Multi-source Data Fusion Approach Based on Improved Evidence Theory

Yongwei Wang
Zhengzhou Information Science and Technology Institute, Zhengzhou, China
Email: wywyongweip@126.com

Kaiguo Yuan
Beijing University of Posts and Telecommunications, Beijing, China
Email: flyingdreaming@gmail.com

Yunan Liu, Hongyong Jia, Wei Qiu
Zhengzhou Information Science and Technology Institute, Zhengzhou, China
Email: {lyn37465@163.com, Jhyzz1975@163.com, 251197915@qq.com}

***Abstract*—The classical evidence theory can result in paradox in the process of information fusion. To resolve this problem, a multi-source data fusion method based on dissimilarity matrix and evidence theory is proposed. First, using the weighted Euclidean distance, evidence dissimilarity matrix is constructed. Second, dissimilarity between the evidences is measured. Third, using dissimilarity matrix, supporting degree, credibility and weight of evidence are calculated, and the original evidences are modified. Finally, using the improved combination rule, the information fusion is completed. Experimental results show that new method is superior to the existing typical methods in accuracy, discrimination and accuracy of fusion results.**

***Index Terms*—dissimilarity matrix, multi-source heterogeneous, combination rule, information fusion**

I. INTRODUCTION

With the diversification and intelligent of network intrusion methods, the network intrusion detection, network firewall, anti-virus systems, and terminal monitoring system are established based on the depth of defense system[1], but at the same time, it brings some problems. On one hand, because of the high false alarm rate, alarm overlap, omissions and weak semantic issues of intrusion detection system, great difficulties for timely identification of network intrusion, analysis and response are brought. On the other hand, with the increasing of information system scale, all kinds of alarm information and log information are growing with magnitude speed. What is more, as the original alarms are underlying message, they are too simple and high redundancy, and there are false alarm problem. In a word, the effective heterogeneous information analysis technology is urgently needed [2][3].

Evidence theory is a theory of uncertain reasoning, which was first proposed by Dempster in 1967. His student Shafer further developed it. Therefore, evidence theory is also called D-S evidence theory. Evidence theory can effectively represent and process uncertainty and imprecision information [4][5][6]. It has been widely used in the field of information fusion [7][8][9][10]. But in the actual information fusion system, there are often conflicting sensors reports due to the interference of natural environment or other reasons. The classical Dempster-Shafer evidence theory cannot deal the integration of conflict information effectively. When there are conflicts between the evidences, if using the Dempster's combination rule to integrate evidences directly, the result is often contrary to the true situation [11]. Therefore, when the degree of conflict between evidences is high, how to achieve integration effectively becomes an urgent problem. As to solve this problem, researchers have proposed many improved methods. Murphy [12] proposed an average method to deal with the evidence confliction based on modified model, and it has a faster convergence rate. The inadequacy of this method is only premeditating the simple averaging of evidence fusion, without considering their mutual relevance. Deng [13] improved Murphy's method, he assigned different weights for each evidences according to the mutual support between the evidences, and it is better able to suppress interference and faster convergence. Considering the consistency of evidence synthesis and application fields, Liang [14] proposed an evidence combination rule of absorption consistent evidence conflict. Tazid [15] proposed a combination method by alternating multiplicative strategy to adding strategy, but its conflicting process method is so moderate that the convergence rate is too slow.

In summary, the existing improved methods of evidence theory can be roughly divided into two

National High Technology Research and Development 863 Program of China under Grant: No. 2012AA012704; National Basic Research Program of China: No. 2011CB311801

Corresponding Author: Yongwei Wang (1977—): Male, Ph D. His major research interests include information security, computer network and situation awareness.

doi:10.4304/jcp.8.11.2864-2872

categories [16]. In the first category, the methods modify the evidences, which can reduce the impact of unreliable evidence on the combination conclusion. This process method can solve the paradox problem to some extent, but the modification of the evidence sources may cause loss of information, making the human factors involve in combination of evidence. It may distort the intention of the evidence itself. In the second category, eliminating the impact of various paradoxes by reassigning the conflict confidence to the power set space or relevant focus elements. However, in the circumstances of multi-focal element and evidences, taking the conflict as a measure of the relationship between evidences is not accurate, and the valuation of the conflict is often high[17]. Therefore, the results of these improved methods are unsatisfactory.

