Title	Two-probabilities focused combination in recommender systems
Author(s)	Nguyen, Van-Doan; Huynh, Van-Nam
Citation	International Journal of Approximate Reasoning, 80: 225-238
Issue Date	2016-09-20
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/15435
Rights	Copyright (C)2016, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license (CC BY-NC-ND 4.0). [http://creativecommons.org/licenses/by-nc-nd/4.0/] NOTICE: This is the author's version of a work accepted for publication by Elsevier. Changes resulting from the publishing process, including peer review, editing, corrections, structural formatting and other quality control mechanisms, may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Van-Doan Nguyen, Van-Nam Huynh, International Journal of Approximate Reasoning , 80, 2016, 225-238, http://dx.doi.org/10.1016/j.ijar.2016.09.005
Description	

Two-Probabilities Focused Evidence Combination in Recommender Systems

Van-Doan Nguyen, Van-Nam Huynh

Japan Advanced Institute of Science and Technology (JAIST), Japan.

Abstract

In this paper, we develop a new method, called 2-probabilities focused combination, for combining pieces of evidence of users' preferences in recommender systems based on Dempster-Shafer theory. This method focuses on significant focal elements in focal sets of mass functions only; whereas, the remaining focal elements are considered as noise and then transferred to the whole set element. With the new method, in the systems, users' preferences are represented as special mass functions called 2-probabilities focused mass functions; and for evidence combination, Dempster's rule of combination is applied to combine 2-probabilities focused functions, and the combination results are transformed into corresponding 2-probabilities focused mass functions. To evaluate as well as demonstrating the advantage of the new method, a baseline method called 2-points focused combination is selected for performance comparison in a range of experiments on two recommender systems using Movielens and Flixster data sets.

Keywords: Evidence Combination, Uncertain Reasoning, Dempster-Shafer Theory, Recommender Systems, Collaborative Filtering.

1. Introduction

On the one hand, in doing online business, online providers make efforts to suggest suitable items (products or services) to each specific online user (customer) to increase sales growth. On the other hand, while doing shopping on the Internet, online users want to not only share their opinions with one

This paper is a significantly extended and revised version of the conference paper presented at IUKM-2015 [35].

another but also be recommended the items related to what they are looking for. During the last two decades, recommender systems (RSs) [2, 11, 30, 40] have been developed to satisfy both online suppliers and online users, as well as being widely integrated in e-commerce applications [3, 29, 41, 43, 44]. From the viewpoint of online providers, a challenge of RSs is how to generate suitable recommendations among many potential items whereas evidence of users' preferences is commonly uncertain, imprecise or incomplete.

As observed, in RSs, there are two main methods to collect information about users' preferences from different sources. The first method is to obtain the information explicitly from user profiles [31, 39, 49] or user ratings on items [12, 18, 26]. The second method is to gather the information implicitly by monitoring users' behavior [13, 17, 27, 33, 37], or by extracting from context information [1, 14, 16, 36, 50] or social networks [6, 28, 34]. Naturally, the more sources of information about user's preferences are available, the more effective recommendation decisions will be.

It is known from the literature that Dempster-Shafer theory (DST) [19, 42] is a general framework to model uncertain, imprecise, and incomplete information. In addition, this theory provides a powerful tool to combine information from different sources. So far, DST has been applied in a variety of applications [20, 21, 23, 25, 32, 38] including RSs [24, 34, 36, 48, 50]. In RSs based on DST, users' preferences or ratings on items are modeled by using mass functions [19, 42], and evidence combination tasks play a significant role as well as being used frequently.

Currently, Dempster's rule of combination [19] is employed in almost all of RSs based on DST. However, in the systems, when several or a large number of mass functions are combined together by using this method, in the focal set of the combination result, usually, most of focal elements have very low probabilities whereas a few focal elements have high probabilities. Especially, in the case where rating domains contain many elements, for example 10 elements as in Flixster data set [34], the number of focal elements with very low probabilities can be numerous. It can be seen that when combining two mass functions by using Dempster's rule of combination, the focal elements with very low probabilities can lead to (1) poor performance in computational time and (2) the unsatisfactory results [46, 52] in case these mass functions are highly conflicting. Besides, in these systems, highly conflicting ratings are common because of the diversity of users.

In [5, 8, 9], the authors have developed an evidence combination method, known as 2-points focused combination, that is capable of distinguishing focal elements with high probabilities from the ones with very low probabilities. But, with this method, when a focal set contains several focal elements with the same high probabilities, users hesitate to select focal elements into the corresponding focal element triplet [5, 8, 9] that consists of two focal elements with the highest probabilities and the whole set element. The reason is that different selections may lead to different combination results.

In this paper, we develop a so-called 2-probabilities focused combination for combining evidence in RSs based on DST. This new method concentrates only on significant focal elements defined as the ones with probabilities in top two highest probabilities in the corresponding mass functions, and ignores the other focal elements excluding the whole set element. With this characteristic, the new method helps the systems improve computational time as well as avoiding the unsatisfactory results. Furthermore, from a given mass function, we can induce only one 2-probabilities focused mass function; thus, we can get only one combination result when combining mass functions by using 2-probabilities focused combination method. That means the new method is also able to overcome the weakness of 2-points focused combination method by cause of non-uniqueness combination results.

In the experiments, to demonstrate the effectiveness and efficiency of 2-probabilities focused combination method, it is integrated in two RSs based on DST using Movielens and Flixster data sets. The experimental results show that, regarding to recommendation accuracy, this method is better than 2-points focused combination; additionally, the computational time of 2-probabilities focused combination can be comparable to that of 2-points focused combination whose time complexity is linear [7].

The rest of this paper is organized as follows. In the second section, DST is briefly introduced first, and then related work is provided. Next, in the third section, 2-probabilities focused combination method is described. After that, in the fourth section, RSs based on DST, that employ 2-probabilities focused combination method, are shown. In the fifth section, experiments on two different data sets are presented. Finally, in the last section, conclusion remarks are shown.

2. Background and related work

2.1. Dempster-Shafer theory

In the context of this theory, a problem domain is represented by a finite set $\Theta = \{\theta_1, \theta_2, ..., \theta_L\}$ of mutually exclusive and exhaustive hypotheses, called the frame of discernment [19]. Each proposition θ_i with $i = 1, ..., L$, also known as a singleton, denotes the lowest level of discernible information.

A function $m : 2^\Theta \to [0, 1]$ is called a basic probability assignment (BPA) or a mass function if it satisfies $m(\emptyset) = 0$ and $\sum_{A \subseteq \Theta} m(A) = 1$. A subset $A \subseteq \Theta$, with $m(A) > 0$, is called a focal element; and the set of all focal elements is called the focal set. Mass function m is considered to be vacuous if $m(\Theta) = 1$ and $m(A) = 0, \forall A \subset \Theta$. In case the information source providing mass function m has a probability of $\delta \in [0, 1]$ of reliability, mass function m is discounted by a discount rate $1 - \delta$ as below

$$\begin{aligned} m^\delta(A) &= \delta \times m(A), \text{ for } A \subset \Theta; \\ m^\delta(\Theta) &= \delta \times m(\Theta) + (1 - \delta), \text{ for } A = \Theta. \end{aligned} \tag{1}$$

Based on mass function m, a belief function $Bel : 2^\Theta \to [0, 1]$ and a plausibility function $Pl : 2^\Theta \to [0, 1]$ are defined by

$$\begin{aligned} Bel(A) &= \sum_{\emptyset \neq B \subseteq A} m(B), \text{ for } A \subseteq \Theta; \\ Pl(A) &= \sum_{A \cap B} m(B), \text{ for } A \subseteq \Theta. \end{aligned} \tag{2}$$

Besides, mass function m can be mapped to a pignistic probability function Bp [45, 47] defined by

$$Bp(\theta_i) = \sum_{A \subseteq \Theta, \theta_i \in A} \frac{m(A)}{|A|}. \tag{3}$$

In [19], the author introduced a method, called Dempster's rule of combination, for combining pieces of evidence, representing as mass functions. When combining two mass functions m_1 and m_2 by using this method, the combination result, also called the orthogonal sum of m_1 and m_2, is denoted by $(m_1 \oplus m_2)$ and defined as follow

$$\begin{aligned} (m_1 \oplus m_2)(\emptyset) &= 0; \\ (m_1 \oplus m_2)(A) &= \frac{1}{1 - K} \sum_{\{C, D \subseteq \Theta | C \cap D = A\}} m_1(C) m_2(D), \end{aligned} \tag{4}$$

where $K = \sum_{\{C, D \subseteq \Theta | C \cap D = \emptyset\}} m_1(C) m_2(D) \neq 0$, and K represents the basic probability mass associated with conflict. If $K = 1$, then the orthogonal sum $m_1 \oplus m_2$ does not exist.

