❒ 884

A doctor recommender system based on collaborative and content filtering

Qusai Y. Shambour[1], Mahran M. Al-Zyoud[2], Abdelrahman H. Hussein[3], Qasem M. Kharma[1]
[1]Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan
[2]Department of Networks and Information Security, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan
[3]Department of Computer Science, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

Article Info

ABSTRACT

The volume of healthcare information available on the internet has exploded in recent years. Nowadays, many online healthcare platforms provide patients with detailed information about doctors. However, one of the most important challenges of such platforms is the lack of personalized services for supporting patients in selecting the best-suited doctors. In particular, it becomes extremely time-consuming and difficult for patients to search through all the available doctors. Recommender systems provide a solution to this problem by helping patients gain access to accommodating personalized services, specifically, finding doctors who match their preferences and needs. This paper proposes a hybrid content-based multi-criteria collaborative filtering approach for helping patients find the best-suited doctors who meet their preferences accurately. The proposed approach exploits multi-criteria decision making, doctor reputation score, and content information of doctors in order to increase the quality of recommendations and reduce the influence of data sparsity. The experimental results based on a real-world healthcare multi-criteria (MC) rating dataset show that the proposed approach works effectively with regard to predictive accuracy and coverage under extreme levels of sparsity.

Corresponding Author:

Qusai Y. Shambour
Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University
Amman, Jordan
Email: q.shambour@ammanu.edu.jo

1. INTRODUCTION

With the rapid expansion of Internet applications and services, a large amount of professional knowledge from a variety of domains can be accessed at any time to provide assistance decisions to internet users anywhere. In the healthcare domain, for example, a number of healthcare platforms, such as RateMDs (ratemds.com), Vitals (vitals.com), Tebcan (tebcan.com), and Healthgrades (healthgrades.com), provide detailed information about doctors that can be utilized by patients to search for satisfactory doctors [1]. Due to the advancement of online healthcare information, it becomes extremely time-consuming and difficult for patients to search through all the available doctors, and are likely to depend on word-of-mouth recommendations from relatives and friends to find the best-suited doctors [1]–[3]. To address the above issue, recommender systems can be regarded as an effective solution due to their ability to reduce the barriers of information overload. By utilizing personalized recommender systems, patients can find doctors that best meet their preferences [1]–[9].

Over the last decades, recommender systems have been successfully employed in various fields, such as recommending music, books, e-government services, business partners, online advertising, tourism trips, hotels, movies, e-learning materials, jobs, jokes, software requirements, websites, and scientific research papers to deal with the information overload problem. Recommender systems can be considered as personalized decision support systems that assist users in choosing needed services or items in accordance with their preferences from an extensive range of possible options. Content-based, neighborhood-based collaborative filtering (CF), and hybrid-based approaches are all common recommendation techniques. Each recommendation approach has its advantages and drawbacks; for instance, CF has data sparsity and cold-start issues, whereas content-based has overspecialized recommendations. Hybrid recommendation approaches, which combine the best qualities of two or more recommendation approaches, have been developed to overcome their challenges [10]–[12].

Content-based recommendation approaches suggest items that are similar to items that a user has previously preferred. It works by first determining the main common features, that can be utilized to recommend items, by analyzing the descriptions of the items favored by a certain user. Then, the features of potential items are compared with the favored features of the user to decide whether an item should be recommended or not [13]. CF-based approaches, also known as memory-based approaches, were amongst the earliest algorithms developed for recommender systems. These approaches are based on the fact that alike users will demonstrate similar behavior in rating and alike items should receive similar ratings. Essentially, there are two types of neighborhood-based CF approaches: user-based CF and item-based CF approaches. This study focuses on the item-based CF approach that is extensively implemented in real-world applications. In the item-based CF approach, the ratings provided by an active user a are used to generate recommendations for him/her. To predict the rating of the target item x by user a, similarity measures are used to identify a set of k-nearest neighbors of items, that are very similar to the target item x, and the ratings provided by user a for these similar items are utilized for computing the rating prediction [14]. Moreover, neighborhood-based CF methods face various limitations because of data sparsity. This is because users typically rate a small number of available objects, resulting in a small number of common ratings between each pair of users or items. Therefore, it becomes unlikely to successfully locate k-nearest neighbors, which affects the performance and accuracy of the neighborhood-based CF approaches [15].

