

A Recommender System Based on Hesitant Fuzzy Linguistic Information with MAPPACC Approach

Zeshui XU^{1*}, Hongyu CHEN¹, Yue HE²

¹ Business School, Sichuan University, No. 24 South Section 1 Yihuan Road, Chengdu, 610064, China
xuzeshui@263.net (*Corresponding author), hy_chenhy@qq.com

² West China Second University Hospital, Sichuan University, No. 24 South Section 1 Yihuan Road, Chengdu, 610064, China
hey_heyue@163.com

Abstract: Recommender systems can make contributions to enterprises by meeting the demands of users and improving their satisfaction. However, because of the uncertainty and complexity of users' preferences, the classical techniques are insufficient to sort out the suitable recommendations. Scholars have made progress to address uncertainty in recommender systems, but the existing studies neglected the uncertain linguistic information and failed to use them to efficiently provide personalized recommendations for individuals. Therefore, this paper demonstrates an item-based recommendation program combined with the Hesitant Fuzzy Linguistic Multi-criteria Analysis of Preferences by means of Pairwise Actions and Criterion Comparisons (HFL-MAPPACC) method to analyze hesitant fuzzy linguistic information. To provide personalized recommendations, preference degrees of the used items and tendency of evaluations are considered in the construction of this algorithm. Then, the proposed approach is implemented in a doctor recommender system to show its applicability. Ultimately, the validity of the proposed method and its superiority are discussed by comparative analyses.

Keywords: Recommender systems, Collaborative filtering, Hesitant fuzzy linguistic set, Multi-criteria decision making, MAPPACC.

1. Introduction

As technology develops, the Internet has permeated many aspects of our daily life, and a large amount of information is exchanged consequently. The information exhibited to users is often selected by different rules and techniques. Recommender system, as an effective tool, can customize the information provided for users based on their personal characters and necessities. Since recommender system emerged in the 1990s (Adomavicius & Tuzhilin, 2005), recommender systems have been applied in a variety of domains to tackle the information overload problem and ameliorate customer relationship, such as e-commerce (Guia et al., 2019), e-tourism (Renjith et al., 2020). Apart from creating profits for enterprises, recommender system can even save lives and make social impact (Vercruyssen et al., 2015).

Recommender systems generally contain three main stages, namely gathering information, running the recommendation techniques, and outputting recommendation lists. Based on kinds of information gathered, the recommendation technique varies from each other, and it can determine the quality and efficiency of the recommender system. Referring to subject information, many recommender systems take into consideration only the numerical ratings, which are limited to express the users' opinions and match the new items. A full score may not

represent a perfect performance. Thus, some scholars introduced the method of fuzzification to depict the inner uncertainty of users. And it really improves the performances of recommender systems (Ajoudanian & Abadeh, 2019). However, the rules of transferring ratings into fuzzy numbers are often fixed, failing to express the different extents of uncertainty within individuals. Hence, the linguistic evaluations have been noticed in the area of recommender systems. The study of uncertain linguistic information can reveal the imprecise attitudes of users related to different aspects of products and has been proven to be efficient in a lot of areas (Xu et al., 2019; Lin et al., 2018). By analyzing the linguistic assessment from users with fuzzy approaches, recommender systems can express the uncertainties and derive the personalized recommendations. Moreover, because the uncertain linguistic information reveals details about individual features and preferences, it can improve some drawbacks in the existing recommendation techniques, such as the sparsity problem (Shambour et al., 2016).

To handle the uncertain linguistic information, the hesitant fuzzy linguistic term sets (HFLTSS) and some extensions are used. Accordingly, recommender systems have been improved from some points of view. The distance and similarity measures of HFLTSS (Liao et al., 2014) are proposed and executed to provide

movie recommendations. Then, based on the same linguistic environment, Farhadinia (2016) designed the entropy measures of HFLTSS and applied them to derive the weights of criteria for alternatives. Similarly, the distance measure (Chen & Liu, 2016) and the entropy measure with the VIKOR method (Chen & Liu, 2018) are applied under the multi-granular HFLTSS environment to provide recommendations. These studies have made foundations for computation and further exploration of uncertain linguistic information in recommender systems. But the evaluations are given by experts or professional organizations. These works do not serve for personalized recommendations, and their results are exhibited on the home pages of websites. But in recommender systems, personalisation is indispensable which can increase users' visits of these websites (Kompan & Bielikova, 2013). This paper intends to deal with the uncertain linguistic evaluations from users in recommender systems in order to provide personalized recommendations for individuals. Because of generality, the HFLTSS are used to describe the uncertain linguistic information.

In addition to the type of information, recommendation techniques also exert a considerable influence over the whole system. One commonly used technique is collaborative filtering (CF) approaches (Ajoudanian & Abadeh, 2019). The classical CF recommender system seeks for similar users or items through ratings. Because linguistic evaluations can be extracted as linguistic ratings, it is reasonable to apply CF techniques to analyze the linguistic information. Deshpande and Karypis (2004) designed the item-based top-N scheme to shorten the time of building and implementing the model. The scheme performs better with sparsity data and run faster than user-based CF approach. However, this algorithm has two defects: 1) it makes a loss of the original information; 2) it overlooks the preferences of users for rated items. But in fact, items which are similar to the preferred items have a bigger possibility to be liked than the others. Scholars notice the second problem and make advances with ratings (Wu et al., 2010). However, when it comes to the uncertain linguistic information, the preferences of users are quite obscure and vague. Therefore, this paper applies the MAPPACC method to deal with HFLTSS information and express the preference degrees of the used items

based on multiple criteria. The MAPPACC approach (Matarazzo, 1986) is designed to solve multi-criteria decision-making problem. Because recommender system belongs to a general subclass of Decision Support System (DSS) (Filip, 2020), it is valid to merge the MAPPACC method into the procedure. Moreover, the result of the MAPPACC method contains the rankings with specific preferences, which can facilitate the advancement of the traditional top-N scheme.

The purpose of this paper is to develop an advanced item-based recommendation algorithm for handling hesitant fuzzy linguistic information. Aiming to reflect the customer preferences to their purchased items and to different criteria, the item-based top-N approach is combined with the HFL-MAPPACC method so as to establish the whole program. This study brings novel aspects such as the following:

1. The application of MAPPACC method under HFLTSS environment and recommender systems is original and valid. The HFL-MAPPACC method helps the expression of preferences and guarantees the quality of recommendations. Its efficiency and advantages have been proven through the comparative analysis.
2. By combining the preference degrees of the used items and the similarities among items, the paper provides recommendations that are more appealing to individuals. Meanwhile, the traditional top-N method is advanced to avoid the loss of information.
3. Users often have their evaluation tendency. Some users tend to give good ratings, while others often rate at a low level. Therefore, this paper proposes the improved cosine similarity measure to eliminate the impacts of evaluation tendency.

The rest of this paper is structured as follows: Section 2 introduces the top-N algorithm, basic knowledge of HFLTSS and the traditional MAPPACC method. Section 3 presents the whole process of recommender program, which contains the HFL-MAPPACC approach. In Section 4, the proposed model is demonstrated in a numerical example. Then, to illustrate the effectiveness of the proposed method, comparative analysis is developed in Section 5. Ultimately, the conclusions and future explorations are discussed in Section 6.

2. Preliminaries

2.1 The Item-based Top-N Method in Recommender Systems

As a model-based recommendation algorithm, the top-N method can be completed quickly by precomputed model (Deshpande & Karypis, 2004). The elements in the user-item matrix are binary. If the user has purchased the item, the element is equal to one, otherwise it equals to zero. The latent meaning in the purchased item vector U of the active user is also the same. The parameter k means to sort out the most similar items for each item which can lessen the running time. Through experiments, k ($10 \leq k \leq 30$) can lead to good results and higher values provide little or no improvements. Ultimately, the algorithm recommends N best items. The process is shown in Table 1.

Table 1. The item-based top-N method

INPUT: the $n \times m$ user-item matrix R ; the $m \times 1$ vector U ; parameters k and N
OUTPUT: the $m \times m$ neighbor matrix M ; the $m \times 1$ recommendation strength vector x
Step 1. For each pair of items i and j , compute similarity measure $sim(R_{*,j}, R_{*,i})$ and let $M_{i,j} = sim(R_{*,j}, R_{*,i})$.
Step 2. For each element in M ,
a) If $i = j$, $M_{i,j} = 0$;
b) If $M_{i,j}$ is not among the k largest similarities in $M_{*,j}$, $M_{i,j} = 0$.
Step 3. Do $M \times U$ and name the result as x .
Step 4. For each element in x ,
c) If $U_i \neq 0$, $x_i = 0$;
d) If x_i is not among the N largest recommendation strengths in x , $x_i = 0$.
End.

