Combining conflicting evidences based on Pearson correlation coefficient and weighted graph

Jixiang Deng[a], Yong Deng[a,b,c,*]

[a]*Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, China*
[b]*School of Education, Shannxi Normal University, Xi'an, 710062, China*
[c]*Department of Biostatics, Medical Center, Vanderbilt University, Nashville, TN. 37235, USA*

Abstract

Dempster-Shafer evidence theory (evidence theory) has been widely used for its great performance of dealing with uncertainty. Based on evidence theory, researchers have presented different methods to combine evidences. Dempster's rule is the most well-known combination method, which has been applied in many fields. However, Dempster's rule may yield counter-intuitive results when evidences are in high conflict. To improve the performance of combining conflicting evidences, in this paper, we present a new evidence combination method based on Pearson correlation coefficient and weighted graph. The proposed method can correctly identify the target with a high accuracy. Besides, the proposed method has a better performance of convergence compared with other combination methods. In addition, the weighted graph generated by the proposed method can directly represent the relation of different evidences, which can help researchers to determine the reliability of every evidence. Moreover, an experiment is expounded to show the efficiency of the proposed method, and the results are analyzed and discussed.

Keywords: Dempster-Shafer evidence theory, conflicting evidences combination, Pearson correlation coefficient, weighted graph, target recognition.

*Corresponding author
Email address: `dengentropy@uestc.edu.cn; prof.deng@hotmail.com` (Yong Deng)

1. Introduction

In the past decades, plenty of theories have been developed for expressing and dealing with the uncertainty in the uncertain environment, for instance, probability theory [1], fuzzy set theory [2], Dempster-Shafer evidence theory [3, 4], rough sets [5], and D numbers [6].

Dempster-Shafer evidence theory (evidence theory) has been widely applied in many fields, like uncertainty measurements [7, 8, 9, 10], data fusion[11, 12, 13, 14], decision making [15], complex networks[16, 17, 18], and so on[19, 20]. In evidence theory, Dempster's combination rule is a widely used method for combining different evidences. However, there are some issues about Dempster's combination rule, especially its combination result of conflicting evidence may be illogical. To solve this problem, a lot of improved evidence combination methods have been developed. Murphy [21] proposed a modified combination method by averaging the BPA of every evidence. Then, Deng [22] improved Murphy's method by weighted averaging BPA based on distance of evidence. Taking independent degree as a discounting factor, Yager [23] proposed a improved combination method of belief function. Some other works were also presented[24, 25, 26]. Most of these evidence combinations use the technique of averaging BPA to reduce the influence of conflicting evidences, namely, averaging BPA based on the reliability of every evidence, which can be calculated by relative entropy[11], similarity[27, 28], distance[22] and so on[29].

In statistics, Pearson correlation coefficient [30] is a linear correlation coefficient for measuring the relationship, or association, of two variables, which is developed by Karl Pearson with wide applications [31, 32]. Since Pearson correlation coefficient is based on the covariance, it can be introduced into evidence theory to calculate the reliability of different evidence. Inspired on this, Xu [33] proposed a method based on shearman coefficient and pearson coefficient, which shows a good accuracy of recognizing objects. Nevertheless, Xu's method can not directly reflect which evidences are in conflict. Moreover, its accuracy can be improved.

In discrete mathematics, graph theory is one of the prime research field, and graph is a useful mathematical tool for directly modeling relations between objects. In a certain graph, the objects can be represented by nodes and linked by edges. If the edges have sense of direction, the graph is called directed graph; if not, the graph is called undirected graph. Because of the good performance of representing objects, graphs can be used to abstract many problems, and have been successfully applied in real practice. For example, the nervous system can be abstracted as a graph, where nerve cells and nerve fibers can be respectively represented by nodes and edges [34]. The social relationships of people can also be abstracted by graphs[35].

Recently, based on graph and complex network, a new technique of identifying conflicting evidences is proposed[18], which provides a feasible way to solve the issues of evidence theory with the help of graphs. Based on this technique, Liu proposes a new evidence combination method[26], which shows a great performance of combining conflicting evidences. However, Liu's method does not use averaging technique, and it just uses simple graph to identify conflicting evidence, which can be further modified to enhance the performance of combining evidence in conflict.

As a result, in this paper, considering the problems mentioned above, we propose a new evidence combination method based on Pearson correlation coefficient and weighted graph, which can combine evidences in conflict and correctly recognize the alternative with a high accuracy. Besides, the performance of convergence of the proposed method is better than other common methods. In addition, the proposed method can generates a weighted graph to illustrate the relation of different evidences, which can directly show the reliability of every evidence.

To summarize, the major contributions of this paper are as follows:

(1) A new evidence combination method is proposed based on Pearson correlation coefficient and weighted graph, which can combine evidence in high

conflict and accurately recognize the correct target.

(2) The proposed method has a good performance of convergence, which can better fit the situation in real practice compared with other common methods.

(3) The weighted graph generated by the proposed method can directly show the relationship of different evidences, which can be used to determine the reliability of evidences and identify conflicting evidences

The rest of this paper is organized as follows. In section **2**, some preliminaries are briefly reviewed. In section **3**, based on Pearson correlation coefficient and weighted graph, a new evidence combination method is proposed. In section **4**, an experiment are expounded to illustrate the proposed method. In section **5**, the results of the experiment are discussed. In section **6**, we have a brief conclusion.

In this section, some preliminaries are briefly introduced including Dempster-Shafer evidence theory, Pearson correlation coefficient and graph theory.