Most current neighborhood-based CF recommendation approaches use a single criteria rating, an overall rating, to quantify the user preferences about items, which does not reflect the detailed preference of each aspect of an item. Moreover, many websites currently allow users to rate items in multiple dimensions. In particular, in the healthcare domain, patients can rate their doctors based on more than one aspect, for example: at RateMDs, as shown in Figure 1, patients can rate doctors based on four criteria including staff, punctuality, helpfulness, and knowledge; whereas, at vitals, patients can rate doctors based on seven criteria including promptness, easy appointments, accurate diagnosis, friendly staff, spends time with patients, bedside manner, and appropriate follow-up. Accordingly, it becomes indispensable to design and develop multi-criteria recommender systems that can exploit the additional rating information, to precisely understand the preferences of users, which contribute to more accurate and effectual recommendations. With regard to the recommendation process, multi-criteria based CF can deliver more accurate recommendations by taking into account the knowledge of essential aspects that lead the users to choose a specific item [16], [17].
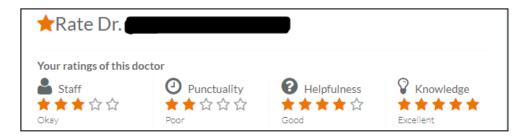

Figure 1. Example of multi-criteria rating of a doctor on RateMDs

In response to the above issues, a hybrid content-based multi-criteria CF approach for helping patients find the best-suited doctors, who meet their preferences accurately, is developed. The proposed approach utilizes multi-criteria decision making, doctor reputation score, and content information of doctors in order to improve the quality of recommendations and reduce the effect of data sparsity. The experimental results based on a real-world healthcare multi-criteria (MC) rating dataset illustrate that the proposed approach works effectively with regard to predictive accuracy and coverage under extreme levels of sparsity.

The rest of this paper is organized as. A summary of related works on the doctor recommendations domain is presented in section 2. Section 3 demonstrates the modules of the proposed approach, while section 4 presents the experimental results in detail. Finally, the study conclusion and directions for future research are illustrated in section 5.

2.   LITERATURE REVIEW

Over the past few years, technological advancements have spawned demands for innovation in all fields. In the healthcare domain, the application of recommender systems has attracted the attention of many researchers [1], [3]–[7], [9], [18], [19]. Although a number of related works regarding the implementation of recommender systems in healthcare have been published, only very limited work has been reported on doctor recommender systems [3].

Narducci *et al.* [18] proposed a semantic-based recommender system for recommending hospitals and doctors to patients based on their profiles. The proposed system first computes semantic similarities between patients, and then produces a ranked list of hospitals and doctors that best fit the patient profile. The proposed recommender system is embedded in the social network named HealthNet, and the main purpose is to share knowledge, find similar patients, and look at their experiences. On the other hand, Zhang *et al.* [1] proposed a healthcare recommender system, called iDoctor, that can utilize patients' ratings and reviews about doctors to provide patients with personalized doctor recommendations. The proposed system performs complete analysis on healthcare crowd-sourced reviews using text sentiment analysis, topic modeling, and hybrid matrix factorization methods in order to provide personalized and accurate doctor recommendations. The experimental results proved that iDoctor provides more accurate recommendations than other CF-based recommendation approaches. In the study of Han *et al.* [19], a hybrid recommender system that provides a list of personalized doctor recommendations to patients is developed. The proposed system utilizes an extensive dataset of consultation histories to model patients' trust in doctors. In addition, it computes similarities among patients and doctors based on their metadata. The proposed system helps a leading European healthcare provider in Portugal to renovate their primary care health service by restructuring their appointment system for family doctors to reduce the search burden for patients. In terms of predictive accuracy, the experimental evaluation confirms the efficacy of the proposed approach when compared to the heuristic baseline and CF-based recommendation approaches.

Waqar *et al.* [3] presented an effective doctor recommender system. The proposed system uses an adoptive algorithm to construct a doctor's ranking function, which is used to transform patients' criteria for choosing a doctor into a numerical base rating. This rating is then exploited by various machine-learning techniques to generate personalized recommendations of doctors to patients. The system has been validated by domain experts, and the results show that the recommendations of doctors are reasonable in that they can match patients' needs effectively. Yang *et al.* [6] proposed a decision support model that recommends proper doctors for patients on haodf.com. The proposed model includes four modules: a transformation module to convert raw data into Intuitionistic fuzzy sets, an integration module to combine interdependent information, a three-cloud presentation module to accommodate patient preferences, and a recommendation module to produce a personalized ranked list of doctors for a target patient. Validation results of the proposed model, on the haodf.com dataset, show the improvements in terms of the diversity and coverage of doctor recommendations when compared to the existing haodf.com approach. Meng and Xiong [5] proposed a doctor recommendation algorithm based on an online healthcare platform. The proposed algorithm uses the textual information of doctor-patient consultations, the latent Dirichlet distribution topic model, and other methods to locate doctors who best suit the needs of patients. The experimental results, using data from a Chinese healthcare website, show the effectiveness of the proposed method.