2.2. Related work

Suppose that there are n pieces of evidence represented by n mass functions defined on the same frame of discernment Θ. Clearly, when these mass functions are combined together by using Dempster's rule of combination, the computational complexity is dominated by the number of elements in Θ, $O(|\Theta|^{n-1})$ in the worst case [9]. Additionally, as mentioned previously, in RSs based on DST, combination tasks are performed very often; therefore, performances of the systems are heavily dependent on these tasks. As observed in the literature, the performances can improve by reducing insignificant focal elements in the corresponding mass functions, but possible answers to questions related to the mass functions are still remained [9]. Over the years, a number of reducing methods for evidence combination have been developed; and they will be briefly presented in the remainder of this section.

2.2.1. Simple and separable support functions.

Let us consider a mass function m defined on the frame of discernment Θ. This mass function is considered to be a simple support mass function focusing on $A \subset \Theta$ if it can be represented in a form as below

$$\begin{aligned} &m(A) = p; \\ &m(\Theta) = 1 - p; \\ &m(B) = 0, \text{ for } A \neq B \subset \Theta; \end{aligned} \tag{5}$$

where $p \in (0, 1]$ is called the degree of support [19]. And, a separable support function is defined as either a simple support function or a combination result of two or more simple support functions [19].

2.2.2. Dichotomous function.

In [4], the author introduced an evidence combination method based on dichotomous mass functions. A mass function m is called a dichotomous if its focal set, denoted by F, consists of only three focal elements A, $\Theta\backslash A$, and Θ with $A \subset \Theta$; in other words, $F = \{A, \Theta\backslash A, \Theta\}$ and $m(A)+m(\Theta\backslash A)+m(\Theta) = 1$. In this case, $m(A)$ is the degree of support for A, $m(\Theta\backslash A)$ is the degree of support for the refutation of A, and $m(\Theta)$ is the degree of the support not assigned for or against the proposition A.

2.2.3. Triplet mass function.

In [5, 8, 9], the authors have developed a new structure known as a focal element triplet to represent pieces of evidence as well as a combination method called 2-points focused combination for evidence combination. Originally, focal element triplets contain singletons; however, this structure can be extended for representing composites.

Let us consider a mass function $m : 2^\Theta \to [0, 1]$ with its focal set contains n elements, denoted by $F = \{A_1, A_2, ..., A_n\}$. Based on this mass function, a focal element triplet is defined as an expression of the form $< X_1, X_2, X_3 >$ where X_1, X_2 and X_3 are defined as follows

$$\begin{aligned}
&X_1 = A_i, \text{ with } m(A_i) = max\{m(A_1), m(A_2), ..., m(A_n)\};\\
&X_2 = A_j, \text{ with } m(A_j) = max\{m(A_k) \in F \backslash A_i\};\\
&X_3 = \Theta.
\end{aligned} \tag{6}$$

The triplet mass function [5, 8, 9] corresponding to this focal element triplet is denoted by $\bar{m}$ and given by

$$\begin{aligned}
&\bar{m}(X_1) = m(A_i);\\
&\bar{m}(X_2) = m(A_j);\\
&\bar{m}(X_3) = 1 - m(A_i) - m(A_j).
\end{aligned} \tag{7}$$

Supposing that, we need to combine two triplet mass functions $\bar{m}_1$ and $\bar{m}_2$ together. With 2-points focused combination method, these two mass functions are combined by using Dempster's rule of combination first; and then the combination result, which can consist of three, four or five different focal elements, is transformed into a corresponding triplet mass function [9].

In some scenarios, however, this method is not effective, as shown in Example 1.

> **Example 1** Assuming that, in a RS based on DST with a rating domain $\Theta = \{1, 2, 3, 4, 5\}$, we need to combine two ratings by using 2-points focused mass functions. These ratings are represented by two mass functions denoted by m_1 and m_2 as well as being depicted in Tables 1 and 2, respectively. When converting into triplet mass functions, mass function m_1 can be one of three different triplet mass functions, called $\bar{m}_1^{(1)}$, $\bar{m}_1^{(2)}$, and $\bar{m}_1^{(3)}$, as shown in Tables 3, 4 and 5, respectively; and mass function m_2 has an only one triplet mass function, denoted by $\bar{m}_2$, described in Table 6. Regarding three triplet mass function options of mass function m_1, when combing two mass functions m_1 and m_2 using

Table 1: Mass function m_1

$m_1(\{1\})$	=	0.30
$m_1(\{3\})$	=	0.30
$m_1(\{4\})$	=	0.04
$m_1(\{5\})$	=	0.30
$m_1(\{1,2,3,4,5\})$	=	0.06

Table 2: Mass function m_2

$m_2(\{1\})$	=	0.40
$m_2(\{2\})$	=	0.10
$m_2(\{3\})$	=	0.07
$m_2(\{4\})$	=	0.40
$m_2(\{1,2,3,4,5\})$	=	0.03

Table 3: Triplet mass function $\bar{m}_1^{(1)}$

$\bar{m}_1^{(1)}(\{1\})$	=	0.30
$\bar{m}_1^{(1)}(\{3\})$	=	0.30
$\bar{m}_1^{(1)}(\{1,2,3,4,5\})$	=	0.40

Table 4: Triplet mass function $\bar{m}_1^{(2)}$

$\bar{m}_1^{(2)}(\{1\})$	=	0.30
$\bar{m}_1^{(2)}(\{5\})$	=	0.30
$\bar{m}_1^{(2)}(\{1,2,3,4,5\})$	=	0.40

Table 5: Triplet mass function $\bar{m}_1^{(3)}$

$\bar{m}_1^{(3)}(\{3\})$	=	0.30
$\bar{m}_1^{(3)}(\{5\})$	=	0.30
$\bar{m}_1^{(3)}(\{1,2,3,4,5\})$	=	0.40

2-points focused combination method, we can achieve three possible results as shown in Tables 7, 8, and 9. We can observe that triplet mass function $\bar{m}^{(3)}$ is significantly different from triplet mass functions $\bar{m}^{(1)}$ and $\bar{m}^{(2)}$. Noticeably, the triplet mass function result of the combination of two mass functions m_1 and m_2 depends on the way we choose the triplet mass function regarding mass function m_1; therefore, 2-points focused combination method is not effective in this scenario.

3. Two-probabilities focused combination method

In RSs based on DST, let us consider a mass function $m : 2^\Theta \to [0,1]$ defined on a frame of discernment or a rating domain $\Theta = \{\theta_1, \theta_2, ..., \theta_L\}$. The focal set of mass function m is denoted by F. Clearly, the number of elements in focal set F is dominated by the number of elements in Θ; and F can contain maximum of $2^{|\Theta|}$ elements.

As mentioned previously, with Dempster's rule of combination, the focal elements with very low probabilities in F lead to not only poor performance in computational time when combining mass function m with another one, but also the unsatisfactory results. In addition, in the focal set F, especially when mass function m is a combination result, usually most of focal elements have infinitesimal probabilities whereas a few focal elements have high probabilities. Under such an observation, we suggest that, in focal set F, only focal elements with their probabilities in top two highest probabilities are retained, and other focal elements excluding the whole set one are treated as noise that may be caused due to superficial rating or resulting from the process of information fusion and then eliminated. Note that, the probabilities

Table 6: Triplet mass function $\bar{m}_2$

$\bar{m}_2(\{1\})$	=	0.40
$\bar{m}_2(\{4\})$	=	0.40
$\bar{m}_2(\{1,2,3,4,5\})$	=	0.20

Table 7: Triplet mass function $\bar{m}^{(1)}$

$\bar{m}^{(1)}(\{1\})$	=	0.53
$\bar{m}^{(1)}(\{4\})$	=	0.25
$\bar{m}^{(1)}(\{1,2,3,4,5\})$	=	0.22

Table 8: Triplet mass function $\bar{m}^{(2)}$

$\bar{m}^{(2)}(\{1\})$	=	0.53
$\bar{m}^{(2)}(\{4\})$	=	0.25
$\bar{m}^{(2)}(\{1,2,3,4,5\})$	=	0.22

Table 9: Triplet mass function $\bar{m}^{(3)}$

$\bar{m}^{(3)}(\{1\})$	=	0.31
$\bar{m}^{(3)}(\{4\})$	=	0.31
$\bar{m}^{(3)}(\{1,2,3,4,5\})$	=	0.38

of the eliminated focal elements are transferred to the whole set element in order to make sure that the achieved mass function is still well-defined.