1.1. Dempster-Shafer evidence theory

Dempster-Shafer evidence theory[3, 4] can be used to deal with uncertainty. Besides, evidence theory satisfies the weaker conditions than the probability theory, which provides it with the ability to express uncertain information directly. Some basic conceptions of evidence theory are given as follows:

Definition 2.1: *Frame of discernment and its power set*

Let Θ, called the frame of discernment, denote an exhaustive nonempty set of hypotheses, where the elements are mutually exclusive. Let the set Θ have N elements, which can be expressed as:

$$\Theta = \{\theta_1, \theta_2, \theta_3, \cdots, \theta_N\} \tag{1}$$

The power set of Θ, denoted as 2^Θ, contains all possible subsets of Θ and *has 2^N elements, and 2^Θ is represented by*

$$\begin{aligned} 2^\Theta &= \{A_1, A_2, A_3, \cdots, A_{2^N}\} \\ &= \{\, \emptyset, \{\theta_1\}, \{\theta_2\}, \cdots, \{\theta_N\}, \{\theta_1, \theta_2\}, \\ &\qquad \{\theta_1, \theta_3\}, \cdots, \{\theta_1, \theta_N\}, \cdots, \Theta \,\} \end{aligned} \tag{2}$$

where the element A_i is called the focal element of Θ, if A_i is nonempty.

Definition 2.2: *Basic probability assignment (BPA)*

A BPA is a mass function mapping m from 2^Θ to $[0,1]$, and it is defined as follows:

$$m : 2^\Theta \to [0,1] \tag{3}$$

which is constrained by the following conditions:

$$\sum_{A \in 2^\Theta} m(A) = 1 \tag{4}$$

$$m(\emptyset) = 0 \tag{5}$$

Definition 2.3: *Dempster's rule of combination*

Given two BPAs m_1 and m_2 from two different evidence sources, the Dempster rule of combination, or the orthogonal sum of m_1 and m_2, is defined as:

$$\begin{cases} m(A) = \frac{\sum_{B \cap C = A} m_1(B) \cdot m_2(C)}{1 - K(m_1, m_2)} & A \neq \emptyset \\ m(\emptyset) = 0 \end{cases} \tag{6}$$

where $K(m_1, m_2)$ is the degree of conflict between m_1 and m_2, and it is defined as follows:

$$K(m_1, m_2) = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C). \tag{7}$$

It worth noting that Dempster's rule of combination can only be used to combine such two BPAs, when $0<K(m_1, m_2)<1$.

1.2. Pearson correlation coefficient

Pearson correlation coefficient is a linear correlation coefficient, which can represent the linear correlation of two variables. The definition of Pearson correlation coefficient is as follows[30]:

Definition 2.4: *Pearson correlation coefficient*

Assume two samples X and Y which can be denoted as vectors: $\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{Y}}$. Each sample contains N sample observations which can be denoted as the components of the vectors, namely, $\overrightarrow{\boldsymbol{X}} = (x_1, x_2, \ldots, x_N)$ and $\overrightarrow{\boldsymbol{Y}} = (y_1, y_2, \ldots, y_N)$. Then the Pearson correlation coefficient of $\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{Y}}$ is defined as:

$$r_{\overrightarrow{\boldsymbol{X}}\overrightarrow{\boldsymbol{Y}}} = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sqrt{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}\sqrt{N\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2}} \tag{8}$$

The main properties of Pearson correlation coefficient is that:

(1) The value range of $r_{\overrightarrow{\boldsymbol{X}}\overrightarrow{\boldsymbol{Y}}}$ is $[-1, 1]$.

(2) If $r_{\overrightarrow{\boldsymbol{X}}\overrightarrow{\boldsymbol{Y}}} > 0$, the relation between $\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{Y}}$ is positive correlation.

(3) If $r_{\overrightarrow{\boldsymbol{X}}\overrightarrow{\boldsymbol{Y}}} < 0$, the relation between $\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{Y}}$ is negative correlation.

(4) If $r_{\overrightarrow{\boldsymbol{X}}\overrightarrow{\boldsymbol{Y}}} = 0$, the linear correlation of $\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{Y}}$ is not obvious.

(5) The greater $\left|r_{\overrightarrow{\boldsymbol{X}}\overrightarrow{\boldsymbol{Y}}}\right|$ is, the higher linear correlation rate of $\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{Y}}$ will be.

1.3. Graph theory

In graph theory, the graph is a useful mathematical tool for dealing with the relationships among objects. Some basic conceptions of graph theory are listed as follows[36]:

Definition 2.5: *Weighted graph*

A weighted graph is defined as $G = (V, E, W)$, where $V = \{v_1, v_2, ..., v_N\}$ is called the node set whose element v_i is node, $E = \{\{v_i, v_j\}|\{v_i, v_j\} \in V \wedge V\}$ is called the edge set whose element $\{v_i, v_j\}$ is edge which connects two nodes v_i and v_j, and $W = \{w_{ij}|i, j = 1, ..., N\}$ is called the weight set whose element w_{ij} is the weight assigned to the edge $\{v_i, v_j\}$.

It should be noted that the weight of the weighted graph describes the relationship between two nodes, such as distance, time, similarity, costs, et al.

Definition 2.6: *Adjacency matrix of weighted graph*

The adjacency matrix A of a weighted graph $G = (V, E, W)$ is defined as a $|V| * |V|$ matrix, whose elements $a_{mn} = w_{mn}$ if and only if $\{v_m, v_n\} \in E$, otherwise, $a_{mn} = 0$.

It worth noting that if a graph is undirected, its adjacent matrix will be symmetric.

2. Proposed method

Since evidence theory has been proposed, different kinds of evidence combination methods have been proposed. Among them, Dempster's method [3] is the most popular evidence combination rule, and it has been widely used. However, when evidences are in high conflict, the result calculated by Dempster's method may be illogical.

To solve this problem, we propose a new method to combine evidence in conflict. The main idea of the proposed method is that different evidence has different reliability. The conflicting evidences should be identified and treated cautiously, and the reliable evidences should be trusted and given a high credibility. To determine the reliability of every evidence, two techniques are applied in the proposed method.

(1) Weighted averaging the BPA of every evidence.