In order to resolve this problem, an evidence combination method based on dissimilarity matrix is proposed. First, using the weighted Euclidean distance, evidence dissimilarity matrix is constructed. Second, the degree of dissimilarity between the evidences is measured. Third, using dissimilarity matrix, the support, credibility and weight of evidence are calculated to modify the original evidences. Finally, using the improved combination rule, the information fusion is completed. Experimental results show that new method is superior to the existing typical methods in discrimination ability, fusion efficiency and accuracy of fusion results.

In section 2, the concepts of evidence theory and its problems are introduced. In section 3, evidence combination method based on dissimilarity matrix is discussed. In section 4, the simulation examples of improved algorithm is conducted, the experimental results and analysis are given out. The final section is the summary of this paper.

II. THE BASIC CONCEPTS OF EVIDENCE THEORY AND ITS PROBLEMS

Definition 1. Framework of Discernment

Framework of discernment (FoD) is a comprehensive collection of Θ. All elements of Θ are mutually exclusive. On an issue, the answer only take an element of Θ at any time, and all subsets of the collection are corresponding to all possible answers of questions. The complete set of mutually exclusive events Θ is called as framework of discernment [18].

Definition 2. Basic Probability Assignment

Basic probability assignment (BPA) is a function from set 2^{Θ} to [0, 1]. A represents any subset of target framework of discernment and it satisfies:

$$m(\phi)=0\ ,\quad \sum_{A\subseteq\Theta} m(A)=1 \tag{1}$$

In formula (1), $m(A)$ is the basic probability assignment function.

Definition 3. Belief Function

Belief function (Bel) is a mapping from set 2^{Θ} to [0, 1], and it is defined as follows [16].

$$Bel(A)=\sum_{B\subseteq A} m(B) \tag{2}$$

$Bel(A)$ is the probability function on Θ. Formula (2) indicates that when the set B is logically implied in A, the confidence of A is the sum of all propositions' confidence which contains B.

Basic probability function is the basis for the measurement of the proposition's uncertainty. In some cases, even if for the same evidence, as the data source is not the same, there will be two or more different basic probability assignment functions. Then, two or more basic probability assignment function needs to be merged into one basic probability assignment function. Therefore, Dempster proposed a combination method, in which two or more basic probability assignment functions are needed to orthogonal and computing. The method is called D-S combination rule and it is defined as follows.

Definition 4. Fusion Rule

When $\forall A\subseteq\Theta$, m_1 and m_2 are basic probability assignment function on Θ, and assuming that m_1 and m_2 are two mutually independent basic probability assignment on 2^{Θ}, then the combination rule is defined as follows[20][21].

$$m(A)=m_1\oplus m_2=\begin{cases}\dfrac{\sum_{A_i\cap B_j=A} m_1(A_i)m_2(B_j)}{1-K} & A\neq\phi\\ 0 & A=\phi\end{cases} \tag{3}$$

In formula (3), $K=\sum_{A_i\cap B_j=\Phi} m_1(A_i)m_2(B_j)<1$, K is a regularization factor and it can be used to guarantee that $m(A)$ is a basic probability assignment function. It is the sum of probability assignment values of all non-conflicting combination propositions in the discernment frame. By normalization, the probability of conflicting proposition is reassigned to the non-conflicting proposition. If $k\to 0$, it indicates that there is no conflict between the evidences. When $k\to 1$, it means that there is a serious conflict between the evidences. If the evidence is combating at this time, it will produce a result which is contrary to human's intuition. This phenomenon is called Zadeh's paradox.

The example of Zadeh's paradox is as follows.

Example 1. Assuming that $\Theta=\{A,B,C\}$ is the frame of discernment. The probability of evidences is shown in table I.

