PAPER

A New Similarity Model Based on Collaborative Filtering for New User Cold Start Recommendation

Ruilin PAN[†], Chuanming GE[†], Li ZHANG[†], Wei ZHAO^{†a)}, *Nonmembers*, and Xun SHAO^{††}, *Member*

SUMMARY Collaborative filtering (CF) is one of the most popular approaches to building Recommender systems (RS) and has been extensively implemented in many online applications. But it still suffers from the new user cold start problem that users have only a small number of items interaction or purchase records in the system, resulting in poor recommendation performance. Thus, we design a new similarity model which can fully utilize the limited rating information of cold users. We first construct a new metric, Popularity-Mean Squared Difference, considering the influence of popular items, average difference between two user's common ratings and non-numerical information of ratings. Moreover, the second new metric, Singularity-Difference, presents the deviation degree of favor to items between two users. It considers the distribution of the similarity degree of co-ratings between two users as weight to adjust the deviation degree. Finally, we take account of user's personal rating preferences through introducing the mean and variance of user ratings. Experiment results based on three real-life datasets of MovieLens, Epinions and Netflix demonstrate that the proposed model outperforms seven popular similarity methods in terms of MAE, precision, recall and F1-Measure under new user cold start condition.

key words: recommender systems, collaborative filtering, cold start problem, similarity model

1. Introduction

The development of Internet, World Wide Web and intelligent mobile technologies has made it easy to access an inconceivably huge amount of information, such as millions of web pages and purchase records of e-commerce products [1]. Then, the enormous volume of information often makes it different to search favorite products. The available choices increase exponentially, making it a challenge for users to find out their interested items, which is well-known as the information overload problem [2], [3]. As a powerful tool, RS break such dilemma [4]. RS are information search and decision support tools that use the historical track records of user activities and possibly personal profiles to recommend items (e.g., musics, restaurants and books) of interest to users automatically [5].

In literatures, there are mainly two recommendation methods, i.e., CF [6] and Content-based (CB) method [7].

In general, CF is a technology aiming to retrieve final recommendation results based on user's history of purchases or previous evaluation of items [8]. CB analyzes content information of items to match items and users. For example, keywords of purchased books might be used to find other books that contain the same or similar keywords [9]. In addition, combining the preceding techniques is also a way to exploit their advantages [10].

Currently, CF has been one of the most successful and commonly used personalized recommendation technologies since it only requires the information about user interactions [11]. As a result, many promising algorithms have been proposed based on CF. However, cold start problem is one of the most challenging problems including new user and new item cold start that have greatly captured the attention of researchers. The former is that there are some new users who have only a small number of items interaction or purchase records in the system [12]. The latter arises due to the fact that the new items entered in RS do not usually have initial ratings. Hence, data sparsity arises along with both of two problems.

Similarity measure plays an important role in recommending process. Recommendation quality can be improved giving a reasonable similarity measure under cold start conditions. In this paper, focusing on solving the above issues (especially the new user cold start problem), we propose a new similarity model based on CF. The main contributions can be summarized as follows: (1) we put forward a new similarity metric, Popularity-Mean Squared Difference (PMSD), which is composed of metrics Popularity and Mean Squared Difference (MSD) [13]. Popularity considers the influence of popularity of common rating items. MSD measures the average difference of two user's common ratings. Meanwhile, both of them take account of the non-numerical information of ratings; (2) to dig the information of co-ratings furtherly, the metric Singularity-Difference (SD) is proposed. It can allocate different penalty values based on the relationship between the difference of co-ratings and the median of all ratings in a dataset; (3) to reflect the overall rating preference of users, we introduce the mean and variance of user ratings in a metric Preference-Difference (PD).

We conduct a large number of experiments to validate our similarity model and compare it with other state-of-the-art models in terms of predictive accuracy (i.e., MAE) and recommendation performance (i.e., precision, recall and F1-Measure).

Manuscript received September 25, 2019.