The reliability (or the weight) of every evidence is calculated based on Pearson correlation coefficient. In general, if a evidence is reliable, it is supported by other evidences which means that the Pearson correlation of coefficient them is relatively high. After that, the weight can be used to weighted average the BPA of evidence, which can improve the accuracy of the method and the performance of convergence.

(2) Representing the relationship of every evidence based on weighted graph.

Graph is a tool to represent the relation of objects. The relationship of evidences can illustrated by weighted graph which can help researchers to directly identify the evidences in conflict or the relatively unreliable evidences.

In the rest of this section, firstly, some basic definitions are proposed. And then, a new evidence combination is present.

2.1. Basic definitions

Pearson correlation coefficient is the linear correlation of two samples. When the number of samples that we are dealing with is larger than two, we need a efficient way to reorganize and represent the linear correlation of them. Inspired of this, Pearson correlation coefficient matrix (PCCM) is proposed.

Definition 3.1: *Pearson correlation coefficient matrix (PCCM)*

Assume there are K samples denoted as vectors: $\overrightarrow{\boldsymbol{M_1}}, \overrightarrow{\boldsymbol{M_2}}, ..., \overrightarrow{\boldsymbol{M_K}}$. Then the Pearson correlation coefficient matrix (PCCM) of these samples is defined as:

$$PCCM = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1K} \\ r_{21} & r_{22} & & r_{2K} \\ \vdots & & \ddots & \vdots \\ r_{K1} & r_{K2} & \cdots & r_{KK} \end{bmatrix} \tag{9}$$

where r_{ij} is the Pearson correlation coefficient of sample $\overrightarrow{\boldsymbol{M_i}}$ and sample $\overrightarrow{\boldsymbol{M_j}}$.

PCCM has N^2 Pearson correlation coefficient so that this matrix can describes the linear correlation of all the samples. Assume the samples can be seen as nodes and the Pearson correlation coefficient as weight. Then we can convert the PCCM into a weighted graph and its adjacent matrix.

Definition 3.2: *PCCM-based weighted graph*

A PCCM-based weighted graph is a undirected weighted graph defined as $G_{PCCM} = (V, E, W)$, where $V = \{\overrightarrow{\boldsymbol{M_1}}, \overrightarrow{\boldsymbol{M_2}}, ..., \overrightarrow{\boldsymbol{M_K}}\}$ is the node set, $E =$

$\{\{\overrightarrow{\boldsymbol{M_i}}, \overrightarrow{\boldsymbol{M_j}}\} | \{\overrightarrow{\boldsymbol{M_i}}, \overrightarrow{\boldsymbol{M_j}}\} \in V \wedge V\}$ is the edge set, and $W = \{w_{ij} | i, j = 1, \ldots, K\}$ is the weight set whose element w_{ij} is defined as:

$$w_{ij} = \begin{cases} r_{ij} \;\; (r_{ij} > 0 \textit{ and } i \neq j) \\ 0 \quad (r_{ij} \leq 0 \textit{ or } i = j) \end{cases} \tag{10}$$

where r_{ij} is the Pearson correlation coefficient of PCCM. If $w_{ij} > 0$, node $\overrightarrow{\boldsymbol{M_i}}$ and node $\overrightarrow{\boldsymbol{M_j}}$ are connected, namely, $\{\overrightarrow{\boldsymbol{M_i}}, \overrightarrow{\boldsymbol{M_j}}\} \in E$. If $w_{ij} = 0$, node $\overrightarrow{\boldsymbol{M_i}}$ and node $\overrightarrow{\boldsymbol{M_j}}$ are unconnected, namely, $\{\overrightarrow{\boldsymbol{M_i}}, \overrightarrow{\boldsymbol{M_j}}\} \notin E$.

It should be noted that the PCCM-based weighted graph does not have self-loop. As a result, $w_{ij} = 0$ when $i = j$.

Definition 3.3: *Adjacent matrix of PCCM-based weighted graph*

A adjacent matrix of PCCM-based weighted graph is defined as:

$$A_{PCCM} = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1K} \\ w_{21} & 0 & & w_{2K} \\ \vdots & & \ddots & \vdots \\ w_{K1} & w_{K2} & \cdots & 0 \end{bmatrix} \tag{11}$$

where w_{ij} is the weight of PCCM-based weighted graph.

Because the PCCM-based weighted graph is undirected, its adjacent matrix is symmetric, namely, $w_{ij} = w_{ji}$.

2.2. Evidence combination algorithm

Assume that there are K evidences $m_1, m_2, \ldots, m_K$ and N alternatives $A_1, A_2, \ldots, A_N$. The BPA of these K evidences is $m_i(A_j)\,(i = 1, 2, \ldots, K \;\; j = 1, 2, \ldots, N)$. Then the proposed evidence combination algorithm is detailed as follows:

Step 1: Convert K evidences $m_i\,(\, i = 1, 2, \ldots, K\,)$ into vectors:

$$\overrightarrow{\boldsymbol{M_i}} = (\, m_i(A_1), m_i(A_2), \ldots, m_i(A_N)\,) \tag{12}$$

Step 2: Calculate the Pearson correlation coefficient matrix (PCCM) of K evidence vectors:

$$PCCM = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1K} \\ r_{21} & r_{22} & & r_{2K} \\ \vdots & & \ddots & \vdots \\ r_{K1} & r_{K2} & \cdots & r_{KK} \end{bmatrix} \tag{13}$$

Step 3: Convert PCCM into PCCM-based weighted graph $G_{PCCM} = (V, E, W)$.

