Future Generation Computer Systems ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization

Yin Zhang [a], Min Chen [b,*], Dijiang Huang [c], Di Wu [d], Yong Li [e]

[a] School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, China
[b] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China
[c] Department of Computer Science and Engineering, Arizona State University, USA
[d] School of Information Science & Technology, Sun Yat-Sen University, China
[e] Department of Electronic Engineering, Tsinghua University, China

HIGHLIGHTS

- We propose a topic model based approach to discover user preference distribution and doctor feature distribution.
- We propose an emotion-aware approach to identify emotional offset in user reviews via sentiment analysis.
- We incorporate topic model and emotional offset into the matrix factorization model.
- The experimental results show that the proposed model provides a high-performance healthcare recommendation.

ARTICLE INFO

Article history:
Received 17 June 2015
Received in revised form
26 November 2015
Accepted 3 December 2015
Available online xxxx

Keywords:
Healthcare
Recommendation
Matrix factorization
Sentiment analysis
Topic model

ABSTRACT

Nowadays, crowd-sourced review websites provide decision support for various aspects of daily life, including shopping, local services, healthcare, etc. However, one of the most important challenges for existing healthcare review websites is the lack of personalized and professionalized guidelines for users to choose medical services. In this paper, we develop a novel healthcare recommendation system called *iDoctor*, which is based on hybrid matrix factorization methods. *iDoctor* differs from previous work in the following aspects: (1) emotional offset of user reviews can be unveiled by sentiment analysis and be utilized to revise original user ratings; (2) user preference and doctor feature are extracted by Latent Dirichlet Allocation and incorporated into conventional matrix factorization. We compare *iDoctor* with previous healthcare recommendation methods using real datasets. The experimental results show that *iDoctor* provides a higher predication rating and increases the accuracy of healthcare recommendation significantly.

1. Introduction

With the rapid development of mobile networks such as the fifth generation (5G) system [1,2], a significant amount of professional knowledge from various sectors is available to Internet users at anywhere and can be accessed anytime to provide assistance decision [3,4]. For example, we can choose different movies according to the ratings on IMDB,[1] while the selection of restaurants, hotels and stores can be referred to other users' reviews on Yelp.[2] Similarly, the way that people choose medical service is changing with health related reviews websites, such as Vitals,[3] Healthgrades[4] and RateMDs,[5] etc. Through these websites, detailed information about doctors can be obtained for choosing doctor with an online appointment. This innovative process of medical consultation exhibits high efficiency compared to traditional onsite doctor selection [5]. However, several challenges exist to enable personalized and accurate medical services:

* Corresponding author.
E-mail addresses: yin.zhang.cn@ieee.org (Y. Zhang), minchen2012@hust.edu.cn (M. Chen), Dijiang.Huang@asu.edu (D. Huang), wudi27@sysu.edu.cn (D. Wu), liyong07@tsinghua.edu.cn (Y. Li).

1 http://www.imdb.com/.

2 http://www.yelp.com/.
3 http://www.vitals.com/.
4 http://www.healthgrades.com/.
5 http://www.ratemds.com.

http://dx.doi.org/10.1016/j.future.2015.12.001

Table 1
Emotional offset in user reviews.

Text
easy-going person. It is not always typical nowadays to get a PCP[a] who one feels comfortable speaking with. As for the staff, only two of the doctor's assistants were kind to my mother. The rest such as the lady who sits at the window to collect payments gives an attitude that the patient is a bother as they are never friendly nor show appreciation. The way we are treated is that we are bothering them which is not right. It is truly regretful.

[a] Primary care physician.

- **Personalized and professionalized demand**: It is very common nowadays for patients to search medical service by disease, but the search results may contain too many doctors to meet diverse needs [6]. Furthermore, user experience is always the most important for the system design [7]. For example, some users would much rather find a doctor nearby, while others prefer prescription to injection as treatment. Unfortunately, at present such personalization and professionalization demand cannot be satisfied intelligently according to user and doctor feature [8].
- **Emotional offset**: Like other sectors, it is a great issue for the medical crowd-sourced reviews that rating accuracy is often interfered by users emotion [9]. For example, a doctor's rating is 4 and a review about him is presented in Table 1. It can be concluded that this doctor is not so welcomed and such rating is possibly an encouragement, which has directly influenced the objectivity and accuracy of doctors estimation.