Formally, assuming that $F' = F \backslash \Theta$ and F' contains n elements. After sorting all elements in F' by descending probabilities, we obtain $F' = \{A_1, A_2, ..., A_n\}$, where $m(A_i) = p_i$ with $A_i \subset \Theta$, and $p_1 \geq p_2 \geq p_3 \geq ... \geq p_n$. Based on mass function m, 2-probabilities focused mass function $\bar{m} : 2^\Theta \to [0,1]$ is defined as follows

$$\bar{m}(A) = \begin{cases} m(A), \text{ for } A \subset \Theta \text{ and } (m(A) = p_1 \text{ or } m(A) = p_2); \\ 1 - \sum\limits_{\{B \subset \Theta | m(B) = p_1\}} m(B) - \sum\limits_{\{C \subset \Theta | m(C) = p_2\}} m(C), \text{ if } A = \Theta; \\ 0, \text{ otherwise.} \end{cases} \tag{8}$$

Then, in RSs based on DST, users' preferences or ratings are represented by 2-probabilities mass functions instead of general mass functions.

Let us consider two 2-probabilities focused mass functions $\bar{m}_1$ and $\bar{m}_2$ defined on the same frame of discernment Θ. The method to combine these two 2-probabilities focused mass functions, denoted by $\bar{m} = \bar{m}_1 \uplus \bar{m}_2$ and called 2-probabilities focused combination, contains two steps as shown below

- Firstly, 2-probabilities focused mass functions $\bar{m}_1$ and $\bar{m}_2$ are combined by using Dempster's rule of combination regarding Equation (4). Let $\ddot{m}$ denotes the combination result after performing this step, we have $\ddot{m} = \bar{m}_1 \oplus \bar{m}_2$.

- Secondly, mass function $\ddot{m}$ is converted into corresponding 2-probabilities focused mass function $\bar{m}$ according to Equation (8).

Supposing that we need to combine n 2-probabilities focused mass functions, defined on the same frame of discernment Θ, by using 2-probabilities focused combination method. In the best case, when there is only a maximum of three focal elements in the focal sets of n 2-probabilities focused mass

Table 10: Mass function m'_1

$m'_1(\{1\})$	$=$	0.85
$m'_1(\{1,2\})$	$=$	0.12
$m'_1(\{2\})$	$=$	0.02
$m'_1(\{3\})$	$=$	0.01

Table 11: Mass function m'_2

$m'_2(\{3\})$	$=$	0.01
$m'_2(\{4\})$	$=$	0.05
$m'_2(\{5\})$	$=$	0.94

functions as well as the temporary computation result ones, 2-probabilities focused combination method will be the same as 2-points focused combination method. In addition, as remarked in [7], the time complexity of 2-points focused combination method is linear $O(n)$. Thus, it can be seen that the time complexity of the proposed method is greater or equal than that of 2-points focused combination.

In the worst case known as when the probabilities of the focal elements (excluding the whole set element) in the focal sets of both n 2-probabilities focused mass functions and the temporary computation result ones are in top two highest probabilities, there are no focal elements are eliminated. In this situation, the time complexity of 2-probabilities focused combination is the same as that of Dempster's rule of combination whose time complexity is exponential $O(|\Theta|^{n-1})$ [9]. Therefore, it also can be claimed that the time complexity of 2-probabilities focused combination is less or equal than that of Dempster's rule of combination.

Generally, in RSs based on DST, we can model users' preferences by t-probabilities focused mass functions as well as using t-probabilities focused combination method with t is an integer number ranging from 1 to $2^{|\Theta|} - 2$ for evidence combination.

In the rest of this section, three advantages and a disadvantage of the proposed combination method will be described.

3.1. Advantages

Firstly, 2-probabilities focused combination method helps RSs based on DST improve their computational time. It can be seen that, when reducing the number of focal elements in focal sets of two mass functions, logically, it takes less time to combine them together. Moreover, regarding the experimental results, the computational time of 2-probabilities focused combination method is somewhat worse than that of 2-points focused combination method whose time complexity is linear.

Table 12: 2-probabilities focused mass function $\bar{m}_1$

$\bar{m}'_1(\{1\})$	$=$	0.85
$\bar{m}'_1(\{1,2\})$	$=$	0.12
$\bar{m}'_1(\{1,2,3,4,5\})$	$=$	0.03

Table 13: 2-probabilities focused mass function $\bar{m}_2$

$\bar{m}'_2(\{4\})$	$=$	0.05
$\bar{m}'_2(\{5\})$	$=$	0.94
$\bar{m}'_2(\{1,2,3,4,5\})$	$=$	0.01

Table 14: 2-probabilities focused mass function $\bar{m}$

$\bar{m}'(\{1\})$	$=$	0.214105793
$\bar{m}'(\{5\})$	$=$	0.710327456
$\bar{m}'(\{1,2,3,4,5\})$	$=$	0.075566751

Secondly, with 2-probabilities focused combination method, the systems can void the unsatisfactory results. For example, in a RS based on DST with a rating domain $\Theta = \{1,2,3,4,5\}$, let consider two ratings represented as two mass functions shown in Tables 10 and 11. When combining these two mass functions by using Dempster's rule of combination, $m' = m'_1 \oplus m'_2$, we will get a unsatisfactory result, $m'(\{3\}) = 1$. With 2-probabilities focused combination method, two mass functions m'_1 and m'_2 are transformed into two 2-probabilities focused mass functions shown in Tables 12 and 13 respectively, and the combination result of these two 2-probabilities focused mass functions is more reasonable, as shown in Table 14. Note that, with 2-probabilities focused combination method, the unsatisfactory results are not completely eliminated.

Thirdly, 2-probabilities focused combination method is capable of dealing with the weakness of 2-points focused combination method due to non-uniqueness of triplet mass functions from a given mass function. It is seen that, from a given mass function, we can induce only one 2-probabilities focused mass function; thus, we get only one combination result when combining mass functions by using 2-probabilities focused combination method. Let us consider Example 1 again. Regarding mass function m_1, there is only one 2-probabilities focused mass function $\bar{m}_1$ as depicted in Table 15; and the 2-probabilities focused mass function $\bar{m}_2$ corresponding to mass function m_2 is shown in Table 16. Consequently, after combining two 2-probabilities focused mass functions $\bar{m}_1$ and $\bar{m}_2$ using 2-probabilities focused combination method, we achieve one combination result as illustrated in Table 17.

Table 15: 2-probabilities focused mass function $\bar{m}_1$

$\bar{m}_1(\{1\})$	=	0.30
$\bar{m}_1(\{3\})$	=	0.30
$\bar{m}_1(\{4\})$	=	0.04
$\bar{m}_1(\{5\})$	=	0.30
$\bar{m}_1(\{1,2,3,4,5\})$	=	0.06

Table 16: 2-probabilities focused mass function $\bar{m}_2$

$\bar{m}_2(\{1\})$	=	0.40
$\bar{m}_2(\{2\})$	=	0.10
$\bar{m}_2(\{4\})$	=	0.40
$\bar{m}_2(\{1,2,3,4,5\})$	=	0.10

Table 17: 2-probabilities focused mass function $\bar{m}$

$\bar{m}(\{1\})$	=	0.60
$\bar{m}(\{4\})$	=	0.15
$\bar{m}(\{1,2,3,4,5\})$	=	0.25

3.2. Disadvantage

However, 2-probabilities focused combination method is not associative. So as to evaluate the effect of this weakness on RSs based on DST, we have conducted the experiment with seventeen users, each of them belongs to four overlapping communities. As detailed in the results of this experiment in section 4.2, the combination results are just slightly influenced by the order of inputs.

4. Recommender systems with 2-probabilities focused combination

In this section, we will present about RSs based on DST, that employ 2-probabilities focused combination method for evidence combination. Almost all characteristics of these systems have been introduced in [34, 36, 50].

In the systems, a set of M users and a set containing N items are denoted by $\mathbf{U} = \{U_1, U_2, ..., U_M\}$ and $\mathbf{I} = \{I_1, I_2, ..., I_N\}$, respectively. Supposing that users can rate items with a rating domain containing L preference levels, denoted by $\Theta = \{\theta_1, \theta_2, ..., \theta_L\}$. A rating of user U_i on item I_k is denoted by $r_{i,k}$, and all ratings are represented by a rating matrix $\mathbf{R} = \{r_{i,k}\}$. Note that, originally $r_{i,k}$ is modeled as general mass function $m_{i,k}$; here, $r_{i,k}$ is represented by 2-probabilities focused mass function $\bar{m}_{i,k}$. In addition, let ${}^{I}R_i$ and ${}^{U}R_k$ denote the set of items rated by user U_i and the set of users rated item I_k, respectively.