Step 4: Obtain the adjacent matrix of PCCM-based weighted graph:

$$A_{PCCM} = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1K} \\ w_{21} & 0 & & w_{2K} \\ \vdots & & \ddots & \vdots \\ w_{K1} & w_{K2} & \cdots & 0 \end{bmatrix} \tag{14}$$

Step 5: Calculate the total weight TW_i of evidence m_i based on the adjacent matrix of PCCM-based weighted graph:

$$TW_i = \sum_{j=1}^{K} w_{ij} \tag{15}$$

Step 6: Normalize the total weight to achieve the normalized weight NW_i of evidence m_i:

$$NW_i = \frac{TW_i}{\sum_{i=1}^{K} TW_i} \tag{16}$$

Step 7: Based on NW_i, calculate the weighted average evidence WAE:

$$WAE = \{m(A_j) \mid j = 1, 2, \dots, N\} \tag{17}$$

$$m(A_j) = \sum_{i=1}^{K} m_i\,(A_j)\,NW_i \tag{18}$$

where $m_i\,(A_j)$ is the BPA for evidence m_i of the alternative A_j.

Step 8: Use Dempster's rule to combine the weighted averaged evidence $K-1$ times and get the combination result of K evidences.

3. Experiment and result

In this section, an experiment is used to illustrate the proposed evidence combination.

Assume there are three alternatives $\{A, B, C\}$ and five evidences $m_1, m_2, \ldots, m_5$ in a target recognition system. The BPA reports of 5 evidences are collected in Table 1.

Table 1: The BPA reports of 5 evidences

	$\{A\}$	$\{B\}$	$\{C\}$	$\{A,C\}$
m_1	0.50	0.20	0.30	0.00
m_2	0.00	0.90	0.10	0.00
m_3	0.45	0.20	0.00	0.35
m_4	0.50	0.20	0.00	0.30
m_5	0.45	0.25	0.00	0.30

These five evidences are combined by the proposed evidence combination, and then we can recognize the exact alternative of the three based on the result. The calculating steps are detailed as follows.

Step 1: Convert these five evidences into vectors:

$$\overrightarrow{\boldsymbol{M_1}} = (\,0.50, 0.20, 0.30, 0.00\,)$$
$$\overrightarrow{\boldsymbol{M_2}} = (\,0.00, 0.90, 0.10, 0.00\,)$$
$$\overrightarrow{\boldsymbol{M_3}} = (\,0.45, 0.20, 0.00, 0.35\,)$$
$$\overrightarrow{\boldsymbol{M_4}} = (\,0.50, 0.20, 0.00, 0.30\,)$$
$$\overrightarrow{\boldsymbol{M_5}} = (\,0.45, 0.25, 0.00, 0.30\,)$$

Step 2: Calculate the Pearson correlation coefficient matrix (PCCM) of these five evidence vectors:

$$PCCM = \begin{bmatrix} 1 & -0.146944 & 0.122679 & 0.307692 & 0.213980 \\ -0.146944 & 1 & -0.273408 & -0.257151 & -0.102190 \\ 0.122679 & -0.273408 & 1 & 0.981433 & 0.978284 \\ 0.307692 & -0.257151 & 0.981433 & 1 & 0.984309 \\ 0.213980 & -0.102190 & 0.978284 & 0.984309 & 1 \end{bmatrix} \tag{19}$$

Step 3: Convert PCCM into PCCM-based weighted graph $G_{PCCM} = (V, E, W)$ as Figure 1, where the full lines with number represent the weight w_{ij} of two nodes, and the dash lines indicate that two nodes are unconnected, namely, $w_{ij} = 0$.

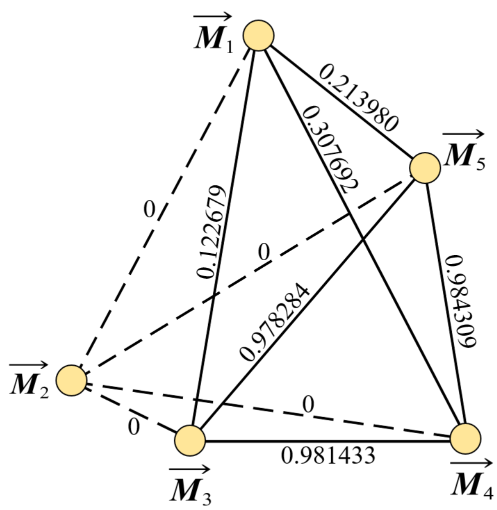


Figure 1: PCCM-based weighted graph

Step 4: Obtain the adjacent matrix of PCCM-based weighted graph:

$$A_{PCCM} = \begin{bmatrix} 0 & 0 & 0.122679 & 0.307692 & 0.213980 \\ 0 & 0 & 0 & 0 & 0 \\ 0.122679 & 0 & 0 & 0.981433 & 0.978284 \\ 0.307692 & 0 & 0.981433 & 0 & 0.984309 \\ 0.213980 & 0 & 0.978284 & 0.984309 & 0 \end{bmatrix} \tag{20}$$

Step 5: Calculate the total weight TW_i of evidence m_i based on the adjacent matrix of PCCM-based weighted graph:

$$\begin{aligned} TW_1 &= 0 + 0 + 0.122679 + 0.307692 + 0.213980 = 0.644351 \\ TW_2 &= 0 + 0 + 0 + 0 + 0 = 0 \\ TW_3 &= 0.122679 + 0 + 0 + 0.981433 + 0.978284 = 2.082396 \\ TW_4 &= 0.307692 + 0 + 0.981433 + 0 + 0.984309 = 2.273434 \\ TW_5 &= 0.213980 + 0 + 0.978284 + 0.984309 + 0 = 2.176573 \end{aligned} \tag{21}$$

Step 6: Normalize the total weight to achieve the normalized weight NW_i of evidence m_i:

$$\begin{aligned} NW_1 &= \frac{0.644351}{7.176754} = 0.089783 \\ NW_2 &= \frac{0}{7.176754} = 0 \\ NW_3 &= \frac{2.082396}{7.176754} = 0.290158 \\ NW_4 &= \frac{2.273434}{7.176754} = 0.316777 \\ NW_5 &= \frac{2.176573}{7.176754} = 0.303281 \end{aligned} \tag{22}$$