To address these challenges, this article proposes a personalized and professionalized doctor recommendation system named iDoctor, which can conduct comprehensive analysis on healthcare crowd-sourced reviews and perform text sentiment analysis, topic model, matrix factorization and other methods. Specifically, this article makes the following contributions: (1) We propose a topic model based approach to discover user preference distribution and doctor feature distribution, which are incorporated into the matrix factorization model to provide more accurate and personalized medical recommendation. (2) We propose an emotion-aware approach to identify emotional offset in user reviews via sentiment analysis, which is incorporated into the matrix factorization model to provide more objective recommendation.

The remainder of this article is organized as follows. Section 2 presents related works of matrix factorization, text sentiment analysis, and topic model. In Section 3, we introduce iDoctor architecture, theories foundation, and technical details. Section 4 analyzes the performance of recommendation provided by iDoctor, and compares it with other recommendations. Finally, we conclude this article in Section 5.

2. Related works

2.1. Matrix factorization

Nowadays, the recommendation based on matrix factorization proposed by Koren et al. in [10], has achieved acceptable result for rating prediction. Through this model, users and items are mapped to a low dimensional latent factor space which is the explanation to users ratings, and a user-item rating matrix is regarded as the product of user and item as presented in Eq. (1).

$$R_{m\times n} = P_{m\times k} * Q_{k\times n} \tag{1}$$

in which k represents the number of selected latent factors, P and Q represent the weights of each user and item for each characteristic in latent factor space, which are the result of rating matrix R factorization and used for rating prediction. Usually, Stochastic Gradient Descent (SGD) [11] is used for calculating P and Q. Remarkably, some latent factors can be ignored, so $k < m, n$. For example, category, director and actor attract more attention than duration and language, which can be ignored in matrix factorization.

Because of the good performance of matrix factorization, many researchers try to extend this work. In [12], Jamali et al. proposed a matrix factorization with trust propagation for recommendation in social networks. In [13], Baltrunas et al. proposed matrix factorization based approach for context aware recommendation.

2.2. Text sentiment analysis

User emotion plays an important role in market analysis, opinion mining and human–computer interaction, so more and more attention has been attracted to emotion recognition [14]. In general, it is complex for emotion recognition through language and facial expression, whereas the development of psychology and linguistics simplifies the process of text sentiment analysis. The text document contains not only topics but also user's emotional features that users expressions always correlate well with the emotion at that time.

At present, quite a few works try to integrate emotional actors into personalized recommendation [15]. In [16], Poirier et al. proposed a collaborative filtering recommendation according to sentiment analysis of user reviews instead of rating. In [17], Ko et al. proposed a hybrid recommendation algorithm based on content and collaborative filtering, in which the vector of item features and user emotion is calculated from item descriptions and user reviews.

2.3. Topic model

In the topic model, a document is regarded as the mixture and combination of multiple topics, and each word in the document is generated by such a procedure that a topic is selected with certain probability from document-topic distribution and then this word is selected with certain probability from topic-word distribution. Currently, the Latent Dirichlet Allocation (LDA) topic model is widely adopted for document topic extraction. Through LDA, topics and their probability distribution can be calculated for analyzing document similarity, which is essential for document classification and personalized recommendation.

Currently, there are many proposals that try to incorporate topic model into matrix factorization. In [18], Agarwal et al. represented item with some words, which are mapped to a multi-topic distribution, and provide recommendation through regression forecasting. In [19], McAuley et al. used topic model to extract item features from user reviews, integrate them with matrix factorization, and verify that the accuracy of this proposal is higher than rating matrix.

3. iDoctor: medical recommendation based on hybrid matrix factorization

In this article, we propose iDoctor to provide user with professionalized and personalized doctor recommendation through mining user emotion and preference from user rating and reviews about doctors. Specifically, it includes the following modules, and the architecture is illustrated in Fig. 1:

- **Sentiment analysis module**, which can calculate user emotional offset from user reviews text.
- **Topic modeling module**, which is used to extract the distribution of user preferences and doctor features.
- **Hybrid matrix factorization module**, which is integrated with two feature distributions extracted by LDA for rating prediction.