Manuscript revised January 21, 2020.

Manuscript publicized March 3, 2020.

[†]The authors are with the School of Management Science and Engineering, Anhui University of Technology, Maanshan, 243032 China.

^{††}The author is with the School of Regional Innovation and Social Design Engineering, Kitami Institute of Technology, Kitami-shi, 090-8507 Japan.

a) E-mail: zhaoweistuart@gmail.com

DOI: 10.1587/transinf.2019EDP7258

The remainder of this paper is organized as follows. In Sect. 2, we review and describe related similarity measures for designing RS. Section 3 presents the mathematical formulation of the new similarity model in detail. The experimental results on three real-life datasets and the respective analysis are presented in Sect. 4. Finally, some conclusions and future work are put forward in Sect. 5.

2. Related Works

Researchers put forward some classic and famous similarity measures based on CF, such as cosine (COS) [14], Pearson correlation coefficient (PCC) [15] and Jaccard [16] similarity measures. Although these traditional recommendation algorithms were developed to improve RS performance, they were known to be vulnerable to the problem of cold start and data sparsity. There were some new similarity measures to deal with the new user cold start problem. Ahn [9] presented a heuristic similarity measure named PIP, which measured three levels: proximity, impact and popularity. However, its weakness was that it only considered the local information of the ratings without considering the global preference of user ratings. The combination of Jaccard and MSD was proposed to address the new user cold start problem [17]. It provided better recommendation results than traditional similarity measures by taking into consideration both numerical and non-numerical information. Bobadilla, et al. [18] proposed a new similarity measure based on neural network, which combined the numerical information and the distribution information of ratings. Liu, et al. [19] defined a similarity measure which used both local context information of user ratings and global preference of user behaviors. Furthermore, Barjasteh, et al. [20] introduced a decoupled similarity measure based on matrix completion that simultaneously exploited the similarity information among users and items.

Besides exploiting information from the direct relations among users or items (i.e., co-ratings), a number of research efforts dealt with the cold start problem by considering the indirect information. Physical dynamics based approaches including heat conduction [21], random walks [22], and mass diffusion [23], [24] were applied in many recommendation applications. Adding information of social network and location to traditional CF method is a new additional direction for improving the accuracy of rating prediction and recommendation quality [25]–[28].

The traditional approaches, e.g., COS, PCC and Jaccard, are widely used and the basis of proceeding methods with some limitations as follows: (I) overlook the influence of popular items on the similarity between users; (II) underutilize the available information of user's ratings including numerical and non-numerical information; (III) ignore the preference of co-ratings difference when using the co-ratings to calculate the similarity; (IV) fail to consider the individual preferences of each user.

3. Proposed Model

To improve performance, a new method should explore the limited rating information sufficiently from multiply aspects: (1) items' popularity, ratings' numerical information (detailed and local information) and non-numerical information (overall and global information); (2) the influence of co-rating difference, and the different sides of two ratings; (3) rating preference of different users. Considering the above factors, a new similarity model, PSP, is proposed, which is composed of three similarity metrics, PMSD, SD and PD. The similarity $PSP(u, v)$ between users u and v is calculated as Eq. (1):

$$PSP(u, v) = PMSD(u, v) \times SD(u, v) \times PD(u, v). \quad (1)$$

3.1 PMSD

In order to overcome limitations (I) and (II), a metric PMSD is proposed. Popularity in Eq. (3) can control the influence of popular items on similarity between users. The contribution of co-rated items to similarity opposites item popularity. Log function can make the data smooth and weaken the collinearity and heteroscedasticity of the model. Besides, it can fit the popularity in practice increasing with the rated times of items. For instance, the popularity of an item often increases quickly when it attracts attention initially. People will be accustomed to an item gradually with the increase of rated times, and then its popularity will increase slowly. To make full use of the available information of user's ratings, MSD in Eq. (4) considers the average rating difference between two users. Moreover, both Popularity and MSD take account of the non-numerical information of ratings, i.e., the number of common and total ratings of two users. Finally, Popularity and MSD are combined to form a new similarity metric, i.e., PMSD in Eq. (2).