Even though a limited number of doctor recommender systems have been reported in the literature, they are still suffering from sparse rating data due to the lack of rating information that is inherent in the healthcare domain. Furthermore, to the best of our knowledge, there is currently no published research on the application of multi-criteria recommender systems in the doctor recommendation domain. Accordingly, the development of an effective doctor recommender system that utilizes the multi-criteria ratings of doctors and addresses the sparsity challenge is essentially required to be considered in the healthcare domain.

3.   THE PROPOSED METHOD

In this section, the framework of the proposed approach is explained. The proposed approach for making recommendations consists of three modules: i) the MC Item-based CF module, ii) the item-based content module, and iii) the hybrid prediction module. Henceforward, patients are adverted to as users, and doctors are adverted to as items.

3.1. The MC item-based CF module

Suppose that there are m patients represented as $P = \{p_1, p_2, \dots p_m\}$; and $D = \{d_1, d_2, \dots d_n\}$ be a set of n doctors rated by patients in P. In addition, let $\{c_1, c_2, \dots, c_k\}$, be a set of evaluation criteria upon which a doctor d is rated upon, each criterion is an aspect of a doctor with a rating value. The patient-doctor MC rating matrix, $R = (r_{p,d,c})_{m*n*k}$, represents the MC rating of patient p on criteria c for doctor d.

In this module, at first, an improved metric for MC item-based CF similarity that considers global similarity, local similarity, and structural similarity information is proposed to enhance the accuracy of prediction. In terms of global similarity, the Bhattacharyya coefficient is employed as a similarity measure due to its effectiveness in extracting global information from the sparse rating datasets in respect of the classical item-based CF similarity techniques [20]. Accordingly, the Bhattacharyya coefficient is used to compute the individual similarities between the items d_i and d_j based on each of the rating criteria c as shown in (1):

$$iBC_{d_i,d_j}^c = \sum_{h=1}^{x} \sqrt{\left(\frac{\#h}{\#d_i}\right)\left(\frac{\#h}{\#d_j}\right)} \tag{1}$$

where x is the number of bins, $\#h$ is the number of users who rated the item with rating value h, $\#d_i$ and $\#d_j$ are the numbers of users who rated items d_i and d_j, respectively. Then, the worst-case similarity [21] is used as an aggregation approach to realize the overall similarity value between given items as (2):

$$OBC_{d_i,d_j} = \min_{c=1,\dots,k} iBC_{d_i,d_j}^c \tag{2}$$

where iBC_{d_i,d_j}^c is the value of individual similarity between items d_i and d_j in terms of the criteria c, k is the number of criteria.

In respect of the local similarities between items, we used the Cosine similarity measure [22]. First, the individual similarities between any given pair of items in terms of each of the rating criteria c are calculated as (3):

$$iCos_{d_i,d_j}^c = \frac{\sum_{p \in M} r_{p,d_i}^c \times r_{p,d_j}^c}{\sqrt{\sum_{p \in M} (r_{p,d_i}^c)^2} \times \sqrt{\sum_{p \in M} (r_{p,d_j}^c)^2}} \tag{3}$$

where r_{p,d_i}^c and r_{p,d_j}^c correspond to the ratings of the user p on items d_i and d_j in terms of the criteria c respectively. M denotes the users who rated both items. Then, the average similarity [21] is utilized to aggregate all individual similarities to compute the overall similarity as (4).

$$OCos_{d_i,d_j} = \frac{\sum_{c=1,\dots,k} iCos_{d_i,d_j}^c}{k} \tag{4}$$

Finally, concerning the structural similarity, the percentage of users that have commonly rated both items d_i and d_j is calculated using the Jaccard coefficient [23].

