMW is only around 0 to 25%. By increasing the number of recommenders, the accuracy loss tends to be reduced. The accuracy loss of FA is around 0 to 30% compared with FE, given the real rating equals to 1. The accuracy loss of FA is around half of that of FM. The accuracy of FAW is the best among the 4 proposed method: it is around -10% to 10%. The variance of FAW's accuracy loss is also the smallest among the four proposed methods given the real rating equals to 1.

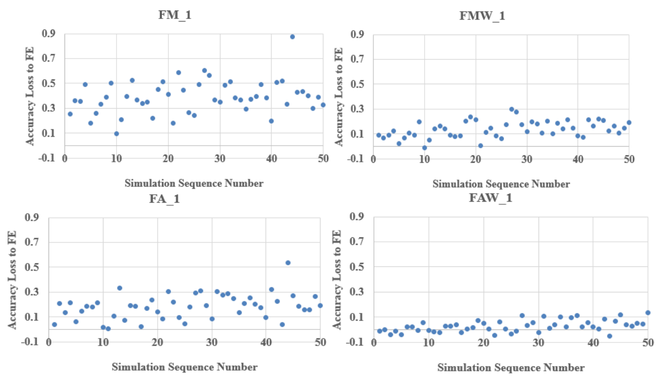

Fig. 3. The comparison of accuracy between existing work and the proposed methods given the real rating on items equals to 1

Fig. 4 shows the accuracy loss of FM, FMW, FA and FAW compared with the prediction accuracy of the existing method FE, given the real ratings equal to 2. The accuracy losses of FM, FMW, FA and FAW given the real ratings equal to 2 are higher than the accuracy losses of their corresponding method given the real ratings equal to 1. Similar as the accuracy losses given the real ratings equal to 1, when the real ratings equal to 2, the weighted recommendation showed better performance than the unweighted recommendations: the accuracy loss of FM is higher than the accuracy loss of FMW, and the accuracy loss of FA is higher than the accuracy loss of FAW. The methods using more recommendations show better performance than the method using fewer recommendations: the accuracy losses of FA and FAW are lower than the accuracy losses of FM and FMW. The performance of FM is still the worst among the proposed 4 methods given the rating real equals to 2. Its accuracy loss is around 50% to 110%. The performance of FAW is still the best among the proposed 4 methods given the rating real equals to 2. Its accuracy loss is around 0% to 30%. And the variance of FAW's accuracy loss is also the smallest among the four proposed methods given the real ratings equal to 2.

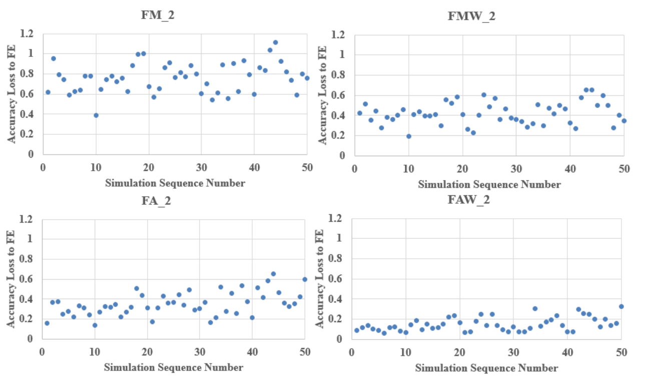

Fig. 4. The comparison of accuracy between existing work and the proposed methods given the real rating on items equals to 2.

Fig. 5 shows the accuracy loss of FM, FMW, FA and FAW compared with the prediction accuracy of the existing method FE, given the real ratings equal to 3. The accuracy losses of FM, FMW, FA and FAW given the real ratings equal to 3 are higher than the accuracy losses of their corresponding method given the real ratings equal to 1 or 2. Similar as the accuracy losses given the real ratings equal to 1 or 2, when the real ratings equal to 3, the weighted recommendation showed better performance than the unweighted recommendations: the accuracy loss of FM is higher than the accuracy loss of FMW, and the accuracy loss of FA is higher than the accuracy loss of FAW. The methods using more recommendations show better performance than the method using fewer recommendations: the accuracy losses of FA and FAW are lower than the accuracy losses of FM and FMW. The performance of FM is still the worst among the proposed 4 methods given the rating real equals to 3. Its accuracy loss is around 60% to 120%. The performance of FAW is still the best among the proposed 4 methods given the rating real equals to 3. Its accuracy loss is around 20% to 40%. And the variance of FAW's accuracy loss is also the smallest among the four proposed methods given the real ratings equal to 3.