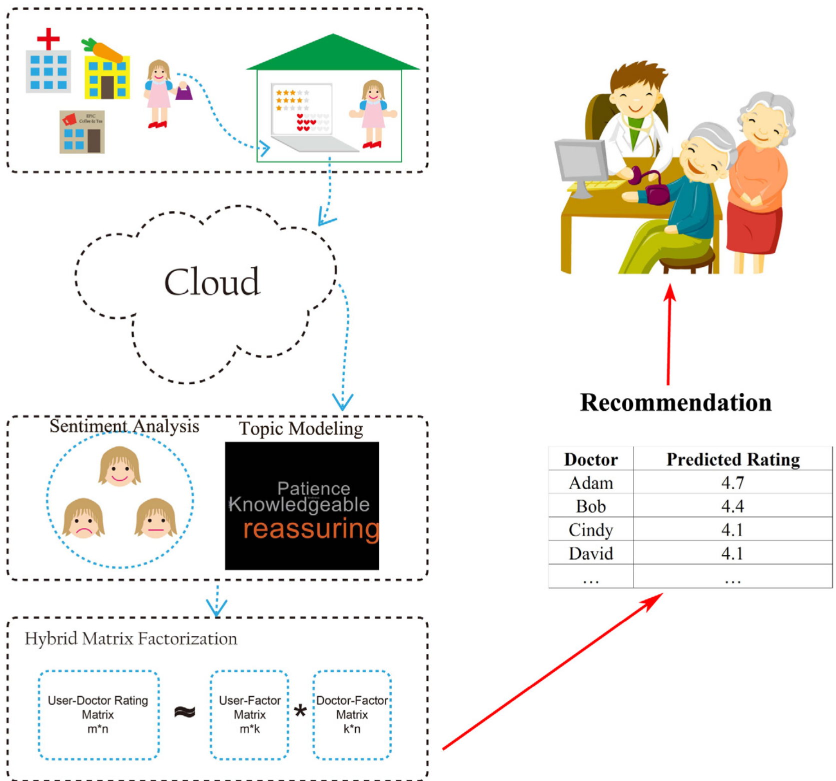


Fig. 1. Architecture of the iDoctor system.

3.1. Sentiment analysis for emotional offset

Considering the emotional offset in user reviews about doctor, sentiment analysis is necessary to calculate the offset for revising the original rating. Specifically, we calculate the emotional offset through non-supervisory learning method based on sentiment lexicon in the following steps:

1. Each user review text is preprocessed, including word segmentation and morphological normalization.
2. After removing stopword and punctuation, we have a collection of words which may involve emotional offset.
3. With sentiment lexicon SentiWordNet 3.0 [20], the emotional offset of each review can be calculated.
4. Furthermore, considering the influence of negative words, it is necessary to split each review in punctuation. In the separated subsentence, if the number of negatives is odd, the polarity (positive "+" or negative "−") of this subsentence should be reversed. Obviously, the overall emotional offset is the sum of all the subsentence.

Assume that there are i words in subsentences sub, we can calculate the sentiment value with Eqs. (2) and (3).

$$Pol(sub) = \begin{cases} -1 & \text{if the number of negative words in } sub \text{ is odd} \\ 1 & \text{if the number of negative words in } sub \text{ is even} \end{cases} \tag{2}$$

$$Sen(sub) = \left(\sum_{w_i \in sub} SentiWordNet(w_i) \times Pol(sub) \right). \tag{3}$$

Wherein, $Pol(sub)$ represents the polarity of sub; $SentiWordNet(w_i)$ represents the sentiment value of word i calculated according to SentiWordNet 3.0.

However, there is significant difference among user review lengths, and the emotional offset of longer review is most likely higher than shorter one. In order to balance the influences that the different length of review exert over overall variance, normalization is essential to calculate the overall emotional offset of user review.

Assume that there are j subsentences in review re including n sentiment words, we can calculate the emotional offset with Eq. (4).

$$Offset(re) = \frac{\sum_{sub_j \in re} Sen(sub_j)}{n}. \tag{4}$$

Obviously, the emotional offset of each user review falls in the range of $[-1, 1]$, which is used to revise the original rating.

3.2. Topic modeling for user preference

Generally, user reviews contain detailed personalized evaluation about doctors, such as their feelings towards medical environment, doctor's ability and attitude. On the one hand, every user preference can be discovered from these details, which is the basis for personalized recommendation. On the other hand, Table 2 describes that doctor features can be summarized from different user's reviews, such as specialty, fee range and prescribing habits, which also may be the basis of doctor selection.

Therefore, LDA model is adopted by iDoctor to extract the topics of user latent preference and doctor features can be extracted from user review comments on doctors, which are involved in matrix factorization for providing more accurate and personalized recommendation.