TABLE I
BASIC PROBABILITY ASSIGNMENT

Serial	$m(A)$	$m(B)$	$m(C)$
1	0.98	0.01	0.01
2	0.00	0.01	0.99
3	0.90	0.01	0.09

It can be seen that evidence 1 and evidence 2 are highly conflicting. They are highly supportive of proposition A and proposition C respectively. Evidence 3 is also highly supportive of proposition A. From intuition, proposition A is the final proper conclusion. Fusing the three evidences based on the combination rule of D-S evidence theory, the results are shown as follows.

$$\begin{cases} m(A) = 0.0000, \\ m(B) = 0.0011, \\ m(C) = 0.9989 \end{cases}$$

It is easy to know that the conclusion is contrary to the common intuition. The probability value of highly supportive proposition A becomes zero after fusion, and the probability value of proposition B becomes 0.9989 after fusion. The combination rule of D-S evidence theory achieves wrong results.

III. IMPROVED EVIDENCE COMBINATION METHOD BASED ON DISSIMILARITY MATRIX

In the fusion of evidence, if the evidence combination rule is used in the combination of relevant evidence directly without considering the relevance of evidence, the combination results will be over estimation. As there is correlation between the evidences, it is necessary to judge the different level between evidence and other evidences. The distance is a method to measure the conflicting of evidence. If the distance is larger, it means that the conflict of evidences is larger and the dissimilarity is greater too, and vice versa. Therefore, an evidence combination method based on dissimilarity is proposed in this paper. The basic idea of this paper is considering the supporting degree of one evidence supported by the others evidences. Due to the weight of different evidence is different, if evidence is supported by other evidences, then it should be more reliable and its weight is correspondingly larger. Conversely, if the conflict is larger, the credibility of evidence will be lower and the weight is smaller.

A. Dissimilarity Measure

There are generally three kinds of methods to measure the dissimilarity of concepts. Euclidean distance, Manhattan distance and Minkowski distance. Euclidean distance is suit for the calculation of numerical objects, and the calculation method is simple. Therefore, this paper calculates the dissimilarity of evidences based on the weighted Euclidean formula. The concept of evidence dissimilarity is defined as follows.

Definition 5. Evidence Dissimilarity

Assuming that E_i and E_j are two instances evidence of the target framework of discernment Θ . m_1 and m_2 are the corresponding basic probability assignment function. A_i and B_j are focal element , then the dissimilarity between E_i and E_j can be expressed as follows.

$$d_{ij}(E_i,E_j)=\begin{cases}0, & C_1=C_2 \\ [(w_1|A_1-B_1|^2+w_2|A_2-B_2|^2+\cdots+w_n|A_n-B_n|^2)^{1/2}, & C_1\neq C_2\end{cases} \tag{4}$$

d_{ij} represents the dissimilarity of the evidence and it can be used to describe the differences degree of evidences. Assuming there are n evidences, then the formula (4) can be used to calculate the dissimilarity between E_i and E_j . What is more, it can be expressed in the form of dissimilarity matrix.

$$D=\begin{pmatrix} d_{11} & d_{12} \cdots & d_{1n} \\ d_{21} & d_{22} \cdots & d_{2n} \\ \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} \cdots & d_{nn} \end{pmatrix} \tag{5}$$

By adding each row of the dissimilarity matrix, we can get the dissimilarity supporting degree of the other evidences to evidence E_i . The dissimilarity supporting degree is defined as follows.

$$DifSup(m_i) = \sum_{j=1,i\neq j}^{n} D(i,j) \qquad i,j=1,2,\cdots,n \tag{6}$$

$DifSup(m_i)$ reflects the overall difference degree between evidence E_i and the other evidences. If one evidence is different to all the others, the dissimilarity supporting degree is higher and the credibility of the evidence is lower. If the dissimilarity supporting degree between one evidence and the others is lower, the supporting degree between the evidence and others is higher. So does the credibility of evidence.