Context information influencing users' preferences is defined as a set containing P concepts, denoted by $\mathcal{C} = \{C_1, C_2, ..., C_P\}$. And, each concept C_p, with $1 \leq p \leq P$, can consist of at most Q_p groups, that means

Figure 1: Context information

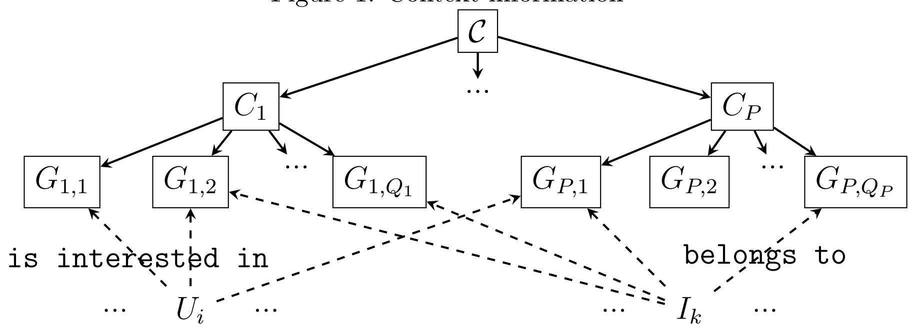
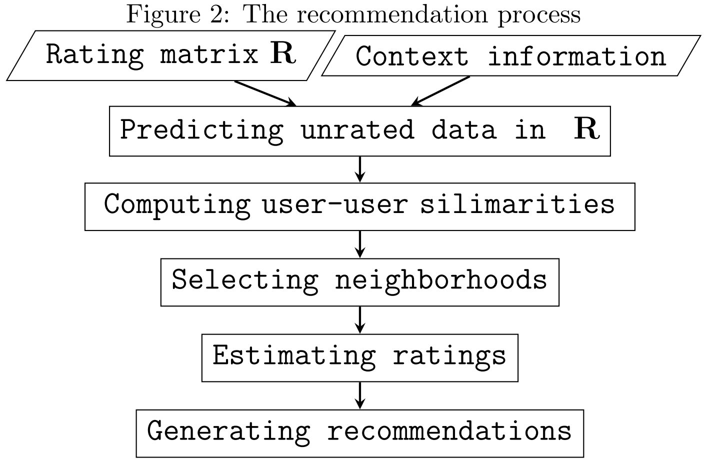


Figure 2: The recommendation process



$C_p = \{G_{p,1}, G_{p,2}, ..., G_{p,Q_p}\}$. Regarding concept C_p, item I_k can belong to some groups, and user U_i can also be interested in several groups, as shown in Figure 1 [36]. Assuming that the groups to which item I_k belongs and the groups in which user U_i is interested can be determined by functions g_p and f_p respectively. Formally, these functions are given by

$$\begin{aligned} g_p : \mathbf{I} &\to 2^{C_p} \\ I_k &\mapsto g_p(I_k) \subseteq C_p; \\ f_p : \mathbf{U} &\to 2^{C_p} \\ U_i &\mapsto f_p(U_i) \subseteq C_p. \end{aligned} \tag{9}$$

The general recommendation process of the RSs consists of 5 steps as illustrated in Figure 2. First, unrated data in rating matrix $\mathbf{R}$ are predicted by using context information. Then, user-user similarities are calculated by employing not only provided but also predicted ratings. Then, for an active user, a neighborhood set according to each unrated item is selected, and then the rating of this user on the item is estimated. Finally, the estimated ratings on unrated items are ranked, and suitable items are selected to recommend for the active user. Note that, in the system developed in [34], users are separated into overlapping communities and the first four steps of the recommendation process are independently applied into each community, after that, the finally estimated ratings are created by combining the estimated ratings in corresponding communities and recommendations are generated based on the finally estimated ratings. In the remainder of this section, details of steps of the recommendation process will be represented.

4.1. Predicting unrated data

The unrated data are mainly predicted based on the assumption that users who are interested in a group are expected to have the same preferences regarding that group. As mentioned previously an item I_k can belong to some groups of a concept C_p; and users' group preference on each group is necessary for generating unrated data on this item. Let consider a group $G_{p,q} \in C_p$ and $G_{p,q} \in g_p(I_k)$, the users' group preference on item I_k regarding this group is denoted by ${}^{G}\bar{m}_{p,q,k} : 2^{\Theta} \to [0,1]$. Each rating, $r_{j,k} = \bar{m}_{j,k}$, of user U_j, who is interested in group $G_{p,q}$, on item I_k is considered to be a piece of evidence of users' group preference on this item regarding group $G_{p,q}$. Thus, the users' group preference on the item regarding group $G_{p,q}$ can be computed by combining the corresponding pieces of evidence as follows

$$ {}^{G}\bar{m}_{p,q,k} = \biguplus_{\{j | I_k \in {}^{I}R_j, G_{p,q} \in f_p(U_j), G_{p,q} \in g_p(I_k)\}} \bar{m}_{j,k}. \tag{10} $$

Supposing that user U_i has not rated item I_k, the process to generate unprovided rating entry $r_{i,k}$ of this user on item I_k contains three steps as below

- Firstly, according to a concept C_p, for each group $G_{p,q} \in f_p(U_i) \cap g_p(I_k)$, users' group preference on item I_k regarding group $G_{p,q}$ is considered to be user U_i's group preference regarding this group as well as a piece of evidence of concept preference of this user regarding concept C_p. Consequently, user U_i's concept preference on item I_k regarding concept C_p, denoted by 2-probabilities focused mass functions ${}^{C}\bar{m}_{p,i,k} : 2^{\Theta} \to [0,1]$, can be computed by combining related users' U_i's group preferences on item I_k as below

$$ {}^{C}\bar{m}_{p,i,k} = \biguplus_{\{q|G_{p,q}\in f_p(U_i), G_{p,q}\in g_p(I_k)\}} {}^{G}\bar{m}_{p,q,k}. \tag{11} $$

- Secondly, if item I_k belongs to at least one group in concept C_p and user U_i is interested in at least one group in concept C_p then the concept preference of user U_i on item I_k regarding concept C_p is considered as a piece of evidence of user U_i's context preference on item I_k. Therefore, context preference of this user on item I_k, denoted by 2-probabilities focused mass function ${}^{\mathcal{C}}\bar{m}_{i,k} : 2^{\Theta} \to [0,1]$, is achieved as follows

$$ {}^{\mathcal{C}}\bar{m}_{i,k} = \biguplus_{p=\overline{1,P}} {}^{C}\bar{m}_{p,i,k}. \tag{12} $$

- Finally, U_i's context preference on item I_k is assigned to unrated entry $r_{i,k}$, as below

$$ r_{i,k} \;=\; {}^{\mathcal{C}}\bar{m}_{i,k}. \tag{13} $$

  Note that, in case the context information does not affect user U_i and item I_k, $\forall p, f_p(U_i) \cap g_p(I_k) = \emptyset$, unrated entry $r_{i,k}$ is assigned by the evidence obtained by combining all 2-probabilities focused mass functions of users who have rated item I_k [36] as follows

$$ r_{i,k} = \biguplus_{\{j|U_j\in {}^{U}R_k\}} \bar{m}_{j,k}. \tag{14} $$

Up to now, all unrated entries in the systems are predicted. Next both predicted and provided ratings will be employed to compute user-user similarities in the following step.