3.3. Hybrid matrix factorization for personalized and professionalized doctor recommendation

Based on the emotional offset calculated in Section 3.1, the original rating is expected to be revised with Eq. (5).

$$RS_{ij} = \rho R_{ij} + (1 - \rho) S_{ij} \tag{5}$$

in which RS_{ij} represents the revised rating on doctor j by user i, R_{ij} represent the original rating, S_{ui} represents the emotional offset,

Table 2
The features included in user reviews.

Text	Feature
Dr. Davis is a gem! Extremely reassuring, knowledgeable & smart. I had been fretting about an ongoing problem, wondering if it could be something else. My PCP just did his prescribing a pill call me in a week thing. I wanted to talk to someone & have an exam! Well Dr. Davis did just that & took his time to explain—what a novel idea! He had some recommendations which are helping me. The main thing was this man responded to me with kindness & concern & did his best to give me peace of mind.	Reassuring, knowledgeable, smart, took time to explain, kindness, concern

and *rho* is used for adjusting the weight between original rating and emotional offset. With Eq. (5), user emotional factor is involved in the matric factorization through the revised user–doctor rating matrix.

In Section 3.2, the preference topic distribution of user i, namely, $U_i = (x_1, x_2, \dots, x_k)$, and the features topic distribution of doctor j, namely, $I_j = (y_1, y_2, \dots, y_k)$, are calculated through LDA. In the above expressions, k represents the number of latent topics.

In the conventional matrix factorization, which is named Basic Matrix Factorization (BMF) in this article, the factorized matrixes cannot represent personalization. Hence, we propose Hybrid Matrix Factorization (HMF) involving user preference and doctor feature for more personalized recommendation, and loss function is defined as presented in Eq. (6),

$$J = \|RS - P \times Q^T\|_F^2 + \alpha\|U - P \times A\|_F^2 + \beta\|I - Q \times A\|_F^2 + \lambda(\|P\|_F^2 + \|Q\|_F^2 + \|A\|_F^2), \quad (6)$$

in which RS represents the rating matrix revised with emotional offset, P represents user latent factor matrix, Q represents doctor latent factor matrix, α is used for adjusting the weight of user preference distribution, while β is a parameter used for adjusting the weight of doctor feature distribution, and λ is used for regulating the weight of regularizing filter.

In order to achieve the best performance for HMF, Eq. (6) is expected to be minimized. After factorization, lower dimension matrixes of P, Q and A represent user, doctor, and the latent-topic-mapping matrix of latent-factor-rating matrix respectively. Finally, rating is available to be predicted through matrix multiply $P \times Q$, and latent factor matrixes P and Q are calculated by SGD, which includes the partial derivatives presented in Eqs. (7), (8), and (9).

$$\frac{\partial J}{\partial P} = -2(R - P \times Q^T)Q - 2\alpha(U - P \times A)A^T + 2\lambda P \quad (7)$$

$$\frac{\partial J}{\partial Q} = -2P^T(R - P \times Q^T) - 2\beta(I - Q \times A)A^T + 2\lambda Q \quad (8)$$

$$\frac{\partial J}{\partial A} = -2\alpha P^T(U - P \times A) - 2\beta Q^T(I - Q \times A) + 2\lambda A. \quad (9)$$

4. Experiments and analysis

4.1. Experimental data and evaluation standards

Experimental data in this article is obtained from a crowd-sourced review website called Yelp.[6] We select rating and reviews about three types doctor (i.e. *Internal Medicine Family Practice* and *Urgent Care*) in *Pittsburgh*, *Charlotte*, *Phoenix*and *Las Vegas* for experiment. All the data are preprocessed by Natural Language Toolkit (NLTK),[7] which is a leading platform for building Python programs to work with human language data, for word segmentation, removal of punctuation and stop word, and lemmatization. In the experiment, around 80% of this data are randomly selected for training, while the others are used for verifying the performance of our proposal. In particular, the experimental data includes 97 618 reviews about 8519 doctors submitted by 12 036 users.

Furthermore, Root Mean Square Error (RMSE) [21] is used for evaluating the accuracy of the proposed recommendation algorithm. With Eq. (10), the RMSE between predicted rating and actual rating can be calculated that the smaller the RMSE is, the better the performance of recommendation is.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - r_i)^2}. \quad (10)$$

In the above equation, n is the number of records in test data set, p_i represents the predicted rating, and r_i represents the actual rating.

Table 3
Parameters to be determined.