$$Jacc_{d_i,d_j} = \frac{|U_{d_i} \cap U_{d_j}|}{|U_{d_i}| + |U_{d_j}| - |U_{d_i} \cap U_{d_j}|} \tag{5}$$

Where $\left|U_{d_i} \cap U_{d_j}\right|$ is the overall number of users that have rated both items d_i and d_j. Eventually, the improved metric for MC item-based CF similarity for any given pair of items is devised as (6).

$$TMCSim_{d_i,d_j} = OBC_{d_i,d_j} \times OCos_{d_i,d_j} \times Jacc_{d_i,d_j} \tag{6}$$

Furthermore, the item reputation score is introduced to improve the approach's capacity to predict unobserved items that are caused by the lack of reliable nearest neighbors due to the sparsity challenge. The item reputation score is computed based on the average variation between its ratings and the users' mean ratings, in addition to the number of connections the item has with other items in the item-item similarity matrix, as revealed.

$$Rep_{d_i} = exp\left(-\frac{\sum_{p \in U_{d_i}}|r_{p,d_i} - \bar{r}_p|}{|U_{d_i}|}\right) \times \sqrt{\frac{|I_{d_i}|}{|I|}} \tag{7}$$

Where r_{p,d_i} is the overall rating of user p on item d_i, $\bar{r}_p$ is the mean rating of user p, and U_{di} is the set of users who rated item d_i. $|I_{di}|$ is the total number of items that have similarity relationships with item d_i, and $|I|$ is the total number of items in the dataset.

For predictions, the mean-based prediction metric [24] is in use to generate MC Item-based predicted ratings as (8).

$$Pred_{p,d_i}^{MCCF} = \begin{cases} \bar{r}_{d_i} + \frac{\sum_{d_j \in NN} TMCSim_{d_i,d_j} \times \left(r_{p,d_j} - \bar{r}_{d_j}\right)}{\sum_{d_j \in NN} TMCSim_{d_i,d_j}}; & if \quad TMCSim_{d_i,d_j} \neq 0 \\ \bar{r}_{d_i} + \frac{\sum_{d_j \in NN} IR_{d_j} \times (r_{p,d_j} - \bar{r}_{d_j})}{\sum_{d_j \in NN} IR_{d_j}}; & if \quad TMCSim_{d_i,d_j} = 0 \end{cases} \tag{8}$$

Where $\bar{r}_{d_j}$ and $\bar{r}_{d_j}$ denote the mean ratings of items d_i and d_j, correspondingly. $TMCSim_{d_i,d_j}$ is the MC item-based CF similarity between items d_i and d_j, and NN is the set of CF-based nearest neighbors for item d_i.

3.2. The item-based content module

Existing patients who have had past consultations with doctors in a particular specialty but want to switch doctors would benefit from the use of the proposed system in finding appropriate doctors in the same specialty thru learning about the preferences of other patients who have visited the same doctors. For instance, a patient with a conflicting schedule with his current pediatrician may benefit from the use of the proposed system in knowing about other pediatricians who are similar to the current pediatrician by utilizing the preferences of other patients who have visited his current pediatrician. Accordingly, the item-based content module takes into account the doctor's specialty as one of the most important attributes for doctors in addition to the patients' ratings, which have been utilized in the previous module, in order to enhance the quality of personalized doctor recommendations.

In this regard, we assume that all doctors are assigned to specified specialty categories. When two doctors have the same specialty, they are assumed to be similar to each other based on the specialty's categories. As a result, the item-based content similarity compares the category representations of doctors rated by an active patient to recommend new doctors who have not visited before. For example, based on Table 1, assume that a patient has already visited and rated doctor *D1*, who is a Dermatologist. Later on, if the patient decided to switch to another doctor in the same specialty, the proposed system will help in finding appropriate doctors by utilizing both the specialty, in this case, *D3* or *D6*, and the preferences of other patients who have visited and rated doctor *D1*.