4.2. Computing user-user similarities

So as to measure the distance between users, the method developed in [15] is adopted. Additionally, based on this method, Wickramarathne et al. [50] have pointed out that the distance between two users U_i and U_j with $i \neq j$, denoted by $D(U_i, U_j)$, can be computed as shown below

$$ D(U_i, U_j) = \sum_{k=1}^{N} \left(\ln \max_{\theta\in\Theta} \frac{Bp_{j,k}(\theta)}{Bp_{i,k}(\theta)} - \ln \min_{\theta\in\Theta} \frac{Bp_{j,k}(\theta)}{Bp_{i,k}(\theta)} \right), \tag{15} $$

where $Bp_{i,k}$ and $Bp_{j,k}$ are the pignistic probability distributions corresponding the preference ratings of user U_i and user U_j on item I_k respectively. In [36], the authors also proposed a new method for computing the distance between two users U_i and U_j as follows

$$D(U_i, U_j) = \sum_{k=1}^{N} \mu(x_{i,k}, x_{j,k}) \Bigg(\ln \max_{\theta \in \Theta} \frac{Bp_{j,k}(\theta)}{Bp_{i,k}(\theta)} - \ln \min_{\theta \in \Theta} \frac{Bp_{j,k}(\theta)}{Bp_{i,k}(\theta)} \Bigg), \tag{16}$$

where $\mu(x_{i,k}, x_{j,k}) \in [0, 1]$ is a reliable function referring to the trust of the evaluation of both user U_i and user U_j on item I_k. Here, $x_{i,k} \in \{0, 1\}$ and $x_{j,k} \in \{0, 1\}$ equal to 1 if $r_{i,k}$ and $r_{j,k}$ are provided rating entries respectively; otherwise, $r_{i,k}$ and $r_{j,k}$ are predicted rating data. The function $\mu(x_{i,k}, x_{j,k})$ can be computed as follows

$$\mu(x_{i,k}, x_{j,k}) = 1 - w_1(x_{i,k} + x_{j,k}) - w_2 x_{i,k} x_{j,k}, \tag{17}$$

where w_1 and w_2 are the reliable coefficients [36].

Either Equation (15) or Equation (16) is selected to apply into the systems. Additionally, the user-user similarity between users U_i and U_j, denoted by $s_{i,j}$, is computed as follows

$$s_{i,j} = e^{-\gamma \times D(U_i, U_j)}, \text{ where } \gamma \in (0, \infty). \tag{18}$$

With the higher value of $s_{i,j}$, the user U_i is closer to user U_j. Finally, the user-user similarities among all users are represented in a matrix $\mathbf{S} = \{s_{i,j} \mid U_i \in \mathbf{U}, U_j \in \mathbf{U}, i \neq j\}$.

4.3. Selecting neighborhoods

Let consider an active user U_i, for each item I_k which has not been rated by this user, a K nearest neighborhood set $\mathcal{N}_{i,k}$ is selected by using the method proposed in [22]. According to this method, the selection process consists of two steps as shown below

- Firstly, a set of users who rated I_k and whose similarities with user U_i are equal or greater than a threshold τ is selected. This set is denoted by $\mathfrak{N}_{i,k}$ and obtained by the following equation

  $$\mathfrak{N}_{i,k} = \{U_j \in \mathbf{U} \mid I_k \in^I R_j, s_{i,j} \geq \tau\}. \tag{19}$$

  In this equation, when I_k is an item, the condition $I_k \in^I R_j$ needs to be removed.

- Secondly, all of members in $\mathfrak{N}_{i,k}$ is descending sorted by s_{ij} and top K members are selected as the neighborhood set $\mathcal{N}_{i,k}$.

4.4. Estimating ratings according to neighborhoods

The rating entries of active user U_i on all unrated items need to be estimated. Supposing that user U_i has not rated item I_k. Let $\hat{r}_{i,k}$ denotes the estimated rating entry of this user on item I_k. It can be seen that the RSs [34, 36, 50] belong to the collaborative filtering category [2]. Thus, the estimated rating $\hat{r}_{i,k}$ is computed based on ratings of user U_i's neighborhoods who are considered to have the similar taste with user U_i on item I_k. Formally, $\hat{r}_{i,k}$ is calculated as follows

$$\hat{r}_{i,k} = r_{i,k} \uplus \tilde{r}_{i,k}, \tag{20}$$

where $\tilde{r}_{i,k}$ is the 2-probabilities focused mass function corresponding to the overall preference of neighborhoods in the neighborhood set $\mathcal{N}_{i,k}$. Let us consider user $U_j \in \mathcal{N}_{i,k}$, and suppose that $s_{i,j}$ is the user-user similarity between user U_i and user U_j. The rating of user U_j on item I_k need to be discounted by a discount rate $1 - s_{i,j}$ [50]. As a result, $\tilde{r}_{i,k}$ is computed as below

$$\begin{aligned} \tilde{r}_{i,k} &= \biguplus_{\{j|U_j \in \mathcal{N}_{i,k}\}} \dot{r}^{s_{i,j}}_{j,k}, \\ \text{where } \dot{r}^{s_{i,j}}_{j,k} &= \begin{cases} s_{i,j} \times r_{j,k}(A), \text{for } A \subset \Theta; \\ s_{i,j} \times r_{j,k}(\Theta) + (1 - s_{i,j}), \text{if } A = \Theta. \end{cases} \end{aligned} \tag{21}$$

4.5. Generating recommendations

So as to generate recommendations for active user U_i, rating entries of this user on all unseen items are estimated, ranked, and then a suitable recommendation list is generated based on the ranked list. Especially, the RSs can offer both hard as well as soft decisions. To generate a hard decision, the pignistic probability is applied, and then the singleton with the highest probability will be selected as the preference label. On the other hand, for a soft decision, the maximum belief with overlapping interval strategy (maxBL) [10] is employed, and the singleton whose belief is greater than the plausibility of any other singleton will be chosen; note that, in case the class label does not exist, a decision will be made based on the favor of the composite class label constituted of the singleton label that has the maximum belief and those singletons having a higher plausibility [34, 36, 50].

5. Experiment

We conducted experiments on two RSs based on DST, that consist of characteristics as described in the previous section. The first system, similar to the system proposed in [36, 50], does not integrate with social networks. In contrast, the second one, the same as the system introduced in [34], is capable of integrating with community information extracted from the social network containing all users. Note that in these two systems, Equation (16) is employed to compute distances between two users.

To evaluate 2-probabilities focused combination method, 2-points focused combination method [5, 8, 9] were selected for the purpose of comparisons in both recommendation performances as well as computational time. In addition, to measure recommendation performances, evaluation methods *DS-MAE* was chosen. Let $\hat{r}_{i,k}$ denotes the estimated rating entry, which will be used for generating recommendations, of user U_i on item I_k; and $\widehat{Bp}_{i,k}$ denotes the pignistic probability distribution of $\hat{r}_{i,k}$. The selected evaluation method is defined as follows

$$\mathit{DS\text{-}MAE}(\theta_j) = \frac{1}{|\,D_j\,|} \sum_{\{(i,k)\in D_j, \theta_l \in \Theta\}} \widehat{Bp}_{i,k}(\theta_l)\ |\,\theta_j - \theta_l\,|, \tag{22}$$

where D_j is the testing set identifying the user-item pairs whose true rating is $\theta_j \in \Theta$.

Since the two systems work with domains with soft ratings, the method suggested in [50] was adopted for generating data sets in the experiments. Regarding this method, data sets with hard ratings are selected first, and then a DS modeling function is applied to transform the hard ratings into corresponding soft ratings. Here, Movielens and Flixster data sets were used in the first and the second systems, respectively. In these data sets, each hard rating, $\theta_l \in \Theta = \{\theta_1, \theta_2, ..., \theta_L\}$, of a user U_i on an item I_k was transformed into the corresponding soft rating which is presented by a 2-probabilities focused mass function $\bar{m}_{i,k}$ by using the DS modeling function [50] as below

$$\bar{m}_{i,k} = \begin{cases} \alpha_{i,k}(1-\sigma_{i,k}), & \text{for } A=\theta_l; \\ \alpha_{i,k}\sigma_{i,k}, & \text{for } A=B; \\ 1-\alpha_{i,k}, & \text{for } A=\Theta; \\ 0, & \text{otherwise,} \end{cases} \text{ with } B = \begin{cases} (\theta_1, \theta_2), & \text{if } l = 1; \\ (\theta_{L-1}, \theta_L), & \text{if } l = L; \\ (\theta_{l-1}, \theta_l, \theta_{l+1}), & \text{otherwise;} \end{cases} \tag{23}$$

where $\alpha_{i,k} \in [0, 1]$ and $\sigma_{i,k}$ are a trust factor and a dispersion factor, respectively [50].

Besides, in both Movielens and Flixster data sets, the information about the genres in which a user is interested is not available. Thus, we assume that if a user has rated an item then this user is interested in all genres to which the item belongs.

In the rest of this section, experiments on the first system with Movielens data set as well as those on the second system with Flixster data set are provided. Note that, the values of parameters in these systems are selected mainly based on the analyzed results published in [36, 50].