Step 7: Based on NW_i, calculate the weighted average evidence WAE:

$$\begin{aligned} m(A) &= \sum_{i=1}^{5} m_i\left(A\right) NW_i = 0.470328 \\ m(B) &= \sum_{i=1}^{5} m_i\left(B\right) NW_i = 0.215164 \\ m(C) &= \sum_{i=1}^{5} m_i\left(C\right) NW_i = 0.026935 \\ m(A,C) &= \sum_{i=1}^{5} m_i\left(A,C\right) NW_i = 0.287573 \end{aligned} \tag{23}$$

Step 8: Use Dempster's rule to combine the weighted averaged evidence 4 times and get the final combination result:

$$\begin{aligned} m(A) &= 0.985939 \\ m(B) &= 0.001833 \\ m(C) &= 0.004413 \\ m(A,C) &= 0.007816 \end{aligned} \tag{24}$$

In this experiment, we choose the other four typical evidence combination method, including Dempster's method[3], Murphy's method[21], Liu *et al*'s method[26] and Deng *et al*'s method[22] to compare with the proposed method. The experiment results of these five methods are shown in Table 2.

Table 2: Results of five evidence combination methods

Method	$m(A)$	$m(B)$	$m(C)$	$m(A,C)$	Target
Dempster's method [3]	0.00000000	0.65573770	0.34426230	0.00000000	B
Murphy's method [21]	0.89960650	0.07885140	0.01782472	0.00371738	A
Liu *et al*'s method [26]	0.95446411	0.00795387	0.03758202	0.00000000	A
Deng *et al*'s method [22]	0.96571592	0.01600922	0.01394744	0.00432741	A
Proposed method	0.98593887	0.00183259	0.00441303	0.00781550	A

For the convenience of discussion, the calculating procedure of the BPA $m(A)$ is shown in Table 3. It should be noted that, the total times of combining the five evidences is $5-1=4$, except for Liu *et al*'s method, because it removes the conflicting evidence m_2 and combines the rest of the four evidences by 3 times.

Table 3: The value of $m(A)$ by different times of combining

Method	Times = 1	Times = 2	Times = 3	Times = 4
Dempster's method	0.000000	0.000000	0.000000	0.000000
Murphy's method	0.596448	0.740307	0.836883	0.899607
Liu *et al*'s method	0.733945	0.890125	0.954464	/
Deng *et al*'s method	0.689934	0.844120	0.925710	0.965716
Proposed method	0.772013	0.909264	0.964400	0.985939

In the next section, the five evidence combination methods are analyzed based on the experiment result.

4. Analysis and discussion

In this section, to illustrate the efficiency of the proposed method, comparisons between the proposed method and the other four methods are analyzed and discussed.

In general, Dempster's method[3] is widely used to combine data from sensors based on evidence theory. Murphy's method[21] is an efficient and typical tool to combine conflicting evidences by simple averaging the BPA of evidences. Deng *et al*'s method[22] uses distance of evidence to calculate credibility of every evidence, which is a weighted-averaging-based method for dealing with conflicting evidence. Liu *et al*'s method[26] is a novel evidence combination based on generalized belief entropy, and this method uses graph model to improve the performance of combining evidences. The proposed method is based on Pearson correlation coefficient and weighted graph, which takes both weighted averaging technique and graph model into consideration. The techniques of these five methods are summarized in Table 4.

It can be seen from Table 2 that, the proposed method has the best performance because it successfully recognizes the correct alternative A based on the

Table 4: The techniques of five evidence combination methods

Method	Graph-based method	Averaging-based method
Dempster's method	$\times$	$\times$
Murphy's method	$\times$	$\checkmark$
Liu *et al*'s method	$\checkmark$	$\times$
Deng *et al*'s method	$\times$	$\checkmark$
Proposed method	$\checkmark$	$\checkmark$

conflicting evidence, and the BPA $m(A)$ calculated by the proposed method is the highest (0.985939) compared with other methods.

As is illustrated in Figure 2, although the difference of Murphy's, Liu *et al*'s, Deng *et al*'s and the proposed method is not large, proposed method can also identify the alternatives correctly under the condition that the threshold is 0.97.

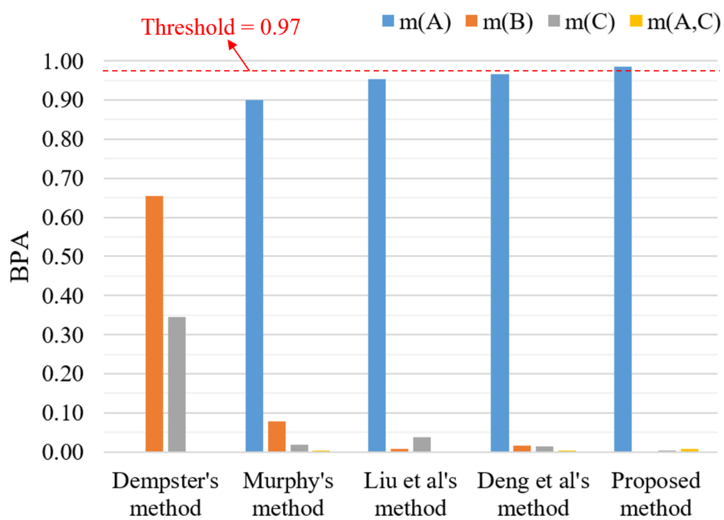


Figure 2: Results of five methods when the threshold is 0.97

When confronting extreme environment, sensors would be influenced by many factors such as radiation, temperature or design defects which cause the sensors to report evidences in high conflict with each other. Under this circumstance, the threshold of identifying target will be higher than common. With the highest accuracy, the proposed method is more reliable to combine the conflicting evidences, at least its result will not worse than the other four methods. As a result, the proposed method has the efficiency to handle conflict in a environment with high uncertainty.