For this purpose, each doctor is initially encoded as a vector of binary numbers [0,1] representing the specialty as shown in (9).

$$\vec{V}_{d_i} = (v_{d_i,1}, v_{d_i,2}, \ldots\ldots, v_{d_i,S}), where$$
$$v_{d_i} = \begin{cases} 1 & If\ Doctor\ d_i\ belongs\ to\ specialty\ S \\ 0 & Otherwise \end{cases} \tag{9}$$

Then, the Driver and Kroeber metric [25] is utilized to express the item-based content similarity among items.

$$CSim_{d_i,d_j} = \frac{A}{\sqrt{(A+B)*(A+C)}}, where$$
$$\begin{cases} A = Frequency\ of\ events\ where\ v_{d_i,s}\ is\ 1\ and\ v_{d_j,s}\ is\ 1 \\ B = Frequency\ of\ events\ where\ v_{d_i,s}\ is\ 0\ and\ v_{d_j,s}\ is\ 1 \\ C = Frequency\ of\ events\ where\ v_{d_i,s}\ is\ 1\ and\ v_{d_j,s}\ is\ 0 \end{cases} \tag{10}$$

For predictions, the mean-based prediction metric used to produce content-based predicted ratings is as (11):

$$Pred_{p,d_i}^{Content} = \bar{r}_{d_i} + \frac{\sum_{d_j \in NN} CSim_{d_i,d_j} \times (r_{p,d_j} - \bar{r}_{d_j})}{\sum_{d_j \in NN} CSim_{d_i,d_j}} \tag{11}$$

where $CSim_{d_i,d_j}$ is the item-based content similarity between items d_i and d_j, and NN is the set of content-based nearest neighbors for item d_i.

Table 1. Example of content-based filtering

ID	Specialty
D1	Dermatologist
D2	Family, Pediatrician
D3	Dermatologist
D4	Gynecologist
D5	Pediatrician
D6	Dermatologist

3.3. The hybrid prediction module

The switching hybridization strategy is employed in this module to obtain the final predicted rating depending on certain conditions, as shown by (12). A weighted harmonic mean aggregation method is used to aggregate the predicted values. This method makes sure that a high total predicted value is only reached if both the multi-criteria (MC) Item-based collaborative filtering (CF) and the item-based content approaches produce high predicted values.

$$Pred_{p,d_i}^{Final} = \begin{cases} 0 & ;if \ Pred_{p,d_i}^{MCCF} = 0 \ and \ Pred_{p,d_i}^{Content} = 0 \\ Pred_{p,d_i}^{MCCF} & ;if \ Pred_{p,d_i}^{MCCF} \neq 0 \ and \ Pred_{p,d_i}^{Content} = 0 \\ Pred_{p,d_i}^{Content} & ;if \ Pred_{p,d_i}^{MCCF} = 0 \ and \ Pred_{p,d_i}^{Content} \neq 0 \\ \frac{2 \times Pred_{p,d_i}^{MCCF} \times Pred_{p,d_i}^{Content}}{Pred_{p,d_i}^{MCCF} + Pred_{p,d_i}^{Content}} & ;if \ Pred_{p,d_i}^{MCCF} \neq 0 \ and \ Pred_{p,d_i}^{Content} \neq 0 \end{cases} \quad (12)$$

4. EXPERIMENTAL RESULTS AND DISCUSSION
4.1. Dataset and evaluation indexes

The RateMDs MC dataset is used for the experimental validation. It is gathered from the *ratemds.com* website, which provides a platform for patients to review doctors on a rating scale from 1 to 5 on four criteria: staff, punctuality, helpfulness, and knowledge. The dataset includes 31,180 multi-criteria ratings of 3,464 patients on 3,118 doctors. The doctors in the dataset have 21 specialties, including pediatricians, dermatologists, family, gynecologists, and physiatrists. The level of the sparsity of the RateMDs dataset is 99.7%.

Three indexes are selected to evaluate the recommendation quality of the proposed approach, including the mean absolute error (MAE), the root mean square error (RMSE), and the prediction coverage. MAE and RMSE are the most frequently used indexes for evaluating the accuracy of recommendation techniques and are computed by comparing the predicted ratings against actual ratings. Note that lower values of MAE and RMSE indicate a higher performance in terms of prediction accuracy. The prediction coverage is the proportion of items for which a recommendation approach can generate a predicted rating [26].

4.2. Comparison algorithms

For the purpose of demonstrating the effectiveness of the proposed approach, three item-based CF benchmark algorithms have been chosen. These algorithms include two standard item-based CF algorithms, namely the single-criteria item-based CF (SC-ICF) [22] and the multi-criteria item-based CF (MC-ICF) [21]. A third state-of-the-art item-based CF algorithm known as the multi-criteria semantic-enhanced CF recommendation algorithm (MC-SeCF) [16] is also included.