In Step 2, the algorithm selects k neighbors of each item, and change non- k similarities into zero, which leads to a loss of information. Besides, binary information only reflects the historical behavior without feedbacks. For this reason, ratings are introduced to express the attitudes of users. To accurately convey the experiences and preferences of users, this paper addresses the HFLTSs. Then, the tradition top-N method is improved by reducing information loss and combining the HFL-MAPPACC method.

2.2 The hesitant fuzzy linguistic term sets

When making assessments, decision makers may hesitate among several linguistic terms, such as “between high and very high” or “lower than good”. To flexibly describe such situations, the concept of HFLTS was introduced by Rodriguez et al. (2012).

Definition 1 (Rodriguez et al., 2012).

$S_1 = \{s_\alpha | \alpha = 0, 1, \dots, \tau\}$ is set to be a linguistic term set. A hesitant fuzzy linguistic term set (HFLTS), H_S , is an ordered subset of consecutive linguistic terms of S .

Referring to the subscript of linguistic term set (LTS), Xu (2005) proposed a symmetric scale: $S_2 = \{s_\alpha | \alpha = -\tau, \dots, -1, 0, 1, \dots, \tau\}$. The similar attitude such as “indifference” or “fair” is denoted as s_0 . The positive and negative subscripts directly show the attitudes of decision makers. Hence, the symmetric way is applied.

The LTS, S , should satisfy the characteristics:

1. if and only if $\alpha > \beta$, $s_\alpha > s_\beta$;
2. there is the negation operator: $neg(s_\alpha) = s_{-\alpha}$.

For $s_\alpha, s_\beta \in S$, some operational laws for LTS are introduced (Xu, 2005):

3. $s_\alpha \oplus s_\beta = s_{\alpha+\beta}$;
4. $\mu s_\alpha = s_{\mu\alpha}$;
5. $\mu(s_\alpha \oplus s_\beta) = \mu s_\alpha + \mu s_\beta$.

The operation results should be within the interval $s_{-\tau}, s_\tau$ (Liao et al., 2014). For example, if $s_{\alpha+\beta} > s_\tau$, then the result of $s_\alpha \oplus s_\beta$ should be s_τ .

Definition 2 (Zhang & Wu, 2014). Let

H_S^i ($i = 1, 2, \dots, N$) be a collection of HFLTSs. The hesitant fuzzy linguistic weighted averaging (HFLWA) operator is

$$\begin{aligned} HFLWA(H_S^1, H_S^2, \dots, H_S^N) &= \bigoplus_{i=1}^N (\omega_i H_S^i) \\ &= \bigcup_{s_{\phi(l_1)} \in H_S^1, s_{\phi(l_2)} \in H_S^2, \dots, s_{\phi(l_n)} \in H_S^N} \left\{ s_{\sum_{i=1}^N \omega_i \phi(l_i)} \right\} \end{aligned} \quad (1)$$

where $\omega = (\omega_1, \omega_2, \dots, \omega_N)^T$ is the weight vector of H_S^i ($i = 1, 2, \dots, N$) with $\omega_i \in [0, 1]$ and $\sum_{i=1}^N \omega_i = 1$. Moreover, if $\omega_i = 1/N$, then the HFLWA operator is reduced to the hesitant fuzzy linguistic averaging (HFLA) operator.

2.3 The Classical MAPPACC Structure

The MAPPACC method is based on a pairwise comparison of alternatives relative to each pair of criteria. To constitute a complete preorder, the relations of preference and indifference (P&I) are introduced. Moreover, by aggregating these preferences, it derives preference indices of alternatives and get their final ranking. Its steps are briefly described as follows:

1. *Comparison*: the first step is to normalize the matrix and get the weights of criteria. Then, the evaluations of every pair of alternatives should be compared under each pair of criteria so as to get the pairwise relations of preference and indifference (P&I).
2. *Calculation*: based on the comparison results, the preference indices can be derived so that the preference matrices related to each pair of criteria can be obtained.
3. *Aggregation*: aggregate the preference matrices to get the global matrix.
4. *Ranking*: by dealing with the global matrix, the ranking results can be determined.

3. The Proposed Model with HFL-MAPPACC

3.1 The Data Processing for HFLTSSs

The normalization of HFLTSSs and the model to derive the weights of the criteria are presented. Different items b_i ($i=1, 2, \dots, n; b_i \in B$) are evaluated by users under certain criteria c_j ($j=1, 2, \dots, m; c_j \in C$). The evaluations are extracted based on the following scale for the LTS ($S = \{s_\alpha | \alpha = -\tau, \dots, -1, 0, 1, \dots, \tau\}$). Then, the original decision matrix would be formed with HFLTSSs. To explicitly express the process, the general form of evaluation matrix is given. The original evaluation matrix of the user A is constituted as follows:

$$(H_s^j)_{n \times m} = \begin{bmatrix} H_s^{11} & H_s^{12} & \dots & H_s^{1m} \\ H_s^{21} & H_s^{22} & \dots & H_s^{2m} \\ \vdots & \vdots & \ddots & \vdots \\ H_s^{n1} & H_s^{n2} & \dots & H_s^{nm} \end{bmatrix} \quad (2)$$

Each row in the matrix represents the evaluation for one item or alternative. Because the pairwise comparison and computation will be implemented,

$b_r, b_s \in B$ are used to represent two different items, where $r=1, 2, \dots, n; s=1, 2, \dots, n$. Similarly, $c_h, c_k \in C; h \neq k; h=1, 2, \dots, m; k=1, 2, \dots, m$.

3.1.1 The Normalization of HFLTSSs

According to the classical MAPPACC method, it is necessary to standardize the elements in the original evaluation matrix. There is the condition that HFLTSSs have different numbers of linguistic terms. To facilitate the operations among HFLTSSs, the shorter HFLTSS is often extended by adding some terms so that they all have the same length. To minimize the influence on the original data and the following operations, the score function result of the shorter H_s is utilized as the element added to itself.

Definition 3. For a HFLTSS $H_s = \{s_{\phi(l)} | s_{\phi(l)} \in S; l=1, 2, \dots, \#H_s\}$. The score function is obtained by Liao et al. (2015):

$$\rho(H_s) = \frac{1}{\#H_s} \sum_{i=1}^{\#H_s} s_{\phi(i)} = s_{\bar{\phi}(l)} = s_{\frac{1}{\#H_s} \sum_{i=1}^{\#H_s} \phi(i)} \quad (3)$$

After the extension, the number of elements in a HFLTSS should be updated, and $\#L$ is used to denote the final length of this HFLTSS. According to the statistical knowledge, neither the score of HFLTSS nor the variance function is affected. The variance function can be obtained as follows:

$$\sigma(H_s) = \frac{1}{\#L} \sum_{l=1}^{\#L} \sqrt{(s_{\phi(l)} - s_{\bar{\phi}(l)})^2} \quad (4)$$

After the extension, these linguistic terms can be arranged in a descending order or in an ascending order by comparing the subscripts of every element in the HFLTSS. Hence, the normalization step is finished and the normalized HFLTSSs (NHFLTSSs) are achieved.

3.1.2 The Model for Deriving the Weights of the Criteria

Considering that the weights of criteria would play a huge influence on the final rankings, it is essential to guarantee its rationality. Because of the limitation of knowledge, sometimes it is irrational for decision makers to determine the whole weights of criteria. Maximizing deviation method (Wang, 1997) is an effective method to reduce subjective impacts. According to this method, if the performances of alternatives under some criteria are quite different than under other

criteria, which can be shown through larger deviation degrees or distances, these criteria play important roles in the decision-making process. So, they should be paid more attention and the weights assigned to them should be relatively larger. The distance measure can be applied to derive the criteria weights.

Hamming distance is often used to measure the distance between two HFLTSSs. With NHFLTSSs, the distance measure can be easily processed. Liao et al. (2014) defined the Hamming distance measure as follows:

$$d(H_S^1, H_S^2) = \frac{1}{\#L} \sum_{l=1}^{\#L} \frac{|\phi(l)_1 - \phi(l)_2|}{2\tau + 1} \quad (5)$$

where $\phi(l)_1$ and $\phi(l)_2$ denote the subscripts of the l -th element in H_S^1 and H_S^2 respectively. The smaller the distance measure is, the smaller the difference between the two NHFLTSSs. It is obvious that $d(H_S^1, H_S^2) = d(H_S^2, H_S^1)$.