Apart from above discussions, in order to show the advantages of the proposed method, more detailed comparisons are expounded in the following three subsections based on the techniques that the method uses.

4.1. Compared with Dempster's method

Dempster's method is neither an averaging-based method nor a graph-based method. When the evidences reported by sensors are in conflict with each other, Dempster's method may yield counter-intuitive results. In this experiment, four evidences support alternative A, while evidence m_2 supports B which is conflicted with other evidences. The result of Dempster's method shows that, even though more evidences support A, Dempster's method supports alternative B ($m(B) = 0.655738$) and is totally against A ($m(A) = 0$), which is illogical.

By contrast, as is both an averaging-based and a graph-based method, the proposed method draws the correct conclusion with high accuracy ($m(A) = 0.985939$), which can be a great alternative of Dempster's method to combine conflicting evidences.

4.2. Compared with averaging-based methods

Murphy's method, Deng *et al*'s method and the proposed method are averaging-based methods. In specific, Murphy's method is a simple-averaging method, namely, every weight of evidence is equal to each other. Deng *et al*'s method is a weighted-averaging method, which means that, the weight of evidence can be modified. The proposed method is actually a weighted-averaging method. Its weight can be changed based on the Pearson correlation coefficient.

All of the three averaging-based methods get the correct conclusion. However, compared with other averaging-based methods, the proposed method is more efficient. The advantages of it are analyzed as follows:

(1) Better performance of convergence.

According to Table 3, the calculating procedure of $m(A)$ based on averaging-based methods is shown in Figure 3. It can be seen in this figure that, obviously, at every time of combining, the BPA $m(A)$ of the proposed method is the highest. Besides, the speed of convergence of the proposed method is the best, since the value of $m(A)$ reaches more than 0.9 only by two times of combining.

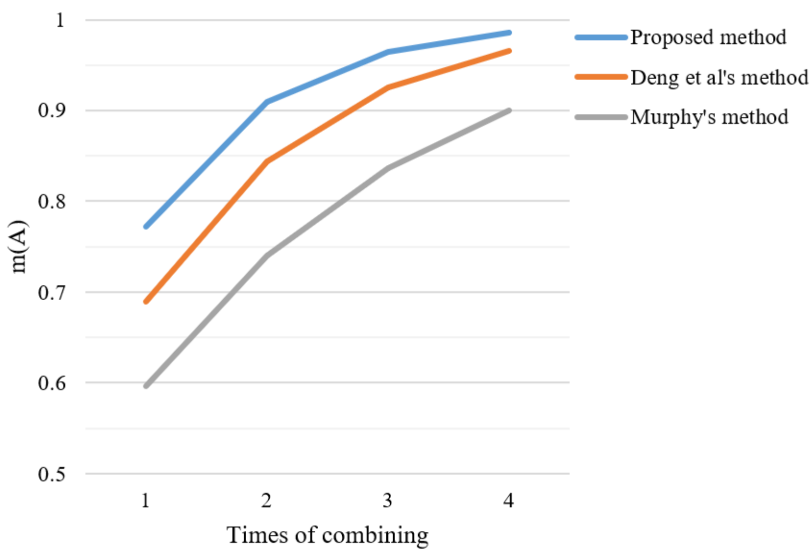


Figure 3: The calculating procedure of $m(A)$ based on averaging-based methods

(2) More efficient to identify the reliability of evidences

Beyond being a weighted-averaging-based method, the proposed method is also a graph-based method. According to Step 5 of the proposed method, the algorithm generates a weighted graph. Compared with other averaging-based methods, the proposed can directly reflect the relationship of the evidences based on the weighted graph.

As is illustrated in Figure 4, the node $\overrightarrow{\boldsymbol{M_2}}$ has no edge connected with other node, which means that m_2 is not supported by other evidences. As a result, we can directly identify that m_2 is the conflicting evidence which should be carefully checked. Besides, the nodes $\overrightarrow{\boldsymbol{M_3}}$, $\overrightarrow{\boldsymbol{M_4}}$ and $\overrightarrow{\boldsymbol{M_5}}$ connect with each other with high weight more than 0.97, which indicates that, m_3, m_4 and m_5 highly support to each other. Hence, we can trust these three evidences. Moreover,

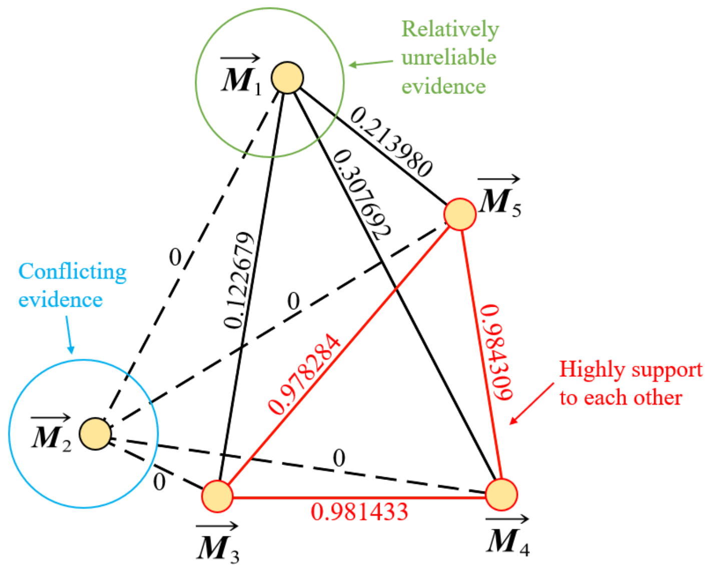


Figure 4: Identify the reliability of evidences based on weighted graph

three weights of the node $\overrightarrow{\boldsymbol{M_1}}$ are relatively low, which alerts us that, m_1 is relatively unreliable, and the BPA reported by m_1 should be taken with a grain of salt.