$$PMSD(u, v) = Popularity(u, v) \times MSD(u, v), \quad (2)$$

$$Popularity(u, v) = \frac{\sum_{i \in I} \frac{1}{\log(1+N_i)}}{|I_u \cup I_v|}, \quad (3)$$

$$MSD(u, v) = \frac{\sum_{i \in I} (r_{u,i} - r_{v,i})^2}{|I|}, \quad (4)$$

where $r_{u,i}$ and $r_{v,i}$ denote the rating of item i by users u and v , respectively; I_u and I_v represent the set of items rated by users u and v , respectively; I represents the set of co-rated items by two users and N_i denotes the number of users who have rated item i .

3.2 SD

If co-ratings are different, their item favor should be different. Smaller difference of co-ratings to items of two users results in a less deviation degree of item favor and a higher similarity. Furthermore, there are several different cases of

co-rating difference, which gives rise to different contributions to the similarity between users. Hence, the distribution of co-rating difference should be considered to adjust the deviation degree. In our model, we consider these factors in Eq. (8) to overcome the limitation (III).

We let R_{max} and R_{min} be the maximum and minimum rating, and R_{med} in Eq. (5) is the median of all ratings in a dataset.

$$R_{med} = \frac{R_{max} + R_{min}}{2}, \quad (5)$$

$$d_{u,v} = \left\{ d = |r_{u,i} - r_{v,i}| \mid r_{u,i} \neq null, r_{v,i} \neq null, i \in I \right\}, \quad (6)$$

$$S_d(u, v) = \sum_{d \in d_{u,v}} (R_{med}^2 - t \times d)^2, \quad (7)$$

$$SD(u, v) = \sum_{d \in d_{u,v}} \frac{S_d(u, v) \times (R_{max} - d)}{\sum_{d \in d_{u,v}} S_d(u, v)}. \quad (8)$$

Equation (6) denotes a set of co-rating difference between two users. The metric $S_d(u, v)$ considers whether the two ratings are in agreement, and gives penalty to ratings in disagreement. Specifically, if two ratings are on the same side (two sides) of median, they are regarded as agreement (disagreement), i.e., if $r_{u,i} > R_{med}$ and $r_{v,i} < R_{med}$ or $r_{u,i} < R_{med}$ and $r_{v,i} > R_{med}$, then agreement ($r_{u,i}, r_{v,i}$) = false and $t = 2$; otherwise, agreement ($r_{u,i}, r_{v,i}$) = true and $t = 1$. Two user's ratings across the median towards an item give rise to a small similarity.

The metric SD in Eq. (8) refines PIP and MJD. $R_{med} - d$ denotes the deviation degree of item favor. The proportion, $\frac{S_d(u, v)}{\sum_{d \in d_{u,v}} S_d(u, v)}$, represents the weight of the similarity degree of co-ratings between two users. Therefore, SD can be more effective to distinguish the item favor of a user.

3.3 PD

It is especially important to design a similarity model to reflect the rating preference since some users prefer giving high ratings, while others tend to give low ratings. Therefore, to resolve the limitation (IV), we introduce the mean and variance of user ratings in Eqs. (9) and (10) to reflect the rating preference, respectively.

$$\mu_u = \frac{\sum_{i \in I_u} r_{u,i}}{|I_u|}, \quad (9)$$

Table 1 The combinations of mean and variance.