Combined the dissimilarity, the entropy in information theory is introduced to measure the importance of evidence. Assuming that ENT_i is the entropy of E_i , and the entropy value is as follows.

$$ENT_i = Sup(m_i)_i \ln(Sup(m_i)) \tag{7}$$

As the information entropy of evidence is proportional to its dissimilarity supporting degree, the credibility of the evidence can be obtained after the normalization of formula (7)'s result. The credibility of m_i is as follows.

$$Confid(m_i)=\frac{1/ENT_i}{\sum_{i=1}^{n} 1/ENT_i} \qquad i,j=1,2,\cdots,n \tag{8}$$

B. Improved Combination Rule

- **The combination rule of two evidences**

The credibility $Confid(m_i)$ reflects the credibility of evidence E_i . Generally, if the supporting degree of evidence supported by the other evidences is higher, the credibility of the evidence is higher too. Conversely, if one is not supported by other evidence, the credibility is lower.

From the formula (8), it can be seen that the sum of $Confid(m_i)$ is equal to 1. That is, $Confid(m_i)$ can be regarded as the weight of m_i . After obtaining the weight of evidence, the evidence can be weighted and the obtained basic confidence assignment values of improved E_i is as follows.

$$m'(A) = Confid(m_i) \cdot m(A), m'(\Theta) = 1 - \sum_{i=1}^{n} m_i'(A) \quad (9)$$

Thus, the improved combination rule of evidence theory is as follows.

$$m(A) = m_1' \oplus m_2' = \begin{cases} \dfrac{\sum_{A_i \cap B_j = A} m_1'(A_i) m_2'(B_j)}{1-K} & A \neq \phi \\ 0 & A = \phi \end{cases} \quad (10)$$

$$= \begin{cases} \dfrac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1-K} \cdot \dfrac{1/ENT_1}{\sum_{i=1}^{n} 1/ENT_i} \cdot \dfrac{1/ENT_2}{\sum_{i=1}^{n} 1/ENT_i} & A \neq \phi \\ 0 & A = \phi \end{cases}$$

- **The combination rule of multiple evidences**

When the number of evidence to be combined is more than two, the new combination rule adopts different method according to evidences' arrival time. If n evidences arrival at the same time, the n evidences should be completely combined once. If n evidences arrival at different time, the n evidences should be combined one by one according to the arrival order.

When n evidences arrival at the same time, the adopted combination rule is as follows.

$$m(A) = m_1' \oplus m_2' \cdots \oplus m_n' = \begin{cases} \dfrac{\sum_{\cap_{i=1}^{n} A_i = A} \prod_{i=1}^{n} m_i'(A_i)}{1-K} & A \neq \phi \\ 0 & A = \phi \end{cases} \quad (11)$$

$$= \begin{cases} \dfrac{\sum_{\cap_{i=1}^{n} A_i = A} \prod_{i=1}^{n} m_i(A_i)}{1-K} \cdot \dfrac{1/ENT_1 \cdot 1/ENT_2 \cdots 1/ENT_n}{(\sum_{i=1}^{n} 1/ENT_i)^n} & A \neq \phi \\ 0 & A = \phi \end{cases}$$

When two evidences are combined, the conflict is equal to the sum of products of focal element probability value. Accordingly, when n evidences are combined, the conflict calculation formula is as follows.

$$k = \sum_{\cap_{i=1}^{n} A_i = \phi} \prod_{i=1}^{n} m_i(A_i) \quad (12)$$

When n evidences arrival at different time, the formula (10) can be used n-1 times in two-two combination mode. The algorithm of two-two combination mode can be expressed as follows.

Algorithm 1 Two-two mode multiple evidences combination algorithm

Input: Evidence vector $E = \{E_1, E_2, ..., E_n\}$)

Output: Combination Conclusion)

```
BEGIN
  //calculate Dissimilarity Matrix using formula (4);
  CalculateDMatrix();
  For i=1 to n
    CalculateENT();
      //calculate confide( m_i ) using formula (8);
    CalculateConfid();
      //calculate m' using formula (9);
    CalculateMtemp();
  end for
  For i=1 to n
    IF n>2
      e_x = e_i ;
      e_y = e_{i+1} ;
      //fusion evidence using formula(10)
      FusionEvidence( e_x , e_y );
      If i==n-1
        return result;
      end if
      i=i+1;
    end if
  end for
END
```

Figure 1. Two-two mode multiple evidences combination algorithm

B. Confidenc Redistribution

In the actual application, decisions are made by single focus element. In order to assure the accuracy of the result, it is necessary to redistribute the confidence of the multi-element to single element in appropriate proportion. $m(A)$ is the confidence assigned to proposition A, which represents the measurement of absolute confidence to proposition A.