To summarize, compared with other averaging based methods, the advantages of the proposed method are the great performance of convergence and the efficiency of identifying evidence in conflict.

4.3. Compared with graph-based methods

Liu *et al*'s method and the proposed method are graph-based methods. Both of the two methods recognize the correct alternative A. However, the proposed method is better than Liu *et al*'s method. The reasons are as follows:

(1) Better performance of convergence.

The calculating procedure of $m(A)$ based on graph-based methods is shown in Figure 5. It worth noting that, since Liu *et al*'s method removes the conflicting evidence m_2 and combine the rest of the 4 evidences, the times of combining of it are $4 - 1 = 3$, while the the times of combining of the proposed method are 4.

It is obviously that, at every time of combining, the BPA $m(A)$ of the proposed method is the higher than that of Liu *et al*'s method, which means that, the performance of convergence of the proposed method is better than Liu *et al*'s method. Although the the BPA $m(A)$ of the proposed method is close to that of Liu *et al*'s method at the the first three times of combining, the proposed method can still enhance the value of $m(A)$ at the $4th$ times, improving the ability to identify the correct target.

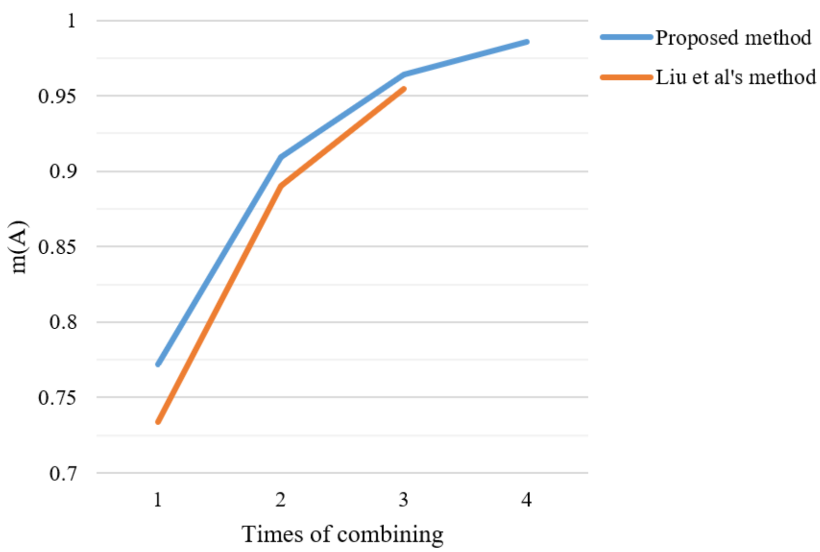
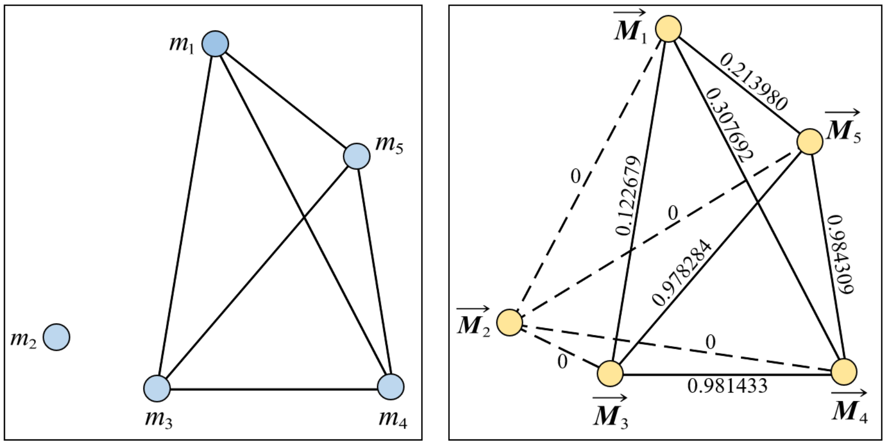


Figure 5: The calculating procedure of $m(A)$ based on graph-based methods

(2) More efficient to identify the reliability of evidences

Both of the two graph-based methods can generate a graph to identify conflicting evidences. As is illustrated in Figure 6, Liu *et al*'s method generates a simple graph whose edge does not have weight, and the connection state is just true (1) or false (0). By contrast, the proposed method generates a weighted graph with weight attached to the edges which can better represent the relationship between nodes compared with simple graph.

For example, in Figure 6 (b) generated by the proposed method, there are three edges connected to $\overrightarrow{\boldsymbol{M_1}}$ (the dash line means two nodes are unconnected); the weight of these edges are 0.122679, 0.307692 and 0.213980 which represents



(a) Simple graph based on Liu et al's method (b) Weighted graph based on proposed method

Figure 6: Comparison between simple graph and weighted graph

that the relation of node $\overrightarrow{\boldsymbol{M_1}}$ between other nodes is not close. While the weights attached to the edges of $\overrightarrow{\boldsymbol{M_3}}$, $\overrightarrow{\boldsymbol{M_4}}$ and $\overrightarrow{\boldsymbol{M_5}}$ are higher than that of $\overrightarrow{\boldsymbol{M_1}}$. As a result, we can draw the conclusion that, evidence m_1 is relatively unreliable compared with m_3, m_4 and m_5, and the BPA reports of evidence m_1 should be treated cautiously. On the contrary, in Figure 6 (a) generated by Liu *et al*'s method, node m_1, m_3, m_4 and m_5 are connected to each other, from which we can not distinguish the difference of reliability of evidence m_1 from m_3, m_4 and m_5.

In conclusion, the proposed method is better than Liu *et al*'s method in terms of performance of convergence and the efficiency of identifying reliability of evidences.