Hence, the distance between the item b_r with the other items b_s related to the criterion c_j is:

$$d(c_j(b_r)) = \sum_{s=1}^n d(H_S^{rj}, H_S^{sj}), \quad r=1, 2, \dots, n; j=1, 2, \dots, m \quad (6)$$

where H_S^{rj} and H_S^{sj} are the evaluations of the items b_r and b_s under the criterion c_j respectively. Let $d(c_j)$ express the total difference which is relative to c_j .

$$d(c_j) = \sum_{r=1}^n d(c_j(a_r)) = \sum_{r=1}^n \sum_{s=1}^n d(H_S^{rj}, H_S^{sj}), \quad j=1, 2, \dots, m \quad (7)$$

As it has been mentioned, the greater the difference is, the more important the criterion should be. Then, the model can be set up:

Model 1

$$\begin{aligned} \max f(\varpi) &= \sum_{j=1}^m w_j d(c_j) = \sum_{j=1}^m \sum_{r=1}^n \sum_{s=1}^n \varpi_j d(H_S^{rj}, H_S^{sj}) \\ \text{s.t. } \sum_{j=1}^m \varpi_j^2 &= 1, \varpi_j \geq 0, j=1, 2, \dots, m \end{aligned} \quad (8)$$

The following Lagrange function is created for solving the model above:

$$\text{Lag}(\varpi, \xi) = \sum_{j=1}^m \sum_{r=1}^n \sum_{s=1}^n \varpi_j d(H_S^{rj}, H_S^{sj}) + \frac{\xi}{2} (\sum_{j=1}^m \varpi_j^2 - 1) \quad (9)$$

After solving the Lagrange function, the weight vector is obtained:

$$\varpi_j = \frac{\sum_{r=1}^n \sum_{s=1}^n d(H_S^{rj}, H_S^{sj})}{\sqrt{\sum_{j=1}^m [\sum_{r=1}^n \sum_{s=1}^n d(H_S^{rj}, H_S^{sj})]^2}} \quad j=1, 2, \dots, m \quad (10)$$

Besides, ϖ_j also needs to be normalized so as to adapt to the traditional usage. So, let

$$w_j = \varpi_j / \sum_{j=1}^m \varpi_j \quad (11)$$

From the above equation, it results

$$w_j = \frac{\sum_{r=1}^n \sum_{s=1}^n d(H_S^{rj}, H_S^{sj})}{\sum_{j=1}^m \sum_{r=1}^n \sum_{s=1}^n d(H_S^{rj}, H_S^{sj})} \quad j=1, 2, \dots, m \quad (12)$$

3.2 The HFL-MAPPACC Method

In this subsection, the comparison scheme should be set and carried out. After that, the pairwise relations of P&I can be obtained and the basic preference indices can be calculated.

Up to now, scholars have developed various methods to compare HFLTSSs. Rodriguez et al. (2012) used the envelopes of HFLTSSs to make comparisons. But, it is an interval-valued linguistic term (Xu, 2005). The interval form may lose the initial information in HFLTSSs. So, other scholars presented some novel methods to overcome the flaw (Wei et al., 2014; Feng et al., 2018). These methods have different basis and advantages. In this paper, the new score function proposed by Wei et al. (2015) is applied. It takes into consideration the average linguistic term and the hesitant degree of decision maker. The greater the mean of linguistic terms in a HFLTS is, the better the HFLTS should be. Besides, if the HFLTS has a higher degree of hesitation, which is shown by more terms of it, its reliance would decline and the score would decrease.

Definition 4 (Wei et al., 2015). To distinguish the new score function from the one mentioned in Definition 2, its name is replaced with the comparison function.

Let $S = \{s_\alpha | \alpha = -\tau, \dots, -1, 0, 1, \dots, \tau\}$ be a LTS. For a HFLTS $H_S = \{s_{\phi(l)} | s_{\phi(l)} \in S; l=1, 2, \dots, \#L\}$, the comparison function is denoted as:

$$\zeta(H_S) = \bar{\phi}(l) - \frac{1}{\#L} \sum_{l=1}^{\#L} \frac{(\phi(l) - \bar{\phi}(l))^2}{\text{var}(\tau)} \quad (13)$$

where $\bar{\phi}(l) = \frac{1}{\#L} \sum_{l=1}^{\#L} \phi(l)$, and

$$\text{var}(\tau) = \frac{(-\tau)^2 + \dots + (-1)^2 + (1)^2 + \dots + (\tau)^2}{2\tau + 1}$$

Hence, the following comparison can be executed:

1. If $\zeta(H_S^1) > \zeta(H_S^2)$, $H_S^1 > H_S^2$;
2. If $\zeta(H_S^1) = \zeta(H_S^2)$, $H_S^1 = H_S^2$.

Then, the evaluations of a pair of items (b_r, b_s) ($b_r, b_s \in B, b_r \neq b_s$) can be compared under the pair of criteria (c_h, c_k) ($c_h, c_k \in C, c_h \neq c_k$), which are denoted as $H_S^{rh}, H_S^{rk}, H_S^{sh}$ and H_S^{sk} . Let $\pi_{hk}(b_r, b_s)$ be the basic preference index relative to the pair of alternatives (b_r, b_s) and to the criterion (c_h, c_k) , so

$$\begin{array}{c} \pi_u(b_r, b_s) = \\ \hline 1 \\ 0 \\ 0.5 \\ \frac{w_h d(H_S^{rh}, H_S^{rh})}{w_h d(H_S^{rh}, H_S^{rh}) + w_k d(H_S^{rh}, H_S^{rh})} \\ \frac{w_h d(H_S^{rh}, H_S^{rh})}{w_h d(H_S^{rh}, H_S^{rh}) + w_k d(H_S^{rh}, H_S^{rh})} \\ \hline \end{array} \quad \begin{array}{c} \pi_u(b_r, b_s) = \\ \hline 0 \\ 1 \\ 0.5 \\ \frac{w_k d(H_S^{rk}, H_S^{rk})}{w_k d(H_S^{rk}, H_S^{rk}) + w_h d(H_S^{rk}, H_S^{rk})} \\ \frac{w_k d(H_S^{rk}, H_S^{rk})}{w_k d(H_S^{rk}, H_S^{rk}) + w_h d(H_S^{rk}, H_S^{rk})} \\ \hline \end{array} \quad \begin{array}{l} \text{if } H_S^{rh} > H_S^{rk} \text{ and } H_S^{sh} > H_S^{sk} \\ \text{if } H_S^{rh} < H_S^{rk} \text{ and } H_S^{sh} < H_S^{sk} \\ \text{if } H_S^{rh} = H_S^{rk} \text{ and } H_S^{sh} = H_S^{sk} \\ \text{if } \begin{cases} H_S^{rh} > H_S^{rk} \text{ and } H_S^{sh} \leq H_S^{sk} \\ H_S^{rh} = H_S^{rk} \text{ and } H_S^{sh} < H_S^{sk} \end{cases} \\ \text{if } \begin{cases} H_S^{rh} \leq H_S^{rk} \text{ and } H_S^{sh} > H_S^{sk} \\ H_S^{rh} < H_S^{rk} \text{ and } H_S^{sh} = H_S^{sk} \end{cases} \end{array} \quad (14)$$

For $\forall b_r, b_s \in A (b_r \neq b_s)$ and $\forall c_h, c_k \in C (c_h \neq c_k)$, the following properties are true:

1. $\pi_{hk}(b_r, b_s) + \pi_{hk}(b_s, b_r) = 1$;
2. $0 \leq \pi_{hk}(b_r, b_s) \leq 1$.

Moreover, the relations (P_{hk}, I_{hk}) of two alternatives b_r, b_s ($b_r, b_s \in A, b_r \neq b_s$) can be obtained with respect to the two criteria c_h, c_k ($c_h, c_k \in C, c_h \neq c_k$).

$$\begin{aligned} b_r P_{hk} b_s &\Leftrightarrow 1 \geq \pi_{hk}(b_r, b_s) > 0.5 > \pi_{hk}(b_s, b_r) \geq 0 \\ b_r I_{hk} b_s &\Leftrightarrow \pi_{hk}(b_r, b_s) = \pi_{hk}(b_s, b_r) = 0.5 \end{aligned}$$

where $b_r P_{hk} b_s$ represents the preference relation that “ b_r is strictly preferred to b_s related to the criteria c_h and c_k ”, and $b_r I_{hk} b_s$ is the indifference relation that “ b_r is indifferent to b_s with reference to the criteria c_h and c_k ”. It should be noted that $\pi_{hk}(b_r, b_s) = \pi_{hk}(b_s, b_r) = 0.5$ is not equal to $H_S^{rh} = H_S^{sh}$ or $H_S^{rk} = H_S^{sk}$. If $w_h d(H_S^{rh}, H_S^{sh}) = w_k d(H_S^{rk}, H_S^{sk}) = 0.5$ happens, it should be taken into account “ b_r is indifferent to b_s with reference to the criteria c_h and c_k ”.