PD	Formula
PD1	$\frac{1}{1+ \mu_u-\mu_v }$
PD2	$\frac{1}{1+\sqrt{ \mu_u-\mu_v }}$
PD3	$\frac{1}{1+\log(\mu_u-\mu_v)}$
PD4	$\frac{1}{(1+ \sigma_u-\sigma_v) \times (1+ \mu_u-\mu_v)}$
PD5	$\frac{1}{(1+\sqrt{ \sigma_u-\sigma_v }) \times (1+\sqrt{ \mu_u-\mu_v })}$
PD6	$\frac{1}{(1+(1+\log(\sigma_u-\sigma_v)))} \times \frac{1}{(1+(1+\log(\mu_u-\mu_v)))}$

$$\sigma_u = \sqrt{\frac{\sum_{i \in I_u} (r_{u,i} - \mu_u)^2}{|I_u|}}. \quad (10)$$

To represent the difference of mean rating preference between two users, we introduce $\mu_u - \mu_v$. Meanwhile, the difference of variance in user's ratings, $\sigma_u - \sigma_v$, measures the rating volatility of two users. To reflect a proper variation degree of preference difference, we apply standard function, square root function and log function to the two differences. Generally speaking, there are six common combinations of mean and variance as listed in Table 1. Likewise, six combinations arise in PSP.

4. Experiments

4.1 Datasets

In experiments, we use three different real-life datasets MovieLens, Epinions and Netflix for three reasons: (1) they have been used by researchers and developers in the RS domain; (2) they are publicly open for research purpose and downloadable, and (3) they are substantially representative with various sparsity as benchmarking. A brief description of these datasets is provided in Table 2.

To create a new user cold start condition, each dataset will be divided into two parts, i.e., cold users and the other users [18], which are selected as validation and training users, respectively. The specific definition of these datasets is provided in Table 3. There are 86 cold users and 5954 other users in MovieLens, 10800 cold users and 14770 other users in Epinions, and 214 cold users and 44984 other users in Netflix.

4.2 Experimental Results

We make the following experimental design focusing on validating prediction accuracy and recommendation performance compared with several common similarity methods in new user cold start conditions. All algorithms are in

Table 2 Description of the datasets used in the experiments.

	MovieLens	Epinions	Netflix
Number of users	6040	49289	45198
Number of items	3952	139738	12422
Number of ratings	1000209	664824	2013034
Rating domain	1-5	1-5	1-5
Sparseness	91.62%	99.99%	99.28%
Availability location	http://www.grouplens.org/	http://www.epinions.com/	http://www.netflixprize.com/

Table 3 Definition of the datasets.

	MovieLens	Epinions	Netflix
Cold users	rated 2-20 items	rated 4-10 items	rated 4-10 items
Other users	rated more than 20 items	rated more than 10 items	rated more than 10 items

Fig. 1 MAE from different combination methods.

Python language and run by Spark cluster in Linux operating system.

We use MAE to measure prediction accuracy, precision, recall and F1-Measure [29] to evaluate recommendation quality. In our experiments, we compare PSP with not only part of the traditional similarity measures such as PCC [15], COS [14] and Jaccard [16], but also with some new similarity measures in CF including PIP [9], JMSD [17], MJD [18] and NHSM [19].

(1) Prediction Quality: MAE

We conduct experiments on MovieLens, Epinions and Netflix to compare the prediction quality of PSP1-6 under the different K-neighbors. In fact, the relation exists between MAE and K-neighbors. On one hand, the similarity of a user's neighbors decreases with the number of its neighbors, however, increases with MAE, especially under new user cold start condition. On the other hand, although MAE and computation complexity increase with the number of neighbors, diverse preferences are included and recommendation range is broadened. Based on a large number of experiments and existing work [19], we set the K-neighbors in the range of [10, 100] with step 10.

Figures 1 (a) and (b) show MAE results in MovieLens and Epinions datasets, giving a best result by PSP4. Figure 1 (c) shows a similar situation on Netflix, except MAE of PSP4 slightly worse than PSP1 in the first three points. It suitably demonstrates standard function reflects the variation extent of the difference to preference. Besides, the variance preforms well from the result pair of (PSP1, PSP4), (PSP2, PSP5) and (PSP3, PSP6), which verifies the importance of variance. Therefore, PSP4 is confirmed as the over-

Fig. 2 MAE in MovieLens, Epinions and Netflix.

all best option to obtain the prediction accuracy in this paper, since it offers a proper tradeoff between user's average and variance.