5.1. Experiment with Movielens data set

MovieLens 100K data set[2] was used in experiments. It contains 943 users and 100,000 hard ratings on 1682 movies with a rating domain containing 5 elements $\Theta = \{1, 2, 3, 4, 5\}$. Each user has rated at least 20 movies. Additionally, in the data set, context information, considered for grouping users, is represented as below

$$\begin{gathered}\mathcal{C} = \{Genre\};\\ Genre = \{Unknown, Action, Adventure, Animation, Children's,\\ Comedy, Crime, Documentary, Drama, Fantasy, Film\text{-}Noir, Horror,\\ Musical, Mystery, Romance, Sci\text{-}Fi, Thriller, War, Western\}.\end{gathered}$$

The values of parameters were selected as follows: $\gamma = 10^{-4}$, $w_1 = 0.3, w_2 = 0.1$, and $\forall(i,k)\{\alpha_{i,k}, \sigma_{i,k}\} = \{0.9, 2/9\}$. Particularly, it is unreasonable to select a fixed value for parameter τ to use in the experiments. The reason is that, with different combination methods, the values of user-user similarities of two specific users are different. Thus, to select value for parameter τ, all values in user-user similarity matrix $\mathbf{S}$ were sorted in ascending, and then, a value of $s_{i,j}$ that can retain top 30% of the highest values in $\mathbf{S}$ was chosen for τ.

Additionally, 10-fold cross validation was used in the experiments. Firstly, ratings in this data set were divided into 10 folds; each fold contains random 10% ratings of each user. Then, the experiments were conducted 10 times; in each time, one of 10 folds was selected as testing data and the remaining ratings were employed as training data. The average results of 10 times will be represented in the remainder of this section.

[2]http://grouplens.org/datasets/movielens/

Figure 3: Overall $DS-MAE$ vs. K (Movielens)

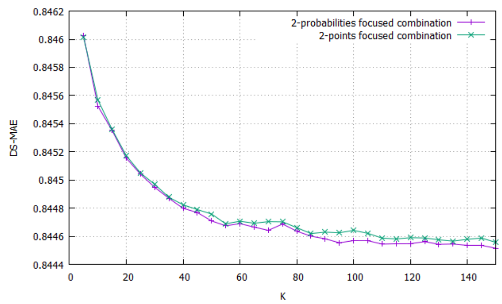


Figure 4: Overall computational time vs. K (Movielens)

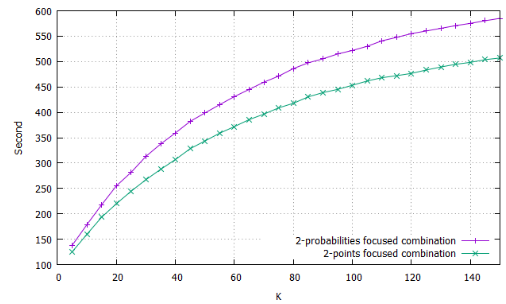


Figure 3 demonstrates overall *DS-MAE* criterion results changes with neighborhood size K. Note that, in this figure, the smaller values are the better ones. And, as can be seen in the figure, with $K \leq 40$ performances of the two methods increase sharply as well as being the same as each other. With $K > 40$, performances of both methods become stable; and especially, 2-probabilities focused combination is slightly better than 2-points focused combination method.

Execution time for the task of estimating ratings varies with neighborhood size K is depicted in Figure 4. As can be seen in this figure, the time taken by 2-probabilities focused is quite effective as well as being comparable to 2-points focused combination.

5.2. Experiment with Flixster data set

Flixster data set [34] consists of 3,827 users with 49,410 friend relationships, and 535,013 hard ratings on 1210 movies. In this data set, each user has rated at least 15 movies with a rating domain containing 10 elements denoted by $\Theta = \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0\}$. Additionally, all the genres considered as context information are represented as below

$$\begin{gathered}\mathcal{C} = \{Genre\};\\ Genre = \{Drama, Comedy, Action\ \&\ Adventure, Television,\\ Mystery\ \&\ Suspense, Horror, ScienceFiction\ \&\ Fantasy,\\ Kids\ \&\ Family, Art\ House\ \&\ International, Romance, Classics,\\ Musical\ \&\ Performing\ Arts, Anime\ \&\ Manga, Animation, Western,\\ Documentary, Special\ Interest, Sports\ \&\ Fitness, Cult\ Movies\}.\end{gathered}$$

All users belong to a social network whose nodes are linked by undirected friendships. In addition, so as to discover overlapping communities in this

Table 18: Overlapping communities in Flixster data set

Community IDs.	Total number of users
16	226
49	377
50	2749
86	712
90	1011
113	460
147	105

Figure 5: Overall $DS-MAE$ vs. K (Flixster)

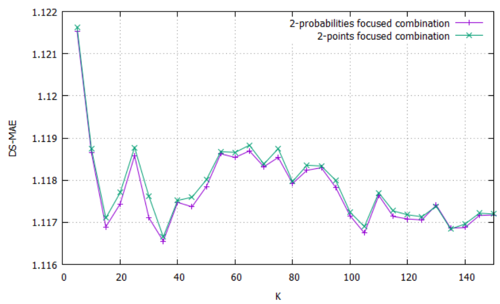
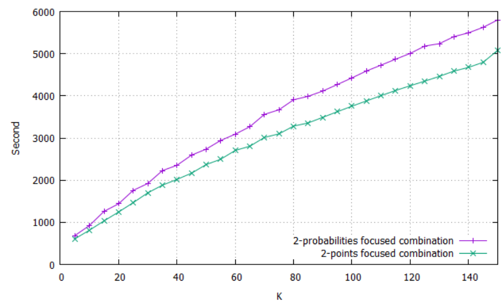


Figure 6: Overall computational time vs. K (Flixster)



social network, Speaker-Listener Label Propagation algorithm (SLPA) [51] was selected; the reason is that this algorithm is capable of not only identifying overlapping communities with the time complexity scaling linearly with the number of edges but also helping to avoid producing small size communities. After executing the SLPA algorithm, 7 overlapping communities were detected and they are depicted in Table 18.

The rating matrix containing all rating data in the Flixster data set was divided into 7 sub-rating matrices according to 7 communities. Each sub-rating matrix consists of the ratings of members in the corresponding community. After that, tasks of predicting unrated data, computing user-user similarities, selecting neighborhood, and estimating rating data were performed in each community independently. The finally estimated rating data for an active user were generated by combining all estimated rating data of this user in the communities to which he/she belongs. The suitable recommendations will be generated based on the finally estimated rating data.

The values of parameters were selected as follows: $\gamma = 10^{-4}$, $w_1 = 0.3, w_2 = 0.1$ and $\forall(i,k)\{\alpha_{i,k}, \sigma_{i,k}\} = \{0.9, 2/9\}$. To choose the value for

Table 19: Users belonging to four overlapping communities

User IDs	Community IDs						
	16	49	50	86	90	113	147
90			✓	✓	✓	✓	
206			✓	✓	✓	✓	
601		✓	✓		✓	✓	
1106		✓	✓		✓	✓	
1523			✓	✓	✓	✓	
1611			✓	✓	✓	✓	
1820		✓	✓		✓	✓	
2302			✓	✓	✓	✓	
2441			✓	✓	✓	✓	
2523		✓	✓		✓	✓	
2825			✓	✓	✓	✓	
3012		✓	✓		✓	✓	
3021		✓	✓		✓	✓	
3024		✓	✓		✓	✓	
3061			✓	✓	✓	✓	
3282			✓	✓	✓	✓	
3481			✓	✓	✓	✓	

parameter τ, all values in user-user similarity matrix $\mathbf{S}$ were sorted in ascending, and then, a value of $s_{i,j}$ that can retain top 50% of the highest values in $\mathbf{S}$. In addition, this data set was separated into two parts, testing data and training data; the first one contains random 5 ratings of each user, and the other consists of the remaining ratings.

Overall *DS-MAE* criterion results varies with neighborhood size K is depicted in Figure 5 . This figure shows that the performances of both combination methods are similar to each other and rise sharply when $K \leq 15$; with K in between 15 and 110, the performances are fluctuated; and then become quite stable when $K > 110$. As observed in this feature, regarding recommendation accuracy, 2-probabilities focused combination is slightly better than 2-points focused combination.

The computation time for the task of estimating ratings changes with neighborhood size K is depicted in Figure 6. As observed, execution time of 2-probabilities focused combination is somewhat worse than but comparable to that of 2-points focused combination. Additionally, this result is consistent with the result illustrated in Figure 4.