According to the basic preference indices, the $N \times N$ square matrix Π_{hk} can be obtained:

$$\Pi_{hk} = [\pi_{hk}(b_r, b_s)], \quad r = 1, 2, \dots, n; \quad s = 1, 2, \dots, n \quad (15)$$

By repeating the procedure above, all the $\frac{1}{2}M(M-1)$ matrices can be obtained, for each pair of criteria (c_h, c_k) ($c_h, c_k \in C, h < k$). Then, all the matrices can be integrated and the final $N \times N$ square matrix Π can be obtained:

$$\Pi = \sum_{hk} \Pi_{hk} \frac{w_h + w_k}{M-1} \quad (h < k) \quad (16)$$

In the global matrix, the element can be written as $\pi(b_r, b_s)$, which is the weighted sum of $\pi_{hk}(b_r, b_s)$. Besides, $\pi(b_r, b_s) + \pi(b_s, b_r) = 1$.

Ultimately, the global preference rankings can be calculated. From the matrix Π , the ranking indices $\delta(b_r)$ and $\bar{\delta}(b_r)$ can be obtained, which are also the preference degrees, $\forall b_r \in B$:

$$\begin{aligned} \delta(b_r) &= \sum_s \pi(b_r, b_s), \quad r \neq s, \quad r = 1, 2, \dots, n \\ \bar{\delta}(b_r) &= \sum_s \pi(b_s, b_r), \quad r \neq s, \quad r = 1, 2, \dots, n \end{aligned} \quad (17)$$

Then, the final rankings are obtained by comparing $\delta(b_r)$ or $\bar{\delta}(b_r)$. If $\delta(b_r) > \delta(b_s)$, then the active user prefers the item b_r to b_s . If $\delta(b_r) = \delta(b_s)$, then graph theories are valid to solve the situation (Matarazzo, 1986).

3.3 The Architecture of the Proposed Item-based CF Recommendation Algorithm

In this subsection, the classical item-based CF recommendation approach is modified and combined with the HFL-MAPPACC method.

Step 1. Establish the evaluation system and obtain the original decision matrix.

The evaluation database should be set up. These assessments are made under the criteria C_j ($j = 1, 2, \dots, M$). Besides, the users and items have to be numbered by the time of admission to the system. Therefore, the set for users is denoted as A_x ($x = 1, 2, \dots, X$) and the set for items is labelled as B_y ($y = 1, 2, \dots, Y$). Because of these principles, the original evaluation matrix for the entire database is:

$$E = (H_{x,y}^j)_{N \times M} \quad (18)$$

where the subscript S of H_S^{ij} is replaced by x, y for convenience. $H_{x,y}^j$ represents that the evaluations are made by the user A_x on the item B_y under the j -th criterion. M is the number of criteria and N is the total number of evaluations in the system. The set of items which the user A_x has used is named as $\Phi(A_x)$.

Example 1. Let us take the user A_1 's profile as example. A_1 has used the items B_1, B_3 and B_6 , so $\Phi(A_1) = \{B_1, B_3, B_6\}$ and the evaluation matrix is shown as follows:

$$A_1 = (H_{1,y}^j)_{N(A_1) \times M} = \begin{bmatrix} H_{1,1}^1 & H_{1,1}^2 & \dots & H_{1,1}^M \\ H_{1,3}^1 & H_{1,3}^2 & \dots & H_{1,3}^M \\ H_{1,6}^1 & H_{1,6}^2 & \dots & H_{1,6}^M \end{bmatrix} \quad (19)$$

Step 2. Normalize, aggregate and derive criteria weights.

With Equation (3), the added terms can be determined. Then, the HFLA operator is used to fuse the pieces of information together in user's profile under each criterion. The result is the average evaluation of every user under the criterion C_j , which can be denoted as $H_{\bar{x}}^j$.

Definition 5. Suppose that $H_{x,y}^j$ represents the evaluation that the user A_x has made on the item B_y under the criterion C_j . $H_{\bar{x}}^j$ can be obtained:

$$H_{\bar{x}}^j = \frac{1}{N(A_x)} \sum_{B_y \in \Phi(A_x)} H_{x,y}^j = \left\{ s \frac{1}{N(A_x)} \sum_{B_y \in \Phi(A_x)} \phi(l)_{x,y}^j \right\} \quad (20)$$

where $N(A_x)$ is the number of items in $\Phi(A_x)$. $\phi(l)_{x,y}^j$ is the subscript of the l -th linguistic terms in $H_{x,y}^j$. Then, Model 1 is used to get the weights for the users' matrices. The weights clearly express the sensitivity of the individuals to each criterion, denoted as w_x^j .

Step 3. Compute the similarity measure between items and gain the neighbors for every item.

In recommender systems, scholars often use the cosine similarity measure to effectively compute the similarity degree within different users or items. The cosine similarity measure for HFLTSs has been defined.

Definition 6 (Liao & Xu, 2015). Given two NHFLTSs $H_{x,r}^j$ and $H_{x,s}^j$, which are related to items B_r and B_s , in the evaluation matrix of A_x , the weighted cosine similarity measure between them under the criterion C_j with the weight w_x^j is

$$sim(B_r, B_s)_x = \frac{\sum_{j=1}^M w_x^j \sum_{l=1}^{\#L} \left(\left| \phi(l)_{x,r}^j \right| \cdot \left| \phi(l)_{x,s}^j \right| \right)}{\sum_{j=1}^M w_x^j \sqrt{\sum_{l=1}^{\#L} \phi^2(l)_{x,r}^j \cdot \sum_{l=1}^{\#L} \phi^2(l)_{x,s}^j}} \quad (21)$$

where $\phi(l)_{x,r}^j$ and $\phi(l)_{x,s}^j$ are the subscripts of the l -th term in $H_{x,r}^j$ and $H_{x,s}^j$.

However, every user has his/her own evaluation tendency. For example, the user A_1 is strict to products and the average linguistic evaluation is approximately s_{-1} . But the user A_2 tends to give evaluations around s_2 and s_3 , whose average linguistic evaluation is about s_2 . If they both assess a product as s_1 , then the latent meaning is different. s_1 indicates that this product seems to be a nice one for the user A_1 , while it is barely satisfactory from the prospect of the user A_2 .

Thus, the similarity measure can be improved by concerning the evaluation tendency, which can be exhibited through the average evaluation.

Definition 7. $H_{x,r}^j$ and $H_{x,s}^j$ are the NHFLTSs for the items B_r and B_s under the criterion C_j given by A_x . $H_{\bar{x}}^j$ denotes the average evaluations of the user A_x under the criterion C_j . So, the improved cosine similarity measure between B_r and B_s from the prospect of A_x is

$$sim^*(B_r, B_s)_x = \frac{\sum_{j=1}^M w_x^j \sum_{l=1}^{\#L} \left[\left(\phi(l)_{x,r}^j - \phi(l)_{\bar{x}}^j \right) \cdot \left(\phi(l)_{x,s}^j - \phi(l)_{\bar{x}}^j \right) \right]}{\sum_{j=1}^M w_x^j \sqrt{\sum_{l=1}^{\#L} \left(\phi(l)_{x,r}^j - \phi(l)_{\bar{x}}^j \right)^2 \cdot \sum_{l=1}^{\#L} \left(\phi(l)_{x,s}^j - \phi(l)_{\bar{x}}^j \right)^2}} \quad (22)$$

where $\phi(l)_{\bar{x}}^j$ is the subscript of the l -th term in $H_{x,r}^j$ and $H_{x,s}^j$. And it is clear that $sim^*(B_r, B_s)_x = sim^*(B_s, B_r)_x$. $sim^*(B_r, B_s)_x$ falls between $(-1, 1)$. If it equals to zero, then the two items have no similarity. If the absolute value of $sim^*(B_r, B_s)_x$ is closer to 1, then these items are more similar. What should be reminded is that -1 denotes an opposite similarity.