4.3. Comparison results

A set of tests were carried out to realize how effective the proposed approach is against the benchmark algorithms mentioned above. First, the proposed approach is compared against the comparison algorithms in terms of predictive accuracy on the RateMDs MC dataset. Then, the proposed approach is compared against the comparison algorithms in terms of predictive accuracy and coverage at different levels of sparsity.

4.3.1. Evaluation on the RateMDs MC dataset

The experimental results are demonstrated in Figures 2 and 3. As shown in Figure 2, the proposed approach has the lowest MAE values in the RateMDs dataset. The proposed approach obtains 89%, 89%, and 85% relative improvements in terms of MAE compared with the SC-ICF, MC-ICF, and MC-SeCF algorithms. Likewise, Figure 3 shows that the proposed approach has the minimum RMSE values in the RateMDs dataset. The proposed approach attains 78%, 78%, and 74% relative improvements in relation to RMSE compared with the SC-ICF, MC-ICF, and MC-SeCF algorithms. Taking into account the extreme level of sparsity of the RateMDs dataset (99.7%), the results confirm the effectiveness and robustness of the proposed approach in comparison to the other algorithms in terms of prediction accuracy.
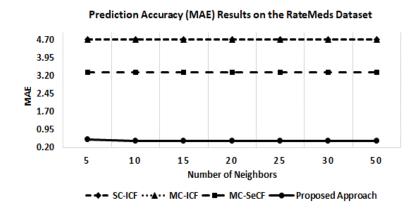

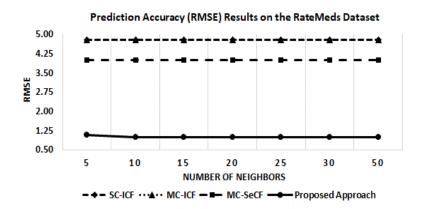Figure 2. Comparison of MAE values under different numbers of neighbors

Figure 3. Comparison of RMSE values under different numbers of neighbors

4.3.2. Evaluation on datasets with varied sparsity levels

Another series of experiments were carried out on several datasets with diverse sparsity levels. Figures 4 and 5 demonstrate the experimental results of the proposed approach and comparison algorithms. As depicted in Figure 4, the proposed approach has the lowest MAE values on all sparse datasets. The proposed approach obtains 67%, 61%, and 31% relative improvements in terms of MAE compared with the SC-ICF, MC-ICF, and MC-SeCF algorithms. In terms of prediction coverage, Figure 5 illustrates that the proposed approach has the maximum prediction coverage percentages on all sparse datasets. The proposed approach attains 57%, 45%, and 14% relative improvements corresponding to prediction coverage compared with the SC-ICF, MC-ICF, and MC-SeCF algorithms.

Once more, when dealing with highly sparse datasets, the proposed approach is remarkably robust. Consequently, the proposed approach greatly reduces the impact of the data sparsity problem in relation to prediction accuracy and coverage. This is due to the utilization of the enhanced item-based similarity metric, item reputation score, and content information of items to increase the quality of recommendations and reduce the influence of data sparsity when adequate rating data is unavailable.
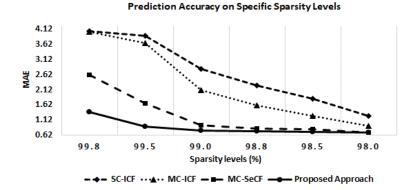
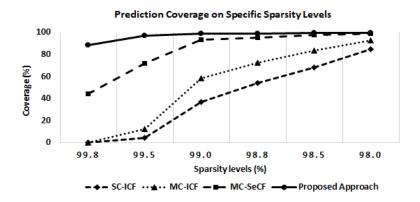Figure 4. Comparison of MAE values under different levels of sparsity

Figure 5. Comparison of coverage percentages under different levels of sparsity

5. CONCLUSION

This research presents a hybrid content-based MC-CF approach that assists patients in selecting the best doctors in accordance with their preferences. The proposed approach incorporates multi-criteria decision making, doctor reputation score, and content information about doctors in order to increase the quality of recommendations and lessen the influence of data sparsity when adequate rating data is unavailable. The experimental results on a real healthcare MC ratings dataset show that the proposed approach can provide highly reliable recommendations in highly sparse data, concerning predictive accuracy and coverage, when compared with other baseline item-based CF-based recommendation algorithms. In future work, we will consider incorporating sentiment analysis of doctor reviews in the process of recommendation to further advance the effectiveness of the proposed approach.