5. Conclusion

In this paper, based on Pearson correlation coefficient and weighted graph, a novel evidence combination method is proposed, which is both a averaging-based combination method and a graph-based method, improving the performance of combining evidence in conflict. In addition, an experiment is expounded, and the results show the efficiency of the proposed method. Moreover, compared with other common evidence combination methods, the advantages of the proposed method are analyzed and discussed which are summarized as follows:

(1) The proposed method can correctly recognize the target among other alternatives, and its accuracy is better than other methods.

(2) Compared with other combination methods, the proposed method has the best performance of convergence.

(3) As a graph-based method, the proposed method can generates a weighted graph to directly reflect the relationship of different evidences, giving us a ideal way to identify the reliability of every evidence.

Acknowledgment

The work is partially supported by National Natural Science Foundation of China (Grant No. 61973332).

References

1. Lee P. Probability theory. *Bulletin of the London Mathematical Society* 1980;12(4):318–9.

2. Zadeh LA. Fuzzy sets. *Information and control* 1965;8(3):338–53.

3. Dempster AP. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics* 1967;38(2):325–39.

4. Shafer G. A mathematical theory of evidence; vol. 1. Princeton university press Princeton; 1976.

5. Pawlak Z. Rough sets. *International journal of computer & information sciences* 1982;11(5):341–56.

6. Liu B, Deng Y. Risk evaluation in failure mode and effects analysis based on d numbers theory. *International Journal of Computers, Communications & Control* 2019;14(5).

7. Luo Z, Deng Y. A matrix method of basic belief assignment's negation in dempster-shafer theory. *IEEE Transactions on Fuzzy Systems* 2019;.

8. Yin L, Deng X, Deng Y. The negation of a basic probability assignment. *IEEE Transactions on Fuzzy Systems* 2018;27(1):135–43.

9. Deng W, Deng Y. Entropic methodology for entanglement measures. *Physica A: Statistical Mechanics and its Applications* 2018;512:693–7.

10. Yager RR, Alajlan N, Bazi Y. Uncertain database retrieval with measure-based belief function attribute values. *Information Sciences* 2019;501:761–70.

11. Song Y, Deng Y. A new method to measure the divergence in evidential sensor data fusion. *International Journal of Distributed Sensor Networks* 2019;15(4):1550147719841295.

12. Li Y, Deng Y. Generalized ordered propositions fusion based on belief entropy. *Int J Comput Commun Control* 2018;13(5):792–807.

13. Xiao F. A multiple-criteria decision-making method based on d numbers and belief entropy. *International Journal of Fuzzy Systems* 2019;21(4):1144–53.

14. Koksalmis E, Kabak O. Sensor fusion based on dempster-shafer theory of evidence using a large scale group decision making approach. *International Journal of Intelligent Systems* 2020;35(7):1126–62.

15. Zhang H, Deng Y. Engine fault diagnosis based on sensor data fusion considering information quality and evidence theory. *Advances in Mechanical Engineering* 2018;10(11):1687814018809184.

16. Bian T, Deng Y. Identifying influential nodes in complex networks: A node information dimension approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 2018;28(4):043109.

17. Wen T, Pelusi D, Deng Y. Vital spreaders identification in complex networks with multi-local dimension. *Knowledge-Based Systems* 2020;:105717.

18. Zhao J, Deng Y. Complex network modeling of evidence theory. *IEEE Transactions on Fuzzy Systems* 2020;:10.1109/TFUZZ.2020.3023760.

19. Xiao F. Generalization of Dempster–Shafer theory: A complex mass function. *Applied Intelligence* 2019;:DOI: 10.1007/s10489–019.

20. Yager RR. A class of belief structures based on possibility measures. *Soft Computing* 2018;22(23):7909–17.

21. Murphy CK. Combining belief functions when evidence conflicts. *Decision support systems* 2000;29(1):1–9.

22. Yong D, WenKang S, ZhenFu Z, Qi L. Combining belief functions based on distance of evidence. *Decision support systems* 2004;38(3):489–93.

23. Yager RR. On the fusion of non-independent belief structures. *International journal of general systems* 2009;38(5):505–31.

24. Fei L, Deng Y. A new divergence measure for basic probability assignment and its applications in extremely uncertain environments. *International Journal of Intelligent Systems* 2019;34(4):584–600.

25. Li H, Xiao F. A method for combining conflicting evidences with improved distance function and tsallis entropy. *International Journal of Intelligent Systems* 2020;.

26. Liu F, Gao X, Zhao J, Deng Y. Generalized belief entropy and its application in identifying conflict evidence. *IEEE Access* 2019;7(1):126625–3.

27. Li P, Fei L. On combination rule in dempster-shafer theory using owa-based soft likelihood functions and its applications in environmental impact assessment. *International Journal of Intelligent Systems* 2019;34(12):3168–89.

28. Pan L, Deng Y. An association coefficient of a belief function and its application in a target recognition system. *International Journal of Intelligent Systems* 2020;35(1):85–104.

29. Xu H, Deng Y. Dependent evidence combination based on decision-making trial and evaluation laboratory method. *International Journal of Intelligent Systems* 2019;34(7):1555–71.

30. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. In: *Noise reduction in speech processing*. Springer; 2009:1–4.

31. Monedero I, Biscarri F, León C, Guerrero JI, Biscarri J, Millán R. Detection of frauds and other non-technical losses in a power utility using pearson coefficient, bayesian networks and decision trees. *International Journal of Electrical Power & Energy Systems* 2012;34(1):90–8.

32. Wang GJ, Xie C, Stanley HE. Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlation-based networks. *Computational Economics* 2018;51(3):607–35.

33. Xu H, Deng Y. Dependent evidence combination based on shearman coefficient and pearson coefficient. *IEEE Access* 2017;6:11634–40.

34. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’networks. *nature* 1998;393(6684):440–2.

35. Hofman JM, Sharma A, Watts DJ. Prediction and explanation in social systems. *Science* 2017;355(6324):486–8.

36. Gross JL, Yellen J. Graph theory and its applications. CRC press; 2005.