Figure 2 displays MAE of similarity measures with different K-neighbors on MovieLens, Epinions and Netflix. According to Fig. 2 (a), our proposal, i.e. the above-mentioned PSP4, achieves stably much lower MAE than all other measures. It verifies the effectiveness of our proposal that PSP can obtain more significant performance compared with the other similarity measures because PSP considers an additional important metric, the item popularity. It achieves obvious improvement compared with other similarity measures with advantage of 43.95% than PCC, 53.06% than PIP, 19.61% than MJD, 61.62% than Jaccard, 63.81% than COS, 55.67% than NHSM and 57.53% than JMSD, respectively.

Figures 2 (b) and (c) present MAE obtained on Epinions and Netflix, respectively. In Fig. 2 (b), we note that our proposal PSP obtains the best result than all other similarities with advantage of at most 50.65% than JMSD, 45.41% than NHSM, 49.95% than PIP, 66.22% than COS, 64.70% than Jaccard, 34.27% than MJD, and 45.51% than PCC, respectively. Moreover, PSP increases approximately with K, which means it has selected the most similar neighbors at K=10. Thus, increasing neighbors combines the diverse opinions of his neighbors, and weakens the influence of similar users. Figure 2 (c) with dataset Netflix shows similar results with Fig. 2 (b).

In general, PSP can achieve lower MAE than all the other similarity measures. This means that our method can predict user's preference more accurately, and further ensure the reliability of recommendation results.

Fig. 3 Precision, recall and F1-Measure on MovieLens.

Fig. 5 Precision, recall and F1-Measure on Netflix.

Fig. 4 Precision, recall and F1-Measure on Epinions.

(2) Recommendation Quality: precision, recall and F1-Measure

Figure 3 displays precision, recall and F1-Measure given Top-N on MovieLens. It is worth noting that PSP can get almost the best precision in addition to PCC in Fig. 3 (a).

However, PSP in Fig. 3 (b) achieves a moderate recall performance compared with other methods in the range [5, 30] and the best performance beyond 30. Figure 3 (c) shows similar performance from viewpoint of F1-measure with Fig. 3 (b).

Figure 4 gives the results of precision, recall and F1-Measure on sparse dataset Epinions. We can see that PSP obtains the best performance on three metrics in general. Besides, the recall of PSP has remarkable advantage over Jaccard, JMSD, PIP and NHSM. Figure 5 displays results on Netflix. Our proposal shows advantage at precision and F1-Measure, but disadvantage at recall.

In short, through the analysis of precision, recall and F1-Measure on three datasets of different sparse degree and data scale, it is clear that PSP can attain overall better recommendation quality. In particularly, precision results have remarkable advantage over most of the similarity measures. Thus, metrics, e.g., popularity, average rating difference, penalty for different rating difference and user rating preference, are important for RS performance.

5. Conclusions

In this paper, a new similarity model has been put forward focusing on the new user cold start problem. First, we present a new similarity metric, PMSD, which is composed of Popularity and MSD. The former considers the influence of popularity of the co-rating items, which can control the influence of popular items on the similarity between users. To make full use of the available information of user's ratings, the latter considers the average rating difference of two

users. Both of them take account of the non-numerical information of ratings. Second, the metric SD is proposed to represent the deviation degree of item favor. Third, to reflect the overall rating preference of users, we introduce the mean and variance of user ratings. Both the experiment results and the theoretical analysis show that the new similarity model can obtain better prediction accuracy and recommendation performance than other similarity measures.