Definition 6. Relative Confidence

The ratio of single elements is defined as relative confidence. It can be expressed as follows.

$$\delta = \frac{m(A)}{m(B)} \quad (13)$$

In formula (13), δ represents the relative confidence of proposition A and proposition B. According to relative confidence, confidence of multi-element is redistributed. The redistribution is conducted in equal proportion. It can be expressed as follows.

$$m'(A) = m(A) + \frac{\delta}{1+\delta} m(A,B) \quad (14)$$

$$m'(B) = m(B) + \frac{1}{1+\delta} m(A,B) \quad (15)$$

$$m'(A,B) = 0 \quad (16)$$

By confidence redistribution, the confidence is only distributed in single focus elements, thus, it is conveniently to make decision, which can improve the discrimination of combination conclusion.

According to the new combination rule in figure 1, the combination result of example 1 is as follows.

$$\begin{cases} m(A) = 0.9049, \\ m(B) = 0.0075, \\ m(C) = 0.0876 \end{cases} \quad (17)$$

The results of new synthetic methods are consistent with the intuition, and it is more reasonable reflecting the distribution of confidence in original evidence. What is more, it can avoid the Zadeh's paradox problem. Because the dissimilarity in the new combination methods can properly measure the conflict degree of evidence, the new combination rule can effetely reduce the neglect effect of the unreliable evidences.

IV. Assessment of Combination Method

Yang et al [22] argue that the performance of evidence theory can be measured from two aspects. The first aspect is whether the combination results are in line with people's logical reasoning. It means that whether it is able to get the expected conclusion. The second aspect is whether the uncertainty of object proposition is reduced after the evidence synthesis. It means that whether the sum of single element's credibility is increased.

A. ZI Attribute

The first measure standard is used to examine whether the combination method can avoid the Zadeh's paradox. It is used to measure the immunity ability to Zadeh's paradox which is called ZI attribute (Zadeh Immunity). The agreement degree between combination conclusions and the logical reasoning results is higher, the ZI attribute is better. Thus, the accuracy of combination method is higher.

B. Discrimination

The second measure standard is used to examine the discrimination ability of the combination results. The sum of single focus element's credibility is larger, the discrimination ability of combination method is stronger, and the performance of the method is better. It is called sum discrimination in this paper and its definition is as follows.

Definition 7. Sum Discrimination

The sum of single focus element's credibility in the combination conclusions is called sum discrimination.

Assuming that E is an evidence in frame of discernment $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, and m_i is the corresponding basic probability assignment function, A_i is the focal element, then the sum discrimination of combination results can be expressed as follows.

$$D_s = \sum m(A_i),\ A_i \text{ is a single focus element} \quad (18)$$

The value of D_s is bigger, the combination method is better. If D_s is larger, it indicates that the spread of confidence to union space is smaller and the discrimination ability is stronger. When D_s = 1, it is easier to make decisions and the sum discrimination is the biggest. When D_s = 0, it cannot make a decision and the sum discrimination is the smallest.

In addition, the difference between the largest and second largest confidence is greater in the combination conclusion, it is easier to make decisions. Therefore, the concept of difference discrimination is introduced as follows.

Definition 8. Difference Discrimination

The difference value between the largest and second largest confidence in combination conclusions is called difference discrimination.