Definition 8. Let $U_{r,s}$ be the set of users who have both assessed items B_r and B_s ($B_r, B_s \in B_y$). $H_{\bar{u}}^j$ denotes the average evaluations of the user A_u ($A_u \in U_{r,s}$) under the criterion C_j . $H_{u,r}^j$ and $H_{u,s}^j$ are the corresponding NHFLTSs given by A_u on these items. The improved similarity measure between the items B_r and B_s ($B_r, B_s \in B_y$) can be computed:

$$sim^*(B_r, B_s) = \frac{\sum_{A_u \in U_{r,s}} \sum_{j=1}^M w_u^j \sum_{l=1}^{\#L} \left[\left(\phi(l)_{u,r}^j - \phi(l)_{\bar{u}}^j \right) \cdot \left(\phi(l)_{u,s}^j - \phi(l)_{\bar{u}}^j \right) \right]}{\sum_{A_u \in U_{r,s}} \sum_{j=1}^M w_u^j \sqrt{\sum_{l=1}^{\#L} \left(\phi(l)_{u,r}^j - \phi(l)_{\bar{u}}^j \right)^2 \cdot \sum_{l=1}^{\#L} \left(\phi(l)_{u,s}^j - \phi(l)_{\bar{u}}^j \right)^2}} \quad (23)$$

where $\phi(l)_{\bar{u}}^j$, $\phi(l)_{u,r}^j$ and $\phi(l)_{u,s}^j$ are the subscripts of l -th term in $H_{\bar{u}}^j$, $H_{u,r}^j$ and $H_{u,s}^j$.

The neighbors of the item B_r are attained by sorting the k largest similarities in $sim^*(B_r, B_s)$ ($r \neq s, s = 1, 2, \dots, Y$). Different from the traditional top-N approach, the proposed method stores the neighbor information without changing non- k similarities into zero.

Step 4. Apply the HFL-MAPPACC method in the user A_x profile and recommend new items.

In this step, the HFL-MAPPACC method can be carried out to gain the preference degrees of the used items in the profile of A_x . Combined with the similarities of neighbors, the preference degrees $\delta(B_s)$ make the recommendation list more personalized.

Definition 9. The item B_r is in the set of neighbors for the items used by A_x and $B_r \notin \Phi(A_x)$, which means that B_r must be a new item for A_x . According to the ranking result of $B_s \in \Phi(A_x)$ and the improved cosine similarity measure between B_r and B_s , the recommendation strength is

$$R(B_r) = \sum_{B_s \in \Phi(A_x)} \delta(B_s) \times \text{sim}^*(B_r, B_s) \quad (24)$$

Finally, the system will recommend the user A_x a set of items, which have the biggest N recommendation strengths among B_r . In next section, an example will be explained to exhibit the program. The whole program is illustrated in Figure 1. Comparative analyses will be carried out in order to testify the superiority of the improvements presented in this paper.

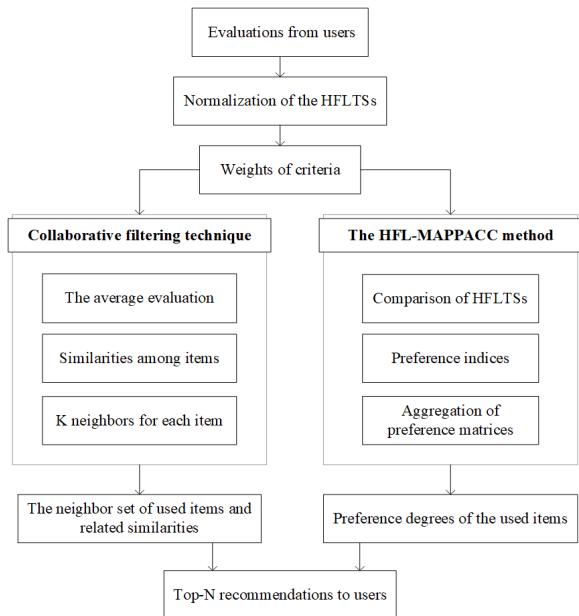


Figure 1. The flow chart for the proposed method

4. An Illustrative Example

In this section, given the stochastic HFLTSS as the evaluations from users, the proposed method is illustrated to recommend suitable doctors for an active user.

4.1 Background and Data Generation

People are forming the habit of getting advices online first when they feel unhealthy. Therefore, companies try to provide high quality service on the mobile platforms. Along with the mutual interface and mechanism, the service quality of doctors plays a great influence on the user satisfaction. When users enter the system, they

tend to select the doctors who are recommended in the lists. However, users often focus on different aspects of doctors. Some feel satisfied if they get quick response, while some users prefer to get clear explanations of the diagnosis and treatment plan even if doctors response more slowly. Hence, it is essential to figure out the preference of users by analyzing their evaluations and recommending suitable doctors.

In the example, the active user A_1 is using the platform to get some advices about diarrhea. There are 50 doctors available and they have been labelled from B_1 to B_{50} . Besides, 500 users, including A_1 , have made effective evaluations about their performance from four aspects: attitude C_1 , response C_2 , professionalism C_3 and treatment effect C_4 . The number of doctors that each user has consulted are stochastic numbers from 3 to 10, and their labels are generated among B_j ($j \leq 50$) without repetition. Then, the evaluations are randomly generated in the form of HFLTSS based on the following linguistic term set:

$$S = \left\{ \begin{array}{l} s_{-3} = \text{very bad}, s_{-2} = \text{bad}, s_{-1} = \text{slightly bad}, s_0 = \text{fair}, \\ s_1 = \text{slightly good}, s_2 = \text{good}, s_3 = \text{very good} \end{array} \right\}$$

The lengths of HFLTSS are randomly different but no more than 3. This method will be illustrated based on the generated data.

4.2 The Numerical Example

Based on the proposed item-based CF method, the recommendations are completed step by step:

Step 1. The profile of A_1 shows that A_1 has consulted four doctors B_4, B_{12}, B_{34} and B_{15} successively. $\Phi(A_1) = \{B_4, B_{12}, B_{34}, B_{15}\}$. The original evaluations in the profile of A_1 is:

$$A_1 = \begin{bmatrix} \{s_0, s_1, s_2\} & \{s_0, s_1\} & \{s_{-1}, s_0\} & \{s_1\} \\ \{s_1, s_2\} & \{s_1, s_2\} & \{s_2\} & \{s_2, s_3\} \\ \{s_0, s_1\} & \{s_2\} & \{s_1, s_2\} & \{s_1, s_2, s_3\} \\ \{s_1, s_2\} & \{s_0\} & \{s_0, s_1\} & \{s_2\} \end{bmatrix}$$

Step 2. Because the lengths of the HFLTSS are different, the normalization step should be carried out. For example, $H_{1,12}^1 = \{s_1, s_2\}$ should be extended by adding its score $s_{\bar{\sigma}(l)} = s_{(1+2)/2} = s_{1.5}$, and $H_{1,34}^1 = \{s_0, s_1\}$ is updated to $\{s_1, s_{1.5}, s_2\}$. Then, the elements are ordered in every HFLTSS. After the normalization, the matrix is changed into:

$$A_1 = \begin{bmatrix} \{s_0, s_1, s_2\} & \{s_0, s_1\} & \{s_{-1}, s_0\} & \{s_1, s_1, s_1\} \\ \{s_1, s_{1.5}, s_2\} & \{s_1, s_2\} & \{s_2, s_2\} & \{s_2, s_{2.5}, s_3\} \\ \{s_0, s_{0.5}, s_1\} & \{s_2, s_2\} & \{s_1, s_2\} & \{s_1, s_2, s_3\} \\ \{s_1, s_{1.5}, s_2\} & \{s_0, s_0\} & \{s_0, s_1\} & \{s_2, s_2, s_2\} \end{bmatrix}$$

Using the HFLA operator, the information can be aggregated and the average evaluation H_1^j of A_1 can be obtained:

$$H_1^j = (\{s_{0.5}, s_{1.125}, s_{1.75}\} \quad \{s_{0.75}, s_{1.25}\} \quad \{s_{0.25}, s_{1.25}\} \quad \{s_{1.5}, s_{1.875}, s_{2.25}\})$$

According to Model 1, the weights of criteria that belong to the user A_1 are obtained. $w_1^j = (0.145, 0.29, 0.352, 0.214)$, which indicates the user is more sensitive to the professionalism of doctors and their speed of response.

Step 3. Actually, the first two steps are also happening when other users come into the system. At the same time, the similarities among doctors will be updated regularly. To concisely show the way of getting the similarity results, the process of similarity computation of the doctor B_1 with the four doctors B_4, B_{12}, B_{34} and B_{15} , who are in the profile of the user A_1 , is introduced. After viewing the dataset, the information of users who have consulted these doctors is extracted. The similarities from the individuals' point of view are obtained by Formula (22), a part of which are shown in Table 2. Besides, the final similarities computed by Formula (23) are also exhibited.