Parameter	Representation	Optimum
η	Learning rate	0.001
λ	Regularization parameter	0.1
N	Iterations	100
ρ	Weight of original rating	0.7
K	Number of latent topic	30
α	Weight of user preference	0.1
β	Weight of doctor feature	0.1

4.2. Experiment design and result analysis

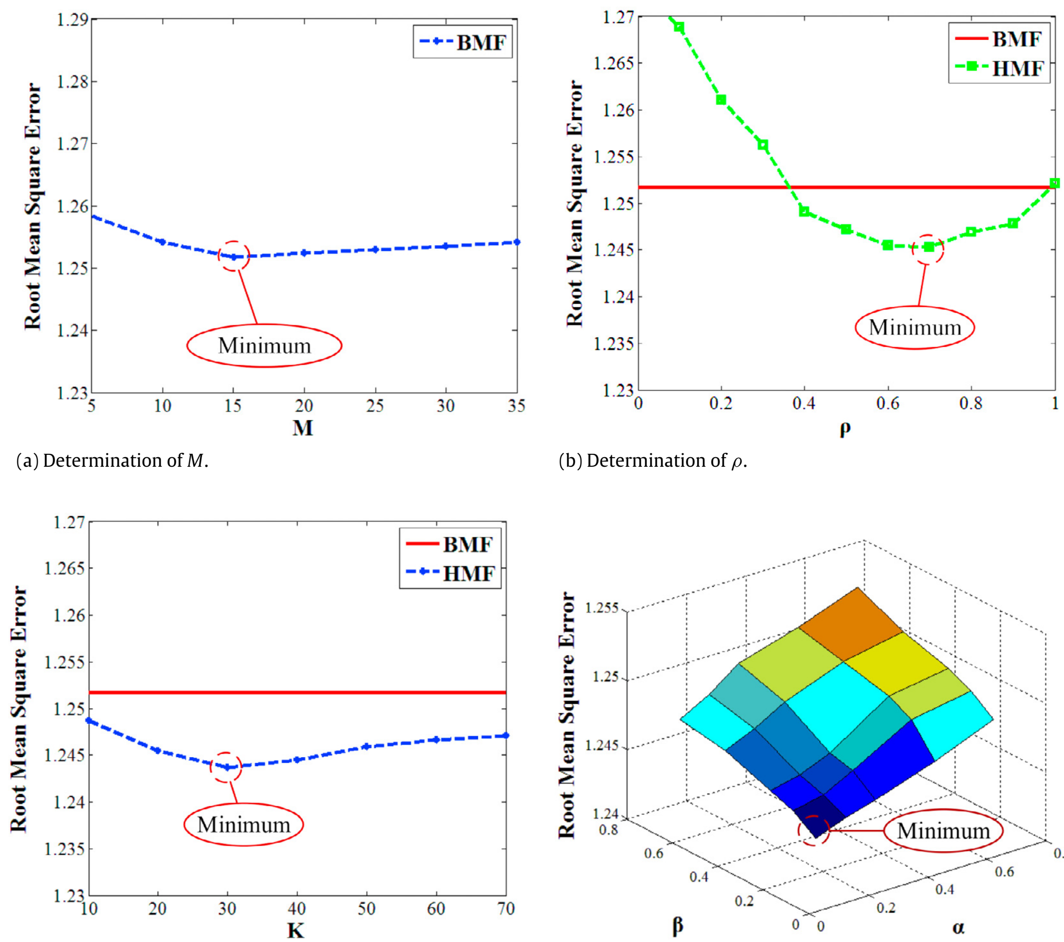
In this article, the experiment is designed for comparing our proposed HMF with BMF proposed in [11], and the hardware environment of our experiments is a computer with Intel Pentium Dual core CPU, 2.7 GHz dominant frequency, 8G primary memory. However, the parameters should be determined before the comparison. Table 3 shows the details of the parameters to be determined.

In Table 3, η, λ and N are the three basic parameters in SGD for matrix factorization. Through simple substitution, the results are satisfactory while $\eta = 0.001$, $\lambda = 0.1$, and $N = 100$. And the others are determined with the following determination of optimum parameters. ρ, K, α and β are only used in HMF, and we determined one parameter through fixing other three ones.