In the future, our research directions are listed in the following points: (1) to combine the long-term or/and short-term interest of users to consider the popularity of items; (2) to utilize fully individual expression related to habits, consumption and tastes of cold start users; (3) to further extend and apply our new similarity model to actual e-business applications to measure the actual effect.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of China under Grant No. 71772002 and 61702006, the Key Laboratory of Multidisciplinary Management and Control of Complex Systems (Department of Education Anhui Province), and Anhui Province Key Laboratory of Affective Computing & Advanced Intelligent Machine under Grant No. ACAIM20180205.

References

- [1] T. Qiu, T.-T. Wang, Z.-K. Zhang, L.-X. Zhong, and G. Chen, "Alleviating bias leads to accurate and personalized recommendation," *EPL (Europhysics Letters)*, vol.104, no.4, p.48007, 2013.
- [2] M.A. Ghazanfar and A. Prügel-Bennett, "Leveraging clustering approaches to solve the gray-sheep users problem in recommender systems," *Expert Systems with Applications*, vol.41, no.7, pp.3261–3275, 2014.
- [3] V. Vijayakumar, S. Vairavasundaram, R. Logesh, and A. Sivapathi, "Effective knowledge based recommender system for tailored multiple point of interest recommendation," *International Journal of Web Portals (IJWP)*, vol.11, no.1, pp.1–18, 2019.
- [4] X. Zhu, H. Tian, P. Zhang, Z. Hu, and T. Zhou, "Personalized recommendation based on unbiased consistence," *EPL (Europhysics Letters)*, vol.111, no.4, p.48007, 2015.
- [5] V. Nikulin, "Hybrid recommender system for prediction of the yelp users preferences," *Industrial Conference on Data Mining*, vol.8557, pp.85–99, 2014.
- [6] M.Y.H. Al-Shamri and K.K. Bharadwaj, "Fuzzy-genetic approach to recommender systems based on a novel hybrid user model," *Expert Systems with Applications*, vol.35, no.3, pp.1386–1399, 2008.
- [7] D. Wang, Y. Liang, X. Dong, X. Feng, and R. Guan, "A content-based recommender system for computer science publications," *Knowledge-Based Systems*, vol.157, pp.1–9, 2018.
- [8] D.H. Park, H.K. Kim, I.Y. Choi, and J.K. Kim, "A literature review and classification of recommender systems research," *International Proceedings of Economics Development & Research*, vol.39, no.11, pp.10059–10072, 2012.
- [9] H.J. Ahn, "A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem," *Information Sciences*, vol.178, no.1, pp.37–51, 2008.
- [10] J.L. Jorro-Aragoneses, B. Díaz-Agudo, and J.A. Recio-García, "Addressing the cold-start problem in facial expression recognition," *International Conference on Case-Based Reasoning*, pp.197–211, Springer, 2015.
- [11] H. Chen, Z. Li, and W. Hu, "An improved collaborative recommendation algorithm based on optimized user similarity," *The Journal of Supercomputing*, vol.72, no.7, pp.2565–2578, 2016.
- [12] M. Zhang, J. Tang, X. Zhang, and X. Xue, "Addressing cold start in recommender systems: A semi-supervised co-training algorithm," *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pp.73–82, ACM, 2014.
- [13] F. Cacheda, V. Carneiro, D. Fernández, and V. Formoso, "Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems," *Acm Transactions on the Web*, vol.5, no.1, pp.1–33, 2011.
- [14] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE transactions on knowledge and data engineering*, vol.17, no.6, pp.734–749, 2005.
- [15] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "Grouplens: an open architecture for collaborative filtering of netnews," *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp.175–186, ACM, 1994.
- [16] G. Koutrika, B. Bercovitz, and H. Garcia-Molina, "Flexrecs: expressing and combining flexible recommendations," *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, pp.745–758, ACM, 2009.
- [17] 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.
- [18] J. Bobadilla, F. Ortega, A. Hernando, and J. Bernal, "A collaborative filtering approach to mitigate the new user cold start problem," *Knowledge-Based Systems*, vol.26, pp.225–238, 2012.
- [19] H. Liu, Z. Hu, A. Mian, T. Hui, and X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering," *Knowledge-Based Systems*, vol.56, no.3, pp.156–166, 2014.
- [20] I. Barjasteh, R. Forsati, D. Ross, A.H. Esfahanian, and H. Radha, "Cold-start recommendation with provable guarantees: A decoupled approach," *IEEE Transactions on Knowledge & Data Engineering*, vol.28, no.6, pp.1–1, 2016.
- [21] J.-G. Liu, T. Zhou, and Q. Guo, "Information filtering via biased heat conduction," *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol.84, no.2, p.037101, 2011.
- [22] M. Jamali and M. Ester, "Trustwalker: a random walk model for combining trust-based and item-based recommendation," *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.397–406, ACM, 2009.
- [23] T. Zhou, Z. Kucsik, J.-G. Liu, M. Medo, J.R. Wakeling, and Y.-C. Zhang, "Solving the apparent diversity-accuracy dilemma of recommender systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol.107, no.10, pp.4511–4515, 2010.
- [24] Y. Qian, Y. Zhang, X. Ma, H. Yu, and L. Peng, "Ears: Emotion-aware recommender system based on hybrid information fusion," *Information Fusion*, vol.46, pp.141–146, 2019.
- [25] W. Carrer-Neto, M.L. Hernández-Alcaraz, R. Valencia-García, and F. García-Sánchez, "Social knowledge-based recommender system. application to the movies domain," *Expert Systems with Applications*, vol.39, no.12, pp.10990–11000, 2012.
- [26] X. Wang, W. Lu, M. Ester, C. Wang, and C. Chen, "Social recommendation with strong and weak ties," *ACM International on Conference on Information and Knowledge Management*, pp.5–14, 2016.
- [27] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, and S. Wang, "Learning graph-based poi embedding for location-based recommendation," *Acm International on Conference on Information & Knowledge Management*, pp.15–24, 2016.
- [28] L. Xin, L. Yong, and X. Li, "Exploring the context of locations for personalized location recommendations," *International Joint Conference on Artificial Intelligence*, pp.1188–1194, 2016.