Table 2. The similarities for pairs of doctors from personal prospects and the public view

	A_{21}	A_{35}	A_{67}	The public
$B_1 \times B_4$	0.05			0.248
$B_1 \times B_{12}$	0.78	0.45	-0.93	-0.316
$B_1 \times B_{34}$		-0.02		-0.307
$B_1 \times B_{15}$		0.33		0.063

As depicted above, individuals may express different opinions about the same doctors. When A_{21} and A_{35} give similar evaluations on B_1 and B_{12} , the user A_{67} expresses an nearly opposite attitude towards them with the similarity degree of -0.93. The final similarity measure between B_1 and B_4 is 0.248, indicating a moderate extent of similarity. The final similarity measure for B_1 and B_{12} is -0.316 while -0.307 is the final similarity measure between B_1 and B_{34} . The negative numbers show that the service provided by the doctor B_1 is at a different level compared with that of the doctors B_{12} and B_{34} . These similarities from the perspectives of individuals and the public can be stored in the database which must be updated regularly with increasing evaluations. Moreover, k equals to 10 in this example. The neighbors of the doctors B_4, B_{12}, B_{34} and B_{15} are shown in Table 3. The neighbor sets are also stored in the system. When users search, these data can be used to get the recommendations.

Step 4. When the active user A_1 is looking for help on the platform, the preference degrees of the used items should be computed by the HFL-MAPPACC approach. The evaluations of four doctors should be compared under these criteria by Function (13). The results are shown in Table 4.

Table 4. The comparison scores ζ

	C_1	C_2	C_3	C_4
B_4	0.833	0.4375	-0.5625	1
B_{12}	1.4583	1.4375	2	1.8333
B_{34}	0.4583	2	1.4375	2.4583
B_{15}	1.4583	0	0.4375	2

Table 3. The neighbor sets

Doctors	The k neighbors	The union of neighbors
B_4	$\{B_1, B_2, B_{26}, B_{27}, B_{29}, B_{35}, B_{36}, B_{38}, B_{40}, B_{41}\}$	$\left\{ B_1, B_2, B_3, B_6, B_8, B_9, B_{10}, B_{13}, B_{14}, B_{16}, B_{17}, B_{18}, B_{19}, B_{21}, B_{25}, B_{26}, B_{27}, B_{29}, B_{30}, B_{33}, B_{35}, B_{36}, B_{37}, B_{38}, B_{39}, B_{40}, B_{41}, B_{45}, B_{46}, B_{47}, B_{48}, B_{49}, B_{50} \right\}$
B_{12}	$\{B_3, B_{13}, B_{14}, B_{17}, B_{30}, B_{40}, B_{41}, B_{45}, B_{48}, B_{49}\}$	
B_{34}	$\{B_3, B_9, B_{16}, B_{18}, B_{19}, B_{33}, B_{39}, B_{47}, B_{49}, B_{50}\}$	
B_{15}	$\{B_1, B_6, B_8, B_9, B_{10}, B_{21}, B_{25}, B_{37}, B_{40}, B_{46}\}$	

Because $\zeta(H_{1,4}^1) = 0.8333 < \zeta(H_{1,12}^1) = 1.4583$ and $\zeta(H_{1,4}^2) = 0.4375 < \zeta(H_{1,12}^2) = 1.4375$, it is obvious that B_{12} has a better performance than B_4 regarding to the pair of criteria C_1 and C_2 . Then, $\pi_1^{12}(B_4, B_{12}) = 0$ and $\pi_1^{12}(B_{12}, B_4) = 1$. When comparing the doctor B_4 with B_{34} , it can be seen that $\zeta(H_{1,4}^1) = 0.8333 > \zeta(H_{1,34}^1) = 0.4583$, but $\zeta(H_{1,4}^2) = 0.4375 > \zeta(H_{1,34}^2) = 2$. The result of comparison meets the fourth condition mentioned in Equation (14). Thus, the preference index of the doctor B_4 with B_{34} related to the pair of criteria C_1 and C_2 is derived:

$$\pi_1^{12}(B_4, B_{34}) = \frac{w_1^1 d(H_{1,4}^1, H_{1,34}^1)}{w_1^1 d(H_{1,4}^1, H_{1,34}^1) + w_1^2 d(H_{1,4}^2, H_{1,34}^2)} = 0.143$$

Hence, $\pi_1^{12}(B_{34}, B_4) = 1 - \pi_1^{12}(B_4, B_{34}) = 0.857$. It implies that the doctor B_{34} has a better performance than B_4 considering the attitude of service and the speed of response. Repeating the procedure, the computation for preference indices can be finished. As shown below, the preference indices with 6 pairs of criteria have been accomplished.

$$\begin{aligned} \Pi_1^{12} &= \begin{bmatrix} 0.5 & 0 & 0.143 & 0.667 \\ 1 & 0.5 & 0.5 & 1 \\ 0.857 & 0.5 & 0.5 & 0.8 \\ 0.333 & 0 & 0.2 & 0.5 \end{bmatrix} & \Pi_1^{13} &= \begin{bmatrix} 0.5 & 0 & 0.093 & 0 \\ 1 & 0.5 & 1 & 1 \\ 0.907 & 0 & 0.5 & 0.708 \\ 1 & 0 & 0.292 & 0.5 \end{bmatrix} \\ \Pi_1^{14} &= \begin{bmatrix} 0.5 & 0 & 0.184 & 0 \\ 1 & 0.5 & 0.575 & 0 \\ 0.816 & 0.425 & 0.5 & 0.425 \\ 1 & 1 & 0.575 & 0.5 \end{bmatrix} & \Pi_1^{23} &= \begin{bmatrix} 0.5 & 0 & 0 & 0.292 \\ 1 & 0.5 & 0.548 & 1 \\ 1 & 0.452 & 0.5 & 1 \\ 0.708 & 0 & 0 & 0.5 \end{bmatrix} \\ \Pi_1^{24} &= \begin{bmatrix} 0.5 & 0 & 0 & 0.404 \\ 1 & 0.5 & 0 & 0.753 \\ 1 & 1 & 0.5 & 1 \\ 0.596 & 0.247 & 0 & 0.5 \end{bmatrix} & \Pi_1^{34} &= \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 1 & 0.5 & 0.622 & 0.787 \\ 1 & 0.378 & 0.5 & 1 \\ 1 & 0.213 & 0 & 0.5 \end{bmatrix} \end{aligned}$$

By aggerating all the matrices together, the global matrix of the user is obtained:

$$\Pi_1 = \begin{bmatrix} 0.5 & 0 & 0.058 & 0.227 \\ 1 & 0.5 & 0.542 & 0.8 \\ 0.943 & 0.459 & 0.5 & 0.855 \\ 0.774 & 0.201 & 0.146 & 0.5 \end{bmatrix}$$

By Function (17), the preference degrees can be gained by adding the elements in each row of Π_1 , except for the preference indices of each doctor with themselves.

$$\delta_1(B_4) = 0 + 0.058 + 0.227 = 0.285$$

$$\delta_1(B_{12}) = 1 + 0.542 + 0.78 = 2.322$$

$$\delta_1(B_{34}) = 0.942 + 0.459 + 0.855 = 2.257$$

$$\delta_1(B_{15}) = 0.774 + 0.201 + 0.146 = 1.121$$

Therefore, the final ranking is $B_{12} \succ B_{34} \succ B_{15} \succ B_4$ with ranking indices 2.322, 2.257, 1.121, and 0.285. For the user A_1 , the performance of the doctor B_{12} win a razor-thin victory along with that of B_{34} . B_4 has certainly left a bad impression on this user. With the ranking indices, it is easy to derive the suitable recommendations for the user A_1 . With Equation (24), the ranking indices and the similarity information of the k neighbors with the consulted doctors, the recommendation list is $\{B_3, B_9, B_{33}, B_{49}, B_{16}, B_{39}, B_{17}, B_{29}, B_{50}, B_{37}\}$ in descending order of recommendation strengths.

5. Comparative Analysis

In this section, the validity of HFL-MAPPACC is verified initially. By a comparative analysis with other classical method, the strengths of the HFL-MAPPACC are also conveyed. Then, because some operations in traditional top-N method have been changed, it is necessary to compare the results generated by the traditional method with the results obtained in the method proposed in this paper.