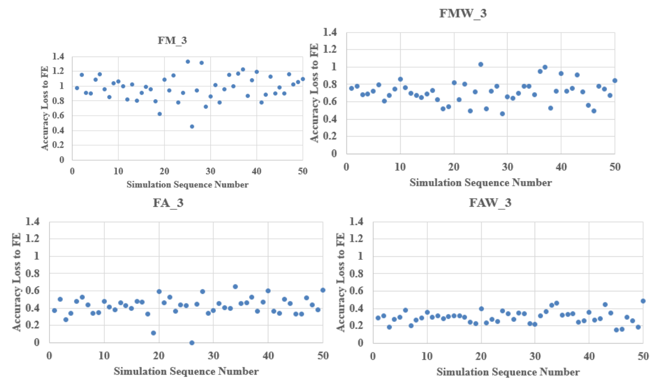

Fig. 5. The comparison of accuracy between existing work and the proposed methods given the real rating on items equals to 3.

Fig. 6 shows the accuracy loss of FM, FMW, FA and FAW compared with the prediction accuracy of the existing method FE, given the real ratings equal to 4. The accuracy losses of FM, FMW, FA and FAW given the real ratings equal to 4 are the highest compared with the accuracy losses of their corresponding method given the real ratings equal to 1, 2, 3 or 5. This is because the rating 4 and rating 5 both means the users like the rated items, and these two ratings are always used alternatively in real applications. So it is hard to gather the very distinct features to differentiate the rating prediction of rating 4 and rating 5. Though the weighted recommendation showed better performance than the unweighted recommendations, their performances are not very distinct. The performance of FAW is still the best among the proposed 4 methods given the rating real equals to 4. Its accuracy loss is around 0% to 30%. And the variance of FAW's accuracy loss is also the smallest among the four proposed methods given the real ratings equal to 4.

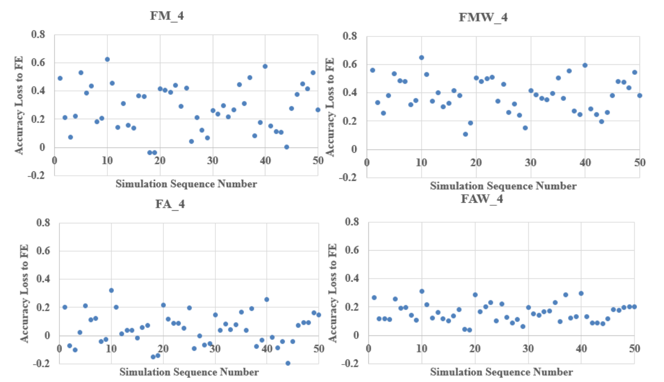

Fig. 6. The comparison of accuracy between existing work and the proposed methods given the real rating on items equals to 4.

Fig. 7 shows the accuracy loss of FM, FMW, FA and FAW compared with the prediction accuracy of the existing method FE, given the real ratings equal to 5. Different as the accuracy loss when given real ratings equal to 1, 2, 3 or 4, the accuracy losses of FM, FMW, FA and FAW given the real ratings equal to 5 tend to be negative numbers. This means prediction accuracy of FM, FMW, FA and FAW tend to be better than the prediction accuracy of the existing work. Different as the cases when the real ratings equal to 1, 2, 3 or

4, the performance of FM is the best among the proposed 4 methods. Its accuracy loss is around -40% to -20%, i.e., its prediction accuracy is 20% to 40% higher than the prediction accuracy of the existing method. The weighted recommendation showed worse performance than the unweighted recommendations, and the methods use fewer recommendations tend to be better than the methods use more recommendations. That is, when the real ratings equal to 5, the rating prediction accuracy of FMW/FAW is better than that of FM/FA respectively, and the rating prediction accuracies of FM and FMW are better than those of FA and FAW. The rating prediction accuracies of FMW and FAW are similar as the rating prediction accuracy of the existing method. And the variance of both FMW and FAW are small.

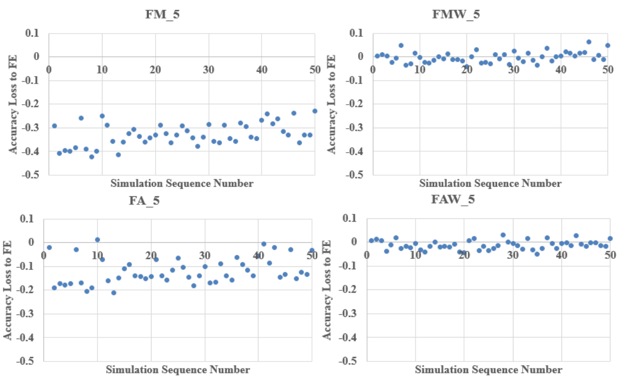

Fig. 7. The comparison of accuracy between existing work and the proposed methods given the real rating on items equals to 5.