To evaluate the weakness of 2-probabilities focused combination method, an experiment was conducted as follows. Seventeen users, each of them belongs to 4 communities concurrently, were selected; and these users as well as their corresponding communities are shown in Table 19. For each user, the estimated ratings on an item in his/her communities are considered as pieces of evidence of the finally estimated rating on the item. Thus, the finally estimated rating is generated by combining corresponding 4 pieces of evidence by using this method. There are 24 combinations of the inputs when combing 4 pieces of evidence. The performances of recommendations regarding 24 combinations were evaluated by using *DS-MAE* evaluation criterion; and the results with $K = 45$ are illustrated in Tables 20 and 21.

In Tables 20 and 21, each column presents the overall *DS-MAE* for one user; and μ and *SD* are means and standard deviations of overall *DS-MAE* over 24 combinations respectively. As observed, the standard deviations are very small (SD is smaller than 0.01 for 4 users, in between 0.01 and 0.1 for 12 users, and about 0.1059 for one user). That means, in the RSs, when combining information by using 2-probabilities focused combination, input order is just minor affected the combination results.

6. Conclusion remarks

Comparing to traditional RSs, the RSs based on DST can have two advantages. The first one is the ability to model users' preferences with uncertain, imprecise and incomplete information. The second advantage is that information about users' preferences from different sources can be combined together easily. It can also be seen that, in these systems, information combination tasks are performed very often and currently Dempster's rule of combination is used in almost all cases. However, when combining information with this rule of combination, focal elements with very low probabilities cause the poor performance in computational time and the unsatisfactory results. With the new method developed in this paper, called 2-probabilities focused combination, these issues are tackled. Additionally, when comparing to an alternative combination method, known as 2-points focused combination, the new method can be comparable in computational time; especially, regarding recommendation accuracy, the new method is more effective because of better and stable combination results.

Table 20: *DS-MAE* varies with twenty four combinations

No.	User IDs								
	90	206	601	1106	1523	1611	1820	2302	2441
1	1.20663	0.78436	1.19467	0.67151	1.09809	0.76991	1.19449	0.51485	1.13954
2	1.15631	0.77841	1.19384	0.88947	1.09985	0.77171	1.11603	0.52434	1.13692
3	1.18254	0.83257	1.19132	0.67529	1.10304	0.77412	1.17816	0.65597	1.13778
4	1.06767	0.74446	1.18026	0.89375	1.10540	0.76072	1.10144	0.71870	1.18033
5	1.04333	0.71440	1.18976	0.89158	1.10707	0.76905	1.12229	0.55391	1.16420
6	1.04686	0.72602	1.17988	0.89321	1.10642	0.75059	1.09904	0.58141	1.12196
7	1.12474	0.73366	1.18902	0.89312	1.10520	0.85901	1.04451	0.81300	1.18558
8	1.13057	0.70265	1.18832	0.89324	1.11029	0.76606	1.04080	0.69762	1.17637
9	1.12119	0.76086	1.18543	0.89281	1.10424	0.84659	1.12487	0.82616	1.17124
10	1.06767	0.74446	1.18026	0.89375	1.10540	0.76072	1.10144	0.71870	1.18033
11	1.05450	0.70331	1.18430	0.89323	1.10712	0.75395	1.10842	0.58004	1.15608
12	1.04686	0.72602	1.17988	0.89321	1.10642	0.75059	1.09904	0.58141	1.12196
13	1.14624	0.68937	1.19437	0.89390	1.10560	0.79425	1.04196	0.60060	1.14811
14	1.13057	0.70265	1.18832	0.89324	1.11029	0.76606	1.04080	0.69762	1.17637
15	1.13490	0.79629	1.19928	0.84851	1.09768	0.76911	1.06677	0.54010	1.13746
16	1.15631	0.77841	1.19384	0.88947	1.09985	0.77171	1.11603	0.52434	1.13692
17	1.05450	0.70331	1.18430	0.89323	1.10712	0.75395	1.10842	0.58004	1.15608
18	1.04333	0.71440	1.18976	0.89158	1.10707	0.76905	1.12229	0.55391	1.16420
19	1.14624	0.68937	1.19437	0.89390	1.10560	0.79425	1.04196	0.60060	1.14811
20	1.12474	0.73366	1.18902	0.89312	1.10520	0.85901	1.04451	0.81300	1.18558
21	1.13490	0.79629	1.19928	0.84851	1.09768	0.76911	1.06677	0.54010	1.13746
22	1.20663	0.78436	1.19467	0.67151	1.09809	0.76991	1.19449	0.51485	1.13954
23	1.12119	0.76086	1.18543	0.89281	1.10424	0.84659	1.12487	0.82616	1.17124
24	1.18254	0.83257	1.19132	0.67529	1.10304	0.77412	1.17816	0.65597	1.13778
μ	1.11796	0.74720	1.18920	0.85247	1.10417	0.78209	1.10323	0.63389	1.15463
SD	0.05274	0.04295	0.00583	0.08275	0.00377	0.03411	0.04882	0.10588	0.02020

Table 21: *DS-MAE* varies with twenty four combinations

No.	User IDs							
	2523	2825	3012	3021	3024	3061	3282	3481
1	1.54115	1.13978	1.50942	1.29970	0.97805	0.58283	1.93319	1.09029
2	1.53108	1.08298	1.49625	1.29963	0.94727	0.64155	1.92527	1.09210
3	1.53921	1.10816	1.43177	1.29970	0.93814	0.57665	1.88074	1.06980
4	1.52362	1.06260	1.45589	1.29946	0.93895	0.70653	1.81822	1.07599
5	1.53234	0.95396	1.49751	1.29937	0.95869	0.75097	1.89421	1.02075
6	1.51809	0.99396	1.46803	1.29937	0.94765	0.74716	1.80553	1.00148
7	1.53751	1.01736	1.53441	1.29946	1.00114	0.51073	1.83120	1.29152
8	1.53106	1.03701	1.52828	1.29947	1.00664	0.70271	1.83379	1.29183
9	1.53389	1.10749	1.43813	1.29951	0.92185	0.62258	1.80976	1.13714
10	1.52362	1.06260	1.45589	1.29946	0.93895	0.70653	1.81822	1.07599
11	1.53366	1.00921	1.49699	1.29940	0.97276	0.74295	1.84090	0.99981
12	1.51809	0.99396	1.46803	1.29937	0.94765	0.74716	1.80553	1.00148
13	1.54034	1.03765	1.59069	1.29957	1.00905	0.56031	1.87270	1.15649
14	1.53106	1.03701	1.52828	1.29947	1.00664	0.70271	1.83379	1.29183
15	1.54321	1.05644	1.58039	1.29975	1.00663	0.56641	1.92080	1.09462
16	1.53108	1.08298	1.49625	1.29963	0.94727	0.64155	1.92527	1.09210
17	1.53366	1.00921	1.49699	1.29940	0.97276	0.74295	1.84090	0.99981
18	1.53234	0.95396	1.49751	1.29937	0.95869	0.75097	1.89421	1.02075
19	1.54034	1.03765	1.59069	1.29957	1.00905	0.56031	1.87270	1.15649
20	1.53751	1.01736	1.53441	1.29946	1.00114	0.51073	1.83120	1.29152
21	1.54321	1.05644	1.58039	1.29975	1.00663	0.56641	1.92080	1.09462
22	1.54115	1.13978	1.50942	1.29970	0.97805	0.58283	1.93319	1.09029
23	1.53389	1.10749	1.43813	1.29951	0.92185	0.62258	1.80976	1.13714
24	1.53921	1.10816	1.43177	1.29970	0.93814	0.57665	1.88074	1.06980
μ	1.53376	1.05055	1.50231	1.29953	0.96890	0.64261	1.86386	1.11015
SD	0.00720	0.05229	0.04950	0.00013	0.03047	0.08275	0.04570	0.09533

References

[1] Adomavicius, G., Mobasher, B., Ricci, F., and Tuzhilin, A. (2011). Context-aware recommender systems. *AI Mag.*, 32(3):67–80.

[2] Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. Knowl. Data Eng.*, 17(6):734–749.

[3] Al-hassan, M., Lu, H., and Lu, J. (2015). A semantic enhanced hybrid recommendation approach: A case study of e-government tourism service recommendation system. *Decis. Support Syst.*, 72:97–109.

[4] Barnett, J. A. (1981). Computational methods for a mathematical theory of evidence. In *IJCAI 1981*, pages 868–875.

[5] Bell, D. A., Guan, J. W., and Bi, Y. (2005). On combining classifier mass functions for text categorization. *IEEE Trans. Knowl. Data Eng.*, 17(10):1307–1319.