REFERENCES

[1] Y. Zhang, M. Chen, D. Huang, D. Wu, and Y. Li, "iDoctor: personalized and professionalized medical recommendations based on hybrid matrix factorization," *Future Generation Computer Systems*, vol. 66, pp. 30–35, Jan. 2017, doi: 10.1016/j.future.2015.12.001.

[2] Q. Han, I. M. D. R. De Troya, M. Ji, M. Gaur, and L. Zejnilovic, "A collaborative filtering recommender system in primary care: towards a trusting patient-doctor relationship," in *IEEE International Conference on Healthcare Informatics*, Jun. 2018, pp. 377–379, doi: 10.1109/ICHI.2018.00062.

[3] M. Waqar, N. Majeed, H. Dawood, A. Daud, and N. R. Aljohani, "An adaptive doctor-recommender system," *Behaviour and Information Technology*, vol. 38, no. 9, pp. 959–973, Jun. 2019, doi: 10.1080/0144929X.2019.1625441.

[4] M. Sridevi and R. R. Rao, "Finding right doctors and hospitals: a personalized health recommender," in *Lecture Notes in Networks and Systems*, vol. 40, Springer Singapore, 2019, pp. 709–719.

[5] Q. Meng and H. Xiong, "A doctor recommendation based on graph computing and lda topic model," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, pp. 808–817, 2021, doi: 10.2991/ijcis.d.210205.002.

[6] Y. Yang, J. Hu, Y. Liu, and X. Chen, "Doctor recommendation based on an intuitionistic normal cloud model considering patient preferences," *Cognitive Computation*, vol. 12, no. 2, pp. 460–478, Mar. 2020, doi: 10.1007/s12559-018-9616-3.

[7] H. Iftikhar, S. M. Anwar, and M. Majid, "A doctor recommendation system using patient's satisfaction analysis," in *Communications in Computer and Information Science*, vol. 1198, Springer Singapore, 2020, pp. 201–209.

[8] L. Guo, B. Jin, C. Yao, H. Yang, D. Huang, and F. Wang, "Which doctor to trust: a recommender system for identifying the right doctors," *Journal of Medical Internet Research*, vol. 18, no. 7, Jul. 2016, doi: 10.2196/jmir.6015.

[9]   Y. F. Huang, P. Liu, Q. Pan, and J. S. Lin, "A doctor recommendation algorithm based on doctor performances and patient preferences," in *International Conference on Wavelet Active Media Technology and Information Processing*, Dec. 2012, pp. 92–95, doi: 10.1109/ICWAMTIP.2012.6413447.

[10]   J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: a survey," *Decision Support Systems*, vol. 74, pp. 12–32, Jun. 2015, doi: 10.1016/j.dss.2015.03.008.

[11]   Q. Shambour, S. Fraihat, and M. Hourani, "The implementation of mobile technologies in higher education: a mobile application for university course advising," *Journal of Internet Technology*, vol. 19, no. 5, pp. 1327–1337, 2018, doi: 10.3966/160792642018091905005.

[12]   Q. Y. Shambour, A. H. Hussein, Q. M. Kharma, and M. M. Abualhaj, "Effective hybrid content-based collaborative filtering approach for requirements engineering," *Computer Systems Science and Engineering*, vol. 40, no. 1, pp. 113–125, 2022, doi: 10.32604/csse.2022.017221.

[13]   M. J. Pazzani and D. Billsus, "Content-based recommendation systems," in *The Adaptive Web*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 325–341.

[14]   C. C. Aggarwal, "Neighborhood-based collaborative filtering," in *Recommender Systems*, 2016, pp. 29–70.

[15]   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, Jun. 2005, doi: 10.1109/TKDE.2005.99.

[16]   Q. Shambour, M. Hourani, and S. Fraihat, "An item-based multi-criteria collaborative filtering algorithm for personalized recommender systems," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 8, 2016, doi: 10.14569/IJACSA.2016.070837.

[17]   Q. Shambour, "A user-based multi-criteria recommendation approach for personalized recommendations," *International Journal of Computer Science and Information Security*, vol. 14, no. 12, pp. 657–663, 2016.

[18]   F. Narducci, C. Musto, M. Polignano, M. de Gemmis, P. Lops, and G. Semeraro, "A recommender system for connecting patients to the right doctors in the healthnet social network," in *Proceedings of the 24th International Conference on World Wide Web*, May 2015, pp. 81–82, doi: 10.1145/2740908.2742748.