To sum up, the rating prediction performance of FAW is the best among the proposed 4 methods. Its total accuracy loss is around 5% to 10% compared with the the exiting work, while it can reduce 80% to 88% of the prediction calculation. Its variance is also the smallest among the proposed 4 methods. Compared with the loss in the prediction loss, the gain in the prediction calculation reduction is significant. This means FAW is suitable to be used if only limited number of recommenders are available in TARS.

V. CONCLUSIONS

TARS predict ratings based on the trust relationship between the active user and the recommenders. It can reach recommenders via the trust propagation among users. In the existing work, TARS involve as many recommenders as possible to predict ratings. However, in the large scale real applications, the trust propagation may be computational complexity. This paper tries to find the method to predict ratings with only limited number of recommenders. This work proposes 4 different rating prediction mechanism using few recommender: FM and FMW both use the most reliable recommender from the active users to predict the rating; FA and FAW both use all recommenders with the shortest trust propagation distance from the active user. The difference between FM/FA and FMW/FAW is that FM/FA uses the recommendation as the predicted rating directly, while FMW/FAW weights the recommendations based on the trust relationship between the active user and the recommender. The experimental results show that FA has the best prediction performance among the proposed 4 methods: it can greatly reduce the calculation of rating prediction with only 5%-10% accuracy loss compared with the existing work. The proposed method has great value in the real applications especially when the scale of the network is very large. It greatly reduces the computational complexity with reasonable prediction accuracy.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No. 61672284). It was also supported by China Postdoctoral Science Foundation (No. 2016M591841) and Zayed University Research Cluster Award # R18038. Prof. Donghai Guan is the corresponding author of this paper.

REFERENCES

[1] S. Liu, C. Jiang, and Z. Lin, "Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach," Information Sciences, vol. 306, pp. 34-52, 2015.

[2] S.R. Yan, X. L. Zheng, and Y. Wang, "A graph-based comprehensive reputation model: Exploiting the social context of opinions to enhance trust in social commerce," Information Sciences, vol. 318, pp. 51-72, 2015.

[3] W. Jiang, G. Wang, and M.Z.A. Bhuiyan, "Understanding graph-based trust evaluation in online social networks: methodologies and challenges," ACM Computing Surveys (CSUR), vol. 49, pp. 1-10, 2016.

[4] G. Guo, J. Zhang J, and N. Yorke-Smith, "A novel recommendation model regularized with user trust and item ratings," IEEE Transactions on Knowledge and Data Engineering, vol. 28(7), pp. 1607-1620, 2016.

[5] X. Yang, J. Wu, and Y. Dang, "A Product Recommendation Approach Based on the Latent Social Trust Network Model for Collaborative Filtering," IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), pp. 178-185, 2016

[6] W. Fan, J. Pei, and S. Ding,"A Novel Trust Inference Framework for Web-Based Scenarios Harnessed by Social Network and Web of Trust–a Heuristic Approach," Informatica, vol. 27(2), pp. 405-432, 2016.

[7] J. H. Choi, D. O. Kang, and J. Y. Jung, "Estimating Social Tie Strength for Autonomous D2D Collaborations," International Journal of Future Computer and Communication, 2015, vol. 4, pp. 1- 7, 2015.

[8] P. Nohuddin, F. Coenen, and R. Christley, "The application of social network mining to cattle movement analysis: introducing the predictive trend mining framework," Social Network Analysis and Mining, vol. 6(1), pp. 1-17, 2016.

[9] C. Jiang, S. Liu, and Z. Lin, "Domain-aware trust network extraction for trust propagation in large-scale heterogeneous trust networks," Knowledge-Based Systems, vol.111, pp. 237-247, 2016.

[10] W. Yuan, D. Guan, and T. Kim, "OS_Searching: Optimized Skeleton based Recommender Searching Mechanism for the Trust-Aware Recommender System," Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication. ACM, 2016: 98.

[11] W. Yuan, D. Guan, and S. Lee, "Skeleton Searching Strategy for Recommender Searching Mechanism of Trust-Aware Recommender Systems," The Computer Journal, vol. 58(9), pp. 1876-1883, 2015.

[12] H. Shirgahi, M. Mohsenzadeh, and H. Javadi, "Trust estimation of the semantic web using semantic web clustering,". Journal of Experimental & Theoretical Artificial Intelligence, vol. 29(3), pp. 537-556, 2017.

[13] W. Yuan, D. Guan, Y. K. Lee, "Improved trust-aware recommender system using small-worldness of trust networks," Knowledge-Based Systems, vol. 23(3), pp. 232-238, 2010.

[14] http://www.trustlet.org/episnions.html