[6] Bernabé-Moreno, J., Tejeda-Lorente, Á., Porcel, C., Fujita, H., and Herrera-Viedma, E. (2015). CARESOME: A system to enrich marketing customers acquisition and retention campaigns using social media information. *Knowl.-Based Syst.*, 80:163–179.

[7] Bi, Y. (2008). An efficient triplet-based algorithm for evidential reasoning. *Int. J. Intell. Syst.*, 23(4):483–516.

[8] Bi, Y. (2012). The impact of diversity on the accuracy of evidential classifier ensembles. *Int. J. Approx. Reasoning*, 53(4):584–607.

[9] Bi, Y., Guan, J., and Bell, D. A. (2008). The combination of multiple classifiers using an evidential reasoning approach. *Artif. Intell.*, 172(15):1731–1751.

[10] Bloch, I. (1996). Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recogn. Lett.*, 17(8):905–919.

[11] Bobadilla, J., Ortega, F., Hernando, A., and Gutiérrez, A. (2013). Recommender systems survey. *Knowl.-Based Syst.*, 46:109–132.

[12] Bobadilla, J., Serradilla, F., and Hernando, A. (2009). Collaborative filtering adapted to recommender systems of e-learning. *Knowl.-Based Syst.*, 22(4):261–265.

[13] Cao, N., Shi, C., Lin, W. S., Lu, J., Lin, Y., and Lin, C. (2016). Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Trans. Vis. Comput. Graph.*, 22(1):280–289.

[14] Champiri, Z. D., Shahamiri, S. R., and Salim, S. S. B. (2015). A systematic review of scholar context-aware recommender systems. *Expert Syst. Appl.*, 42(3):1743–1758.

[15] Chan, H. and Darwiche, A. (2005). A distance measure for bounding probabilistic belief change. *Int. J. Approx. Reasoning*, 38(2):149–174.

[16] Colombo-Mendoza, L. O., Valencia-García, R., González, A. R., Alor-Hernández, G., and Samper-Zapater, J. J. (2015). Recommetz: A context-aware knowledge-based mobile recommender system for movie showtimes. *Expert Syst. Appl.*, 42(3):1202–1222.

[17] Costa-Montenegro, E., Barragáns-Martínez, A. B., and Rey-López, M. (2012). Which app? A recommender system of applications in markets: Implementation of the service for monitoring users' interaction. *Expert Syst. Appl.*, 39(10):9367–9375.

[18] Damiani, E., Ceravolo, P., Frati, F., Bellandi, V., Maier, R., Seeber, I., and Waldhart, G. (2015). Applying recommender systems in collaboration environments. *Comput. Hum. Behav.*, 51:1124–1133.

[19] Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.*, 38:325–339.

[20] Denoeux, T. (2013). Maximum likelihood estimation from uncertain data in the belief function framework. *IEEE Trans. Knowl. Data Eng.*, 25(1):119–130.

[21] Denoeux, T. and Masson, M. (2012). Evidential reasoning in large partially ordered sets - application to multi-label classification, ensemble clustering and preference aggregation. *Ann. Oper. Res.*, 195(1):135–161.

[22] Herlocker, J. L., Konstan, J. A., Borchers, A., and Riedl, J. (1999). An algorithmic framework for performing collaborative filtering. In *SIGIR '99*, pages 230–237. ACM.

[23] Huynh, V., Tri, N. T., and Le, C. A. (2010). Adaptively entropy-based weighting classifiers in combination using dempster-shafer theory for word sense disambiguation. *Comput. Speech Lang.*, 24(3):461–473.

[24] Iglesias, J., Bernardos, A. M., and Casar, J. R. (2013). An evidential and context-aware recommendation strategy to enhance interactions with smart spaces. In *HAIS 2013*, pages 242–251.

[25] Kanjanatarakul, O., Sriboonchitta, S., and Denoeux, T. (2014). Forecasting using belief functions: An application to marketing econometrics. *Int. J. Approx. Reasoning*, 55(5):1113–1128.

[26] Langseth, H. and Nielsen, T. D. (2012). A latent model for collaborative filtering. *Int. J. Approx. Reasoning*, 53(4):447–466.

[27] Lee, S. K., Cho, Y. H., and Kim, S. H. (2010). Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations. *Inf. Sci.*, 180(11):2142–2155.

[28] Li, Y., Chou, C., and Lin, L. (2014). A social recommender mechanism for location-based group commerce. *Inf. Sci.*, 274:125–142.

[29] Linden, G., Smith, B., and York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Comput.*, 7(1):76–80.

[30] Lu, J., Wu, D., Mao, M., Wang, W., and Zhang, G. (2015). Recommender system application developments: A survey. *Decis. Support Syst.*, 74:12–32.

[31] Martínez-Cruz, C., Porcel, C., Bernabé-Moreno, J., and Herrera-Viedma, E. (2015). A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling. *Inf. Sci.*, 311:102–118.

[32] Masson, M. and Denoeux, T. (2011). Ensemble clustering in the belief functions framework. *Int. J. Approx. Reasoning*, 52(1):92–109.

[33] Mohanty, B. K. and Passi, K. (2010). Agent based e-commerce systems that react to buyers' feedbacks - a fuzzy approach. *Int. J. Approx. Reasoning*, 51(8):948–963.

[34] Nguyen, V.-D. and Huynh, V.-N. (2014). A community-based collaborative filtering system dealing with sparsity problem and data imperfections. In *PRICAI 2014*, pages 884–890.

[35] Nguyen, V.-D. and Huynh, V.-N. (2015a). Evidence combination focusing on significant focal elements for recommender systems. In *IUKM 2015*, pages 290–302.

[36] Nguyen, V.-D. and Huynh, V.-N. (2015b). A reliably weighted collaborative filtering system. In *ECSQARU 2015*, pages 429–439.

[37] Núñez-Valdéz, E. R., Lovelle, J. M. C., Martínez, O. S., García-Díaz, V., de Pablos, P. O., and Marín, C. E. M. (2012). Implicit feedback techniques on recommender systems applied to electronic books. *Comput. Hum. Behav.*, 28(4):1186–1193.

[38] Quost, B., Masson, M., and Denoeux, T. (2011). Classifier fusion in the Dempster-Shafer framework using optimized t-norm based combination rules. *Int. J. Approx. Reasoning*, 52(3):353–374.

[39] Rashid, A. M., Karypis, G., and Riedl, J. (2008). Learning preferences of new users in recommender systems: An information theoretic approach. *SIGKDD Explor.*, 10(2):90–100.

[40] Ricci, F., Rokach, L., Shapira, B., and Kantor, P. B., editors (2011). *Recommender Systems Handbook*. Springer.

[41] Schafer, J. B., Konstan, J. A., and Riedl, J. (2001). E-commerce recommendation applications. *Data Min. Knowl. Discov.*, 5(1/2):115–153.

[42] Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press.

[43] Shambour, Q. and Lu, J. (2011). A hybrid trust-enhanced collaborative filtering recommendation approach for personalized government-to-business e-services. *Int. J. Intell. Syst.*, 26(9):814–843.

[44] Shambour, Q. and Lu, J. (2015). An effective recommender system by unifying user and item trust information for B2B applications. *J. Comput. Syst. Sci.*, 81(7):1110–1126.

[45] Smets, P. (1999). Practical uses of belief functions. In *UAI 1999*, pages 612–621. Morgan Kaufmann Publishers Inc.

[46] Smets, P. (2007). Analyzing the combination of conflicting belief functions. *Inf. Fusion*, 8(4):387–412.

[47] Smets, P. and Kennes, R. (1994). The transferable belief model. *Artif. Intell.*, 66(2):191–234.

[48] Troiano, L., Rodríguez-Muñiz, L. J., and Díaz, I. (2015). Discovering user preferences using Dempster-Shafer theory. *Fuzzy Set. Syst.*, 278:98–117.

[49] Vozalis, M. G. and Margaritis, K. G. (2007). Using SVD and demographic data for the enhancement of generalized collaborative filtering. *Inf. Sci.*, 177(15):3017–3037.

[50] Wickramarathne, T. L., Premaratne, K., Kubat, M., and Jayaweera, D. T. (2011). CoFiDS: A belief-theoretic approach for automated collaborative filtering. *IEEE Trans. Knowl. Data Eng.*, 23(2):175–189.

[51] Xie, J. and Szymanski, B. K. (2012). Towards linear time overlapping community detection in social networks. In *PAKDD 2012*, pages 25–36.

[52] Zadeh, L. A. (1984). Book review: A mathematical theory of evidence. *AI Mag.*, 5(3):81–83.