Assuming that E is an evidence in frame of discernment $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, and m_i is the corresponding basic probability assignment function, A_i is the focal element. The difference discrimination of combination results can be expressed as follows.

$$D_d = \mathrm{Set}_1 - \mathrm{Set}_2 \quad (19)$$

Set_1 represents the biggest confidence value of a single focal element. Set_2 represents the second largest confidence value.

$$\mathrm{Set}_1 = \max\{m_1(A_1), \cdots m(A_i), \cdots m_n(A_n)\} \quad (20)$$

$$\mathrm{Set}_2 = \max\{\{m_1(A_1), \cdots, m_n(A_n)\} - \max\{m_1(A_1), \cdots, m_n(A_n)\}\} \quad (21)$$

The value of D_d is bigger, the combination method's performance is better. If D_d is larger, it indicates that the extraction ability of combination method is stronger. When D_d = 1, it is easier to make decisions and the difference discrimination is the biggest. When D_d = 0, it cannot make a decision and the difference discrimination is the smallest.

C. Fusion Efficiency

The evidence theory is an important method to handle multi-source data fusion. When the number of evidence sources is large, the faster the combination method gets the right conclusion, the excellent the combination method is. Thus, the concept of fusion efficiency is introduced as follows.

Definition 9. Fusion Efficiency

It can be represented by the smallest number of evidence which can draw combination conclusions.

Assuming that $\{E_1, E_2, \cdots E_i \cdots, E_n\}$ is the set of n evidences in the frame of discernment $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$. If the i_{th} evidence arrives, we can get the synthetic conclusion. Thus, the fusion efficiency of combination method can be expressed by i.

If the value of fusion efficiency is smaller, it indicates that the ability of combination algorithm's information caption ability is stronger and the fusion effect is better.

V. EXPERIMENTS AND ANALYSIS

In order to verify the performance of the new method in this paper, five typical fusion methods are chosen and compared in experiments. The chosen methods are Dempster's method, Dubois&Prade's method, Murphy's method, Deng's method and Tazid's method.

A. Experimental Data

For the convenience of comparison, we choose the existing classic example to compare in the literature [15]. The configuration of discernment frame is $\Theta = \{A = Fighter, B = BombingPlane, C = Airliner\}$. The basic probability assignment of evidences is as follows.

TABLE II
BASIC PROBABILITY ASSIGNMENT SET

Serial Number	$m(A)$	$m(B)$	$m(C)$
1	0.50	0.20	0. 30
2	0.00	0.90	0.10
3	0.55	0.10	0.35
4	0.55	0. 10	0.35
5	0.60	0.10	0.30

B. Experimental Results

The dissimilarity matrix can be obtained by equation (4).

$$D = \begin{pmatrix} 0.0000 & 0.8832 & 0.1225 & 0.1225 & 0.1414 \\ 0.8832 & 0.0000 & 1.0025 & 1.0025 & 1.0198 \\ 0.1225 & 1.0025 & 0.0000 & 0.0000 & 0.0707 \\ 0.1225 & 1.0025 & 0.0000 & 0.0000 & 0.0707 \\ 0.1414 & 1.0198 & 0.0707 & 0.0707 & 0.0000 \end{pmatrix} \quad (22)$$

The confidence vector can be obtained by dissimilarity matrix (17) and equation (6), (7) and (8). The result of confidence vector is as follows.

$$C = \{0.2095,\ 0.0119,\ 0.2971,\ 0.2971,\ 0.1843\} \quad (23)$$

In order to demonstrate the absorption capacity of evidence combination method to new evidence clearly, according to the algorithm provided in Figure1, we use two-two combination method to fusion the evidence in the experiment. The results of two-two combination are shown in table III.