1. M: Both in BMF and HMF, M is expected to be determined. Based on the definition of loss function of BMF in [11], we evaluate RMSE with $M = [5, 10, \dots, 30, 35]$. As shown in Fig. 2(a), it is obvious that when $M = 15$, RMSE is minimum. Hence, the number of latent factors in HMF is determined to be 15 either.
2. ρ: We set $K = 20$, $\alpha = 0.1$, $\beta = 0.1$, and try to minimize RMSE with $\rho = [0, 0.1, \dots, 0.9, 1]$. In particular, $\rho = 0$ means to directly use emotional offset to predict rating, while $\rho = 1$ means to directly original rating. As shown in Fig. 2(b), which indicates that the weight of emotional offset should be relatively smaller than original rating, because the original rating is usually submitted after serious consideration and it can represent adequately user review.
3. K: We substitute $\rho = 0.7$ and set $\alpha = 0.1$, $\beta = 0.1$. Fig. 2(c) illustrates the experimental results for RMSE calculation with $K = [10, 20, \dots, 60, 70]$ that RMSE is minimized when $K = 30$. That is because if K is too small, the topic distributions of user preference and doctor feature cannot be represented sufficiently in HMF. Conversely, the convergence rates of HMF is too low to minimize RMSE.

6 http://www.yelp.com/dataset_challenge/.
7 http://www.nltk.org/.

Please cite this article in press as: Y. Zhang, et al., iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization, Future Generation Computer Systems (2016), http://dx.doi.org/10.1016/j.future.2015.12.001



(a) Determination of M.

(b) Determination of ρ.

(c) Determination of K.

(d) Determination of α and β.

Fig. 2. Determination of optimum parameters.

Table 4
Comparison between HMF, BMF, User-based and Item-based CF.

RMSE	*M*						
	5	10	15	20	25	30	35
Item-based CF	1.561	1.561	1.561	1.561	1.561	1.561	1.561
User-based CF	1.562	1.561	1.561	1.561	1.561	1.561	1.561
BMF	1.254	1.253	1.252	1.252	1.253	1.254	1.254
HMF	1.248	1.246	1.244	1.245	1.246	1.248	1.249

4. α and β: According to the representation, we try to find the optimum of $\alpha, \beta \in [0, 1]$ by grid search [22], and Fig. 2(d) shows that RMSE is minimized when $\alpha = 0.1$, $\beta = 0.1$. It indicates that the weights of user preference and doctor feature, which are expected to revise BMF for more accurate prediction, should be relative small.

With determined optimum parameters, RMSE of HMF is lower than that of BMF which means HMF-based iDoctor based can provide more accurate recommendation. Furthermore, considering the similarity between the number of latent factors in matrix factorization (i.e. M) and the number of neighborhood in collaborative filtering (CF) which is one of the most representative techniques used by recommendation systems, item-based CF and user-based CF are available to be involved into the comparison. As shown in Table 4, RMSE of HMF is significantly lower than that of other models, and it proves that iDoctor can provide more accurate doctor recommendation.

5. Conclusions

It is challenging to address the problem that user choose doctor online without sufficient personalized and personalized instruction. To solve the problem, proposed iDoctor to (1) discover emotional rating from user reviews to revise user original rating; (2) discover topic distributions of user preference and doctor feature to improve conventional matrix factorization. Our experiment results proved that the prediction of proposed HMF is better than BMF, item-based CF, and user-based CF, and iDoctor can provide considerable accurate recommendation. In our future work, we plan to take the time-varying possibility of user preferences into account and import data from social networks to further increase the recommendation accuracy, and try to develop a general system like iDoctor to provide more personalized, professionalized and objective recommendation.

Acknowledgments

The research reported here was supported by China National Natural Science Foundation under Grants 61572220. Prof. Di Wu's work was supported in part by the National Natural Science Foundation of China under Grant Nos. 61272397 and 61572538, and the Guangdong Natural Science Funds for Distinguished Young Scholar under Grant No. S20120011187.

Please cite this article in press as: Y. Zhang, et al., iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization, Future Generation Computer Systems (2016), http://dx.doi.org/10.1016/j.future.2015.12.001

References

[1] K. Zheng, L. Zhao, J. Mei, M. Dohler, W. Xiang, Y. Peng, 10 gb/s hetsnets with millimeter-wave communications: access and networking-challenges and protocols, IEEE Commun. Mag. 53 (1) (2015) 222–231.

[2] X. Ge, H. Cheng, M. Guizani, T. Han, 5g wireless backhaul networks: challenges and research advances, IEEE Netw. 28 (6) (2014) 6–11.

[3] M.S. Hossain, G. Muhammad, Cloud-assisted speech and face recognition framework for health monitoring, Mob. Netw. Appl. (2015) 1–9.