[19]   Q. Han, M. Ji, I. M. D. R. De Troya, M. Gaur, and L. Zejnilovic, "A hybrid recommender system for patient-doctor matchmaking in primary care," in *IEEE 5th International Conference on Data Science and Advanced Analytics, DSAA 2018*, Oct. 2019, pp. 481–490, doi: 10.1109/DSAA.2018.00062.

[20]   B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi, "A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data," *Knowledge-Based Systems*, vol. 82, pp. 163–177, 2015, doi: 10.1016/j.knosys.2015.03.001.

[21]   G. Adomavicius and Y. O. Kwon, "New recommendation techniques for multicriteria rating systems," *IEEE Intelligent Systems*, vol. 22, no. 3, pp. 48–55, May 2007, doi: 10.1109/MIS.2007.58.

[22]   B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the tenth international conference on World Wide Web - WWW '01*, 2001, pp. 285–295, doi: 10.1145/371920.372071.

[23]   Q. Shambour and J. Lu, "A hybrid multi-criteria semantic-enhanced collaborative filtering approach for personalized recommendations," in *IEEE/WIC/ACM International Conference on Web Intelligence*, Aug. 2011, vol. 1, pp. 71–78, doi: 10.1109/WI-IAT.2011.109.

[24]   J. Herlocker, J. A. Konstan, and J. Riedl, "An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms," *Information Retrieval*, vol. 5, no. 4, pp. 287–310, 2002, doi: 10.1023/A:1020443909834.

[25]   R. Todeschini, V. Consonni, H. Xiang, J. Holliday, M. Buscema, and P. Willett, "Similarity coefficients for binary chemoinformatics data: Overview and extended comparison using simulated and real data sets," *Journal of Chemical Information and Modeling*, vol. 52, no. 11, pp. 2884–2901, Nov. 2012, doi: 10.1021/ci300261r.

[26]   C. C. Aggarwal, "Evaluating recommender systems," in *Recommender Systems*, 2016, pp. 225–254.

BIOGRAPHIES OF AUTHORS

Qusai Y. Shambour 🔵 🟦 🆂🅲 ⟳ holds a B.Sc. degree in Computer Science from Yarmouk University, Jordan, in 2001, an M.Sc. degree in Computer Networks from the University of Western Sydney, Australia, in 2003, and the Ph.D. degree in Software Engineering from the University of Technology Sydney, Australia, in 2012. Currently, he is a Professor at the Department of Software Engineering, Al-Ahliyya Amman University, Jordan. His research interests include information filtering, recommender systems, VoIP, machine learning, and data science. He can be contacted at email: q.shambour@ammanu.edu.jo.

Mahran M. Al-Zyoud 🔵 🟦 🆂🅲 ⟳ is an Assistant Professor of Networks and Cybersecurity at Al-Ahliyya Amman University. He received a B.Sc. in Computer Science and an M.Sc. in Computer Information Systems from The University of Jordan in 2004 and 2012. He received a Ph.D. in Computer Science from The University of Alabama, USA, in 2019. His research interests include data privacy and IoT security. He can be reached by email at: m.zyoud@ammanu.edu.jo.

Abdelrahman H. Hussein ⓘ ⑧ ⑤ ↻ is an associate professor at the Department of Computer Science at Al-Ahliyya Amman University, Jordan. He received his first degree in Computer Science from Jordan University of Science and Technology, Jordan, in July 2000, a master's degree in Computer Science from Jordan University, Jordan in July 2003, and Ph.D. from the Anglia Ruskin University (ARU), UK in 2010. He is currently chairman of the Department of Computer Science, Al-Ahliyya Amman University in Jordan, Amman. His main research interests lie in the areas of VoIP, Mobile Ad-Hoc Networking, and E-Learning. He can be contacted at email: a.husein@ammanu.edu.jo.

Qasem M. Kharma ⓘ ⑧ ⑤ ↻ received the Ph.D. in Computer Science from Florida International University in 2005. He is an associate Professor in Software Engineering Department at Al-Ahliyya Amman University. His research interests are information systems, software engineering, databases, and using technology in education. He worked as a director of the E-Learning Center and Chairman of the Computer Science Department. Currently, he is the Head of the Software Engineering Department. He can be contacted at email: q.kharma@ammanu.edu.jo.