TABLE III
RESULTS OF DIFFERENT COMBINATION METHODS

Serial Number	Focus	$m_{1,2}$	$m_{1,2,3}$	$m_{1,2,3,4}$	$m_{1,2,3,4,5}$
Dempster's method	A	0.0000	0.0000	0.0000	0.0000
	B	0.8571	0.6316	0.3288	0.2228
	C	0.1429	0.3684	0.6712	0.8772
Dubois& Prade's method	A	0.0000	0.2800	0.3559	0.4211
	B	0.1800	0.0937	0.0519	0.0320
	C	0.0300	0.1319	0.0914	0.0867
	AB	0.4500	0.3529	0.2613	0.1952
	AC	0.0500	0.0448	0.1864	0.2388
	BC	0.2900	0.0967	0.0531	0.0262
Murphy's method	A	0.1543	0.3500	0.6027	0.7958
	B	0.7469	0.5224	0.2627	0.0932
	C	0.0988	0.1726	0.1346	0.1110
Deng's method	A	0.1543	0.4861	0.7773	0.8909
	B	0.7469	0.3481	0.0628	0.0086
	C	0.0988	0.1657	0.1600	0.1005
Tazid's method	A	0.4924	0.7016	0.8075	0.5588
	B	0.0051	0.0059	0.0068	0.1586
	C	0.5025	0.2925	0.1857	0.2816
Our method	A	0.1632	0.4989	0.8207	0.9135
	B	0.7126	0.3287	0.0519	0.0042
	C	0.1242	0.1724	0.1274	0.0823

C. Experiment Analysis

- **ZI comparison**

By intuition, the support degree of the 5 evidences in table II for proposition A is larger, and the proposition A shall be the final conclusion. The confidence of proposition A assigned by different methods is shown in figure 2.

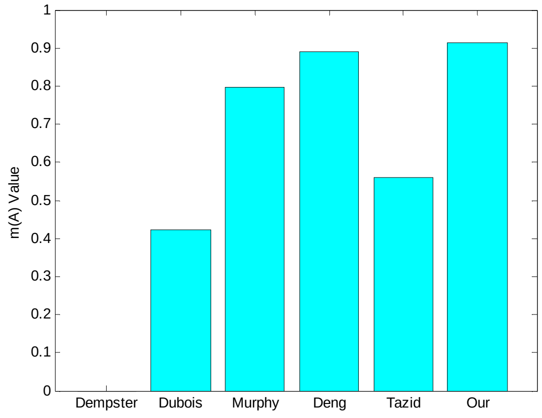

Figure 2. M (A) comparison in different fusion methods.

It can be seen from table III that Dempster's combination rule will obtain unreasonable results when the evidence is conflicting, and it cannot reflect the experts' real opinion. The value of $m(A)$ is always 0 after fusion. Because the supporting degree of evidence 2 to target A is 0, though the later evidences support target A, the final fusion conclusion is not A. This is obviously unreasonable, the reason may be the interference of external environmental factors or human factors, resulting in sensor failure and the conclusion of the evidence 2 does not match with the actual situation. Murphy's method, Dubois's method, Deng's method, Tazid's method and the new method proposed in this paper are able to avoid the Zadeh's paradox, but the degree of support for the proposition A is different. The confidence of proposition A in the newly proposed method is the largest and its accuracy is the highest.

The comparison results of Zadeh Immunity in various combination methods are shown in table IV.

TABLE IV
ZI ATTRIBUTE COMPARISON IN DIFFERENT FUSION METHODS

Serial Number	Method	ZI Attribute
1	Dempster's method	
2	Dubois& Prade's method	★★
3	Murphy's method	★★★
4	Deng's method	★★★★
5	Tazid's method	★★
6	Our method	★★★★★

- **Discrimination comparison**

The sum discrimination and difference discrimination of different combination methods are compared in Figure 3 and Figure 4. In figure 3, the sum discrimination of Dempster's method, Murphy's method, Deng's method, Tazid's method and the new method proposed in this paper are 1. The main reason is that these methods only distribute confidence to single focal element when conflict confidence is redistributed. Therefore, their sum discrimination values are higher. Because the Dubois's method assigns conflict confidence to union space, its sum discrimination value is relatively low, only 0.5398. In Figure 4, the difference discrimination of the newly proposed method is the highest, reaching 0.8312, and its support degree of proposition A is the strongest.