[4] D. Zhang, C.-H. Hsu, M. Chen, Q. Chen, N. Xiong, J. Lloret, Cold-start recommendation using bi-clustering and fusion for large-scale social recommender systems, IEEE Trans. Emerging Top. Comput. 2 (2) (2014) 239–250.

[5] T.R. Hoens, M. Blanton, A. Steele, N.V. Chawla, Reliable medical recommendation systems with patient privacy, ACM Trans. Intell. Syst. Technol. (TIST) 4 (4) (2013) 67.

[6] M. Hossain, Cloud-supported cyber-physical localization framework for patients monitoring, IEEE Syst. J. PP (99) (2015) 1–10. http://dx.doi.org/10.1109/JSYST.2015.2470644.

[7] K. Zheng, X. Zhang, Q. Zheng, W. Xiang, L. Hanzo, Quality-of-experience assessment and its application to video services in lte networks, IEEE Wirel. Commun. 22 (1) (2015) 70–78.

[8] G. Luo, S.B. Thomas, C. Tang, Automatic home medical product recommendation, J. Med. Syst. 36 (2) (2012) 383–398.

[9] Y. Tan, Y. Zhang, M. Zhang, Y. Liu, S. Ma, A unified framework for emotional elements extraction based on finite state matching machine, in: Natural Language Processing and Chinese Computing, Springer, 2013, pp. 60–71.

[10] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer (8) (2009) 30–37.

[11] R. Gemulla, E. Nijkamp, P.J. Haas, Y. Sismanis, Large-scale matrix factorization with distributed stochastic gradient descent, in: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2011, pp. 69–77.

[12] M. Jamali, M. Ester, A matrix factorization technique with trust propagation for recommendation in social networks, in: Proceedings of the Fourth ACM Conference on Recommender Systems, ACM, 2010, pp. 135–142.

[13] L. Baltrunas, B. Ludwig, F. Ricci, Matrix factorization techniques for context aware recommendation, in: Proceedings of the Fifth ACM Conference on Recommender Systems, ACM, 2011, pp. 301–304.

[14] M.S. Hossain, G. Muhammad, B. Song, M.M. Hassan, A. Alelaiwi, A. Alamri, Audio-visual emotion-aware cloud gaming framework 25 (12).

[15] Y. Moshfeghi, J.M. Jose, Role of emotional features in collaborative recommendation, in: Advances in Information Retrieval, Springer, 2011, pp. 738–742.

[16] D. Poirier, F. Fessant, I. Tellier, Reducing the cold-start problem in content recommendation through opinion classification, in: 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Vol. 1, IEEE, 2010, pp. 204–207.

[17] M. Ko, H.W. Kim, M. Yi, J. Song, Y. Liu, Moviecommenter: Aspect-based collaborative filtering by utilizing user comments, in: 2011 7th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), IEEE, 2011, pp. 362–371.

[18] D. Agarwal, B.-C. Chen, flda: matrix factorization through latent dirichlet allocation, in: Proceedings of the Third ACM International Conference on Web Search and Data Mining, ACM, 2010, pp. 91–100.

[19] J. McAuley, J. Leskovec, Hidden factors and hidden topics: understanding rating dimensions with review text, in: Proceedings of the 7th ACM Conference on Recommender Systems, ACM, 2013, pp. 165–172.

[20] S. Baccianella, A. Esuli, F. Sebastiani, Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining, in: LREC, Vol. 10, 2010, pp. 2200–2204.

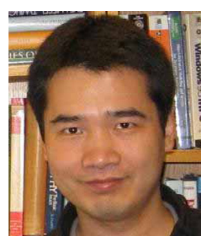
[21] A. Zollanvari, E.R. Dougherty, Moments and root-mean-square error of the Bayesian mmse estimator of classification error in the Gaussian model, Pattern Recognit. 47 (6) (2014) 2178–2192.

[22] T. Nakagawa, Y. Iwahori, M. Bhuyan, Defect classification of electronic board using multiple classifiers and grid search of svm parameters, in: Computer and Information Science, Springer, 2013, pp. 115–127.

Yin Zhang is a faculty member of the School of Information and Safety Engineering, Zhongnan University of Economics and Law. He worked as a Poster-Doctoral Fellow in the School of Computer Science and Technology at Huazhong University of Science and Technology (HUST) for two years. He is a handling Guest Editor for New Review of Hypermedia and Multimedia. He serves as reviewer for IEEE Network, Information Sciences. He is TPC Co-Chair for 5th International Conference on Cloud Computing (CloudComp 2014). He is Local Chair for the 9th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities (TRIDENTCOM 2014), and CloudComp 2013.