5.1 Comparison with Other Decision-Making Methods

The comparative analysis is based on the example of the user A_1 in Section 4. Before comparing the results, the basic backgrounds for these methods are illustrated in Table 5. In addition, if the method does not introduce the means of getting weight vector, the weight vector is supposed to be the same as that derived by the method presented in this article where $w^j = (0.145, 0.29, 0.352, 0.214)$.

Both the BR-method and the LW-method consider multiple criteria but overlook the weights in the specific computation. In the LX-method, ξ affects the element added in the shorter HFLTS in the step of normalization, and p decides the HFL Euclidean distance measure. θ represents the preference of the user over different criteria. ξ expresses the relative importance between the distance of alternatives with the maximum HFLE and the distance of weights with the biggest weight of criterion in the LW-method. δ , σ and μ are the thresholds to carry out the PIR framework in this method. The weights in the DC-method are also derived by maximizing deviation method. But because it develops a new fuzzy envelope and cosine similarity measure, the results of weights are quite different from that in this method, especially for the criteria C_3 and C_4 . The method

Table 5. The basic backgrounds of different decision-making (DM) methods

Methods	The DM method	Parameters
Our method	MAPPACC	$w^j = (0.145, 0.29, 0.352, 0.214)$
BR-method (Beg & Rashid, 2013)	TOPSIS	No parameters and weights
LX-method (Liao et al., 2015)	VIKOR	$\xi = 0.5, p = 1, \theta = 0.5$ $w^j = (0.145, 0.29, 0.352, 0.214)$
LW-method (Liao et al., 2018)	ORESTE	$\xi = 0.5, \delta = 0.002, \sigma = 0.001, \mu = 0.0005$
DC-method (Dong et al., 2018)	QUALIFLEX	$p = q = 1, w^j = (0.175, 0.24, 0.184, 0.403)$

with these two different weights is carried out to see the differences. p and q can influence the BM operator which is an important tool to transform HFLTSS into its envelope in the DC-method.

The ranking results of the five methods are shown in Table 6. They are highly consistent, which can tell that the present method is very effective, while the DC-method gains a moderately distinct ranking.

Table 6. The ranking results of different methods

Methods	The ranking results
Our method	$B_{12} \succ B_{34} \succ B_{15} \succ B_4$
BR-method	$B_{12} \succ B_{34} \succ B_{15} \succ B_4$
LX-method	$B_{12} \succ B_{34} \succ B_{15} \succ B_4$
LW-method	$B_{12} \succ B_{34} \succ B_{15} \succ B_4$
DC-method	$B_{12} \succ B_{15} \succ B_{34} \succ B_4$ with $w^j = (0.175, 0.24, 0.184, 0.403)$
DC-method	$B_{12} \succ B_{34} \succ B_{15} \succ B_4$ with $w^j = (0.145, 0.29, 0.352, 0.214)$

Delving into the original reason, it could be found that the DC-method improves the cosine similarity with the hesitant degree of every HFLTSS. In the present example, HFLTSS are in different lengths, especially under the criterion C_4 , resulting in a bigger weight of the criterion C_4 . The evaluation of B_{34} under the criterion C_4 is $\{s_1, s_2, s_3\}$, but for B_{15} it is only $\{s_2\}$ which has an advantage over the hesitant degree. Hence, with the weights derived by its improved similarity measure, B_{15} behaves better than B_{34} . Since this ranking result has been testified to be effective, their advantages and shortcomings should be discussed:

1. The BR-method: its computations are based on the envelope of HFLTSS. Although it

makes the whole process easier to carry out, information may get lost in dealing with massive data. In the proposed method, the original evaluations are retained and the personal attitude are considered towards different criteria, which makes it more adaptive to the complex environment.

2. The LX-method: the HFL-VIKOR method is a good tool to reach the balance between the total community and individual satisfaction. But the purpose of recommender system is to provide personalized recommendation for individuals, which is the focal point of the method proposed in this paper.
3. The LW-method: this method needs to set some thresholds to form PIR framework and the final strong rankings. When $\sigma = 0.01, \mu = 0.005$, the results show that B_{12} and B_{34} are indifferent. When $\sigma = 0.015, \mu = 0.075$, it turns out that B_{12}, B_{34} and B_{15} have no difference. It is hard to decide the suitable thresholds for individuals. That will be an unstable factor in the system and affect the accuracy of recommendations. Moreover, it requires individuals to give weights in the form of HFLTSS which is difficult to obtain directly from users.
4. The DC-method: it transforms the HFLTSS into its fuzzy envelope so that the lengths of HFLTSS play no role in the computation. Besides, the hesitant degree is taken into consideration throughout the whole system. However, the QUALIFLEX approach should enumerate all the possible ranking conditions. The permutation with largest concordance/discordance indices is the best condition. Apparently, the final two steps cost much time with massive alternatives. Moreover, the indices only manifest the scores of each permutation without showing the concrete advantage of each item over the others, which provides no assistance in the recommendation programs.

5.2 Comparison with Original Top-N Method

The proposed method improves some aspects of the traditional top-N method. In this subsection, the superiority of these advancements is testified by comparison. The first improvement is getting the k neighbors of each item but retaining other similarities without changing them into zero. The second one is using the HFL-MAPPACC method to gain the preference degrees of the used items, and taking the results as weights to aggregate these similarities. To carry out the comparison comprehensively, various conditions are set to see the differences (Table 7). In these conditions, “neighbors” denotes the scheme of choosing items which have the k biggest similarities with used items. “zero” means that the procedure changes the similarities of non- k neighbors into zero. And if the condition applies the HFL-MAPPACC method, it concerns the preferences of users. What should be reminded is that Condition 3 is the present method, and Condition 6 is the traditional item-based top-N method.

It is obvious that the recommendation lists of Conditions 1 and 3 are highly consistent. The results of Conditions 2 and 4 are equal. But the results under the strategy of changing non- k similarities into 0 are quite different from others. By comparing these results, it can be concluded that the strategy of selecting the k neighbors without changing others into zero can remain the accuracy of recommendations. Meanwhile, after applying this strategy, only 33 items are needed to be ranked in the final step rather than all 50 items. If the database contains more items, then the efficiency will also improve.

As shown in Table 7, B_1 is in the lists of Conditions 2, 4 and 6. Besides, it even exceeds B_3 in the

final condition, while B_3 has a good performance under other conditions. Tracing back the similarities of B_1 and B_3 with $\{B_4, B_{12}, B_{34}, B_{15}\}$, it is found that the data related to B_1 is (0.248, -0.316, -0.307, 0.063), which indicates that B_1 is more positively similar to B_4 and B_{15} through the public awareness. And B_1 is the neighbor of B_4 and B_{15} . The relevant similarities of B_3 are (-0.327, 0.053, 0.202, -0.098) whose performance is more akin to those of B_{12} and B_{34} . So, when the algorithm does not apply the HFL-MAPPACC method to manifest the preference attitudes of the active user, B_1 surpasses many doctors by a higher score, especially with the strategy of changing non- k largest similarities into 0. For instance, the final score of B_1 is 0.311 in Condition 6, but B_3 gets 0.255 whose similarities with B_{12} and B_{34} are retained in the computation. However, B_{12} and B_{34} are more likable for the user A_1 based on the rankings. The chance is that A_1 has better impression on B_3 rather than B_1 . Hence, the HFL-MAPPACC method can help the system to derive more personalized lists.

6. Conclusion and Future Work

Aiming to better clarify the preference of users and meet their vague necessities, uncertain linguistic evaluations related to multi-criteria items are paid attention in this paper to accomplish the personalized recommender system. A novel cosine similarity measure of HFLTSs is proposed to mitigate the effect of individual evaluation tendency and ascertain the realistic attitude of the public. Besides, by executing the fuzzy MAPPACC approach, the proposed algorithm demonstrates the preferences of each user for purchased items and combines them with the similarities so as to gain the recommendations.