- [29] Á. Tejeda-Lorente, C. Porcel, E. Peis, R. Sanz, and E. Herrera-Viedma, "A quality based recommender system to disseminate information in a university digital library," *Information Sciences*, vol.261, no.5, pp.52–69, 2014.

Ruilin Pan received his Ph.D. degree in Enterprise Management from Dalian University of Technology in 2010. Currently he is a professor with the School of Management Science and Engineering at Anhui University of Technology, China. His research interests include machine learning, operation management and optimization algorithms.

Chuanming Ge received the B.S. degree in Industrial Engineering from the Anhui University of Technology, in 2017. He is currently working toward the M.Sc. degree in Management Science and Engineering at the School of Management Science and Engineering, Anhui University of Technology. Her research interest includes machine learning.

Li Zhang received the B.S. degree in Engineering Management from the Qingdao University of Technology, in 2016. She is currently working toward the M.Sc. degree in Management Science and Engineering at the School of Management Science and Engineering, Anhui University of Technology. Her research interest includes machine learning.

Wei Zhao received his Ph.D. in applied information science from Tohoku University in 2015. He was an "overseas researcher under Postdoctoral Fellowship of Japan Society for the Promotion of Science" (JSPS) with Prof. Takahiro Hara at Osaka University. He is currently an associate professor at Anhui University of Technology, China. His papers received best paper awards at GLOBECOM 2014 and WCSP 2014. His major research interests include wireless mesh networks and mobile ad hoc networks.

He is a member of IEEE.

Xun Shao received the Ph.D. degree from Osaka University, Japan, in 2013. He was a researcher at the National Institute of Information and Communications Technology (NICT), Japan, from 2013 to 2017. He is currently Assistant Professor with the School of Regional Innovation and Social Design Engineering, Kitami Institute of Technology, Japan. His research interests include computer networks and distributed systems. He is a member of IEICE and IEEE.