Min Chen is a Professor in School of Computer Science and Technology at Huazhong University of Science and Technology (HUST). He is the Director of Embedded and Pervasive Computing (EPIC) lab. He was an Assistant Professor in School of Computer Science and Engineering at Seoul National University (SNU) from Sep. 2009 to Feb. 2012. He worked as a Post-Doctoral Fellow in Department of Electrical and Computer Engineering at University of British Columbia (UBC) for three years. Before joining UBC, he was a Post-Doctoral Fellow at SNU for one and half years. He received Best Paper Award from IEEE ICC 2012, and Best Paper Runner-up Award from QShine 2008. He is a Guest Editor for IEEE Network, IEEE Wireless Communications Magazine, etc. He is Co-Chair of IEEE ICC 2012-Communications Theory Symposium, and Co-Chair of IEEE ICC 2013-Wireless Networks Symposium. He is General Co-Chair for the 12th IEEE International Conference on Computer and Information Technology (IEEE CIT-2012) and Mobimedia 2015. He is General Vice Chair foTridentcom 2014. He is Keynote Speaker for CyberC 2012, Mobiquitous 2012 and Cloudcomp 2015. He has more than 260 paper publications, including 120+ SCI papers, 50+ IEEE Trans./Journal papers, 6 ISI highly cited papers and 1 hot paper. He has published a book on IoT: OPNET IoT Simulation (2015) with HUST Press, and a book on big data: Big Data Related Technologies (2014) with Springer Series in Computer Science. His Google Scholars Citations reached 5400+ with an h-index of 35. His top paper was cited 656 times, while his top book was cited 420 times as of Aug. 2015. He is an IEEE Senior Member since 2009. His research focuses on Internet of Things, Mobile Cloud, Body Area Networks, Emotion-aware Computing, Healthcare Big Data, Cyber Physical Systems, and Robotics, etc.

Dijiang Huang received his Bachelor of Science degree in Telecommunications from Beijing University of Posts & Telecommunications in 1995. He received his Master of Science and Ph.D. degrees from University of Missouri-Kansas City in 2001 and 2004, respectively, both majored in Computer Science and Telecommunications. He joined Computer Science and Engineering Department at ASU in 2005. Dr. Huang's research is supported by NSF, ONR, ARO, HP, and Consortium of Embedded System (CES). He is a recipient of ONR Young Investigator Award and HP Innovative Research Award.



Di Wu is an Associate Professor and Associate Department Head in the Department of Computer Science, Sun Yat-sen University, Guangzhou, China. He received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 2000; the M.S. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2003; and the Ph.D. degree in Computer Science and Engineering from the Chinese University of Hong Kong, Shatin, Hong Kong, in 2007. During 2007–2009, he worked as a postdoctoral researcher in the Department of Computer Science and Engineering, NYU Polytechnic School of Engineering (previously Polytechnic Institute of NYU), advised by Prof. Keith W. Ross. He was also working closely with Prof. Yong Liu and Prof. Joel Wein. He was a co-recipient of IEEE INFOCOM 2009 Best Paper Award. He has served as the Media Streaming IG Chair (2014–2016) of IEEE Communications Society Multimedia Communications Technical Committee, the TPC Co-chair of IEEE GLOBECOM-CCSNA Workshop 2014, Chair of CCF YOCSEF-Guangzhou (2014–2015), Member of the Council of China Computer Federation (2012–2015). He has also served as the Guest Editor of IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on "Visual Computing in the Cloud: Mobile Computing", and the Editor of SCI-indexed journals, such as Springer Telecommunication Systems, Springer Peer-to-Peer Networking and Applications, Wiley Security and Communication Networks. His research interests include Multimedia Networking, Cloud Computing, Big Data Analytics, Internet Measurement, and Network Security.

Yong Li received the B.S. degree in Electronics and Information Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007 and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2012. During July–August 2012 and 2013, he was a Visiting Research Associate with Telekom Innovation Laboratories and The Hong Kong University of Science and Technology, respectively. During December 2013–March 2014, he was a Visiting Scientist with the University of Miami. He is currently a Faculty Member of the Department of Electronic Engineering, Tsinghua University. His research interests are in the areas of networking and communications.