Table 7. Strategy taken and results under these conditions

Condition	neighbors	zero	MAPPACC	the top-ten items
1	-	-	-	$\{B_3, B_9, B_{33}, B_{49}, B_{16}, B_{39}, B_{17}, B_{29}, B_{50}, B_{37}\}$
2	-	-	-	$\{B_{29}, B_9, B_3, B_{40}, B_{37}, B_1, B_{50}, B_{33}, B_{49}, B_{41}\}$
3	√	-	√	$\{B_3, B_9, B_{33}, B_{49}, B_{16}, B_{39}, B_{17}, B_{29}, B_{50}, B_{37}\}$
4	√	-	-	$\{B_{29}, B_9, B_3, B_{40}, B_{37}, B_1, B_{50}, B_{33}, B_{49}, B_{41}\}$
5	√	√	√	$\{B_9, B_{33}, B_3, B_{16}, B_{50}, B_{30}, B_{17}, B_{48}, B_{39}, B_{45}\}$
6	√	√	-	$\{B_9, B_1, B_{33}, B_3, B_{16}, B_{29}, B_{30}, B_{50}, B_{36}, B_{21}\}$

Therefore, the chances are that the final recommendations can be chosen and liked by the active users. The proposed approach has been implemented in the case of recommending doctors on the online platform, which is innovative for the application of recommender systems. By comparative analysis with classical multi-criteria decision-making methods, the proposed method offers stable and personalized recommendations without losing much information.

Considering the hypothesis and theory, the proposed method still has some limitations. For the reason that uncertain linguistic evaluations contain more complex information than numerical numbers, its recognition and computation are more complicated as well. Simultaneously, it consumes much time to accomplish the algorithm, and operation costs are expensive to maintain the system. Besides, the advantages of the new algorithm have been theoretically analyzed

through numerical examples and comparisons. But, in order to gain an excellent performance in practical systems, the method should be adapted to diverse features of these systems.

Thus, we point out some further research orientations as follows:

1. It is essential to delve into the recognition of uncertain linguistic information. With massive data, the precise way to transform the natural language into a computable form can make the system agile and reliable.
2. Many realistic applications focus on the efficiency. Another future task is to set up the mechanism which can cost less time and assure the quality of recommendations.
3. Implementing the proposed algorithm into a practical system is the fourth direction. The parameters to measure the result are also worth discussing.

REFERENCES

- Adomavicius, G. & Tuzhilin, A. (2005). 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*, 17(6), 734-749.
- Ajoudanian, S. & Abadeh, M. N. (2019). Recommending Human Resources to Project Leaders Using a Collaborative Filtering-Based Recommender System: Case Study of GitHub, *IET Software*, 13(5), 379-385.
- Beg, I. & Rashid, T. (2013). TOPSIS for Hesitant Fuzzy Linguistic Term Sets, *International Journal of Intelligent Systems*, 28(12), 1162-1171.
- Chen, X. M. & Liu, Y. Z. (2016). Method of Group Recommender Systems with Unknown Attribute Weights in a Multi-Granular Hesitant Fuzzy Linguistic Term Environment, *Control and Decision*, 31(9), 1631-1637.
- Chen, X. M. & Liu, Y. Z. (2018). Multi-Granular Hesitant Fuzzy Linguistic Term Sets and Their Application in Group Recommendation Based on Entropy Measure and VIKOR Method, *Control and Decision*, 33(1), 111-118.
- Deshpande, M. & Karypis, G. (2004). Item-Based top-N Recommendation Algorithms, *ACM Transactions on Information Systems*, 22(1), 143-177.
- Dong, J. Y., Chen, Y. & Wan, S. P. (2018). A Cosine Similarity Based QUALIFLEX Approach with Hesitant Fuzzy Linguistic Term Sets for Financial Performance Evaluation, *Applied Soft Computing*, 69, 316-329.
- Farhadinia, B. (2016). Multiple Criteria Decision-Making Methods with Completely Unknown Weights in Hesitant Fuzzy Linguistic Term Setting, *Knowledge-Based Systems*, 93, 135-144.
- Feng, X. Q., Tan, Q. Y. & Wei, C. P. (2018). Hesitant Fuzzy Linguistic Multi-Criteria Decision Making Based on Possibility Theory, *International Journal of Machine Learning and Cybernetics*, 9(9), 1505-1517.
- Filip F. G (2020). DSS—A Class of Evolving Information Systems. In: Dzemyda G. Bernatavičienė J., Kacprzyk J. (eds). *Data Science: New Issues, Challenges and Applications. Studies in Computational Intelligence*, 869, 253-277. Springer Cham.
- Guia, M., Silva, R. R. & Bernardino, J. (2019). A Hybrid Ontology-Based Recommendation System in E-Commerce, *Algorithms*, 12(11), 239.
- Kompan, M. & Bielikova, M. (2013). Personalized Recommendation for Individual Users Based on the Group Recommendation Principles, *Studies in Informatics and Control*, 22(3), 331-342. DOI: 10.24846/v22i3y201310
- Liao, H. C., Wu, X. L., Liang, X. D., Xu, J. P. & Herrera, F. (2018). A New Hesitant Fuzzy Linguistic ORESTE Method for Hybrid Multicriteria Decision Making, *IEEE Transactions on Fuzzy Systems*, 26(6), 3793-3807.

- Liao, H. C. & Xu, Z. S. (2015). Approaches to Manage Hesitant Fuzzy Linguistic Information Based on the Cosine Distance And Similarity Measures for HFLTSS and Their Application in Qualitative Decision Making, *Expert Systems with Applications*, 42(12), 5328-5336.
- Liao, H. C., Xu, Z. S. & Zeng, X. J. (2014). Distance and Similarity Measures for hesitant Fuzzy Linguistic Term Sets and Their Application in Multi-Criteria Decision Making, *Information Sciences*, 271, 125-142.
- Liao, H. C., Xu, Z. S. & Zeng, X. J. (2015). Hesitant Fuzzy Linguistic VIKOR Method and Its Application in Qualitative Multiple Criteria Decision Making, *IEEE Transactions on Fuzzy Systems*, 23(5), 1343-1355.
- Lin, M. W., Xu, Z. S., Zhai, Y. L. & Yao, Z. Q. (2018). Multi-attribute Group Decision-Making under Probabilistic uncertain Linguistic Environment, *Journal of the Operational Research Society*, 69(2), 157-170.
- Matarazzo, B. (1986). Multicriterion Analysis of Preferences by Means of Pairwise Actions and Criterion Comparisons (MAPPACC), *Applied Mathematics and Computation*, 18(2), 119-141.
- Renjith, S., Sreekumar, A. & Jathavedan, M. (2020). An Extensive Study on the Evolution of Context-Aware Personalized Travel Recommender Systems, *Information Processing & Management*, 57(1), 102078.
- Rodriguez, R. M., Martinez, L. & Herrera, F. (2012). Hesitant Fuzzy Linguistic Term Sets for Decision Making, *IEEE Transactions on Fuzzy Systems*, 20(1), 109-119.
- Shambour, Q., Hourani, M. A. & Fraihat, S. (2016). An Item-Based Multi-Criteria Collaborative Filtering Algorithm for Personalized Recommender Systems, *International Journal of Advanced Computer Science and Applications*, 7(8), 274-279.
- Vercruyssen, N., Tomozei, C., Furdu, I., Varlan, S. & Amancei, C. (2015). Collaborative Recommender System Development with Ubiquitous Computing Capability for Risk Awareness, *Studies in Informatics and Control*, 24(1), 91-100. DOI: 10.24846/v24i1y201510
- Wang, Y. M. (1997). Using the Method of Maximizing Deviation to Make Decision for Multiindices, *Journal of Systems Engineering and Electronics*, 8(3), 21-26.
- Wei, C. P., Ren, Z. L. & Rodriguez, R. M. (2015). A Hesitant Fuzzy Linguistic TODIM Method Based on a Score Function, *International Journal of Computational Intelligence Systems*, 8(4), 701-712.
- Wei, C. P., Zhao, N. & Tang, X. J. (2014). Operators and Comparisons of Hesitant Fuzzy Linguistic Term Sets, *IEEE Transactions on Fuzzy Systems*, 22(3), 575-585.
- Wu, J. W., Yu, J. H. & Yu, X. H. (2010). An Item-based Weighted Collaborative Recommended Algorithm Based on Screening Users' Preferences. In *2010 International Conference on Computer Application and System Modeling* (pp. 1614-1614). DOI: 10.1109/ICCASM.2010.5619326
- Xu, Y. W., Meng, F. Y. & Wang, N. (2019). Correlation Coefficients of Linguistic Interval Hesitant Fuzzy Sets and Their Application, *Iranian Journal of Fuzzy Systems*, 16(4), 65-81.
- Xu, Z. S. (2005). Deviation Measures of Linguistic Preference Relations in Group Decision Making, *Omega-International Journal of Management Science*, 33(3), 249-254.
- Zhang, Z. M. & Wu, C. (2014). Hesitant Fuzzy Linguistic Aggregation Operators and Their Applications to Multiple Attribute Group Decision Making, *Journal of Intelligent & Fuzzy Systems*, 26(5), 2185-2202.