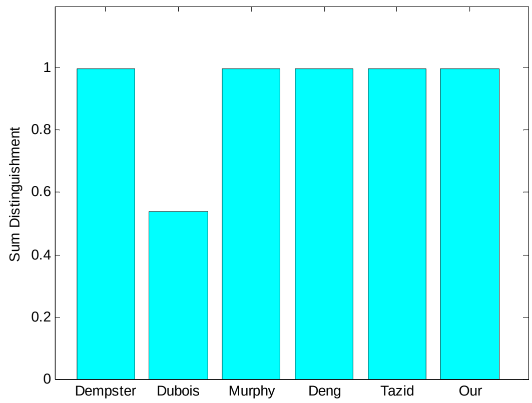

Figure 3. Sum discrimination comparison in different fusion methods.

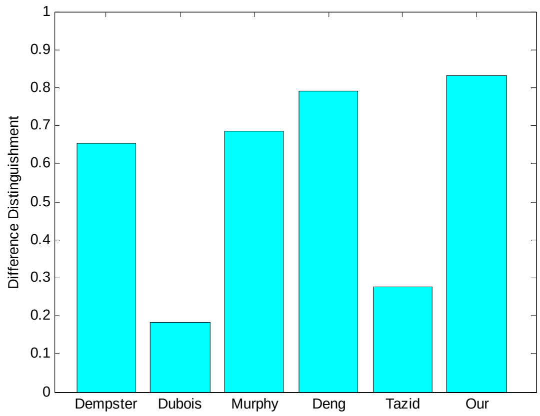

Figure 4. Difference discrimination comparison in different fusion methods.

- **Fusion Efficiency Comparison**

With the increasing of evidence, Murphy's averaging method, Deng's distance method and the new method proposed in this paper can fuse A effectively. But Murphy's method does not consider the correlation between evidences. Only when the fourth evidence reaches, the target A can be identified by this method. As Tazid's method using the additive strategies, its convergence rate of fusion is slow and it is not convenient to get the final decision. When the third evidence is collected, Deng's method and the proposed method can identify the target A effectively, but the proposed method is more sensitive to the conflict of evidences, and its convergence rate is faster. This is mainly due to the use of dissimilarity matrix to measure the dissimilarity between evidences. Thus, the interconnectedness between focal elements attributes and evidences is fully taken into account. It can reduce the confidence and the weight of interference evidence. What is more, it can reduce the impact of "interference evidence" to final fusion results effectively and improve the efficiency and performance of the fusion.

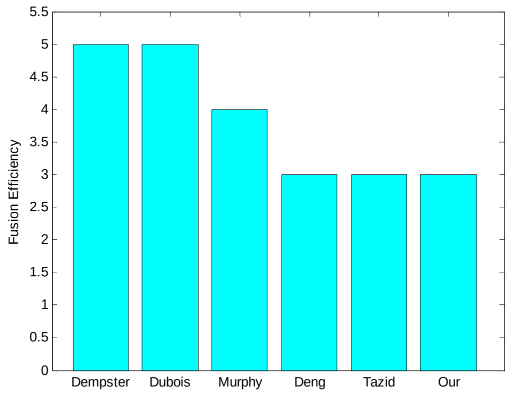

Figure 4. Fusion efficiency comparison in different fusion methods.

VI. CONCLUSIONS

With the increasing of system scale, all kinds of alarm information and log grow in an order of magnitude speed. How to analysis and integrate these huge, simple and high redundancy heterogeneous information has become an urgent problem to be resolved. Evidence theory is a theory of uncertain reasoning, and it has been widely used in the field of information fusion. When the synthesis of evidence is in a high conflict, there will be paradox in the classical theory evidence. In order to resolve this problem, a multi-source information fusion method based on dissimilarity matrix is proposed. First, using the weighted Euclidean distance, evidence dissimilarity matrix is constructed. Second, the dissimilarity of between the evidences is measured. Third, through calculating the supporting degree, credibility and weight of evidence by using dissimilarity matrix, the original evidences are modified. Finally, the improved combination rule is used for information fusion. Experimental results show that the proposed method is superior to the existing typical method in accuracy, discrimination and fusion efficiency of combination results.

ACKNOWLEDGMENT

This work has been supported by the National High Technology Research and Development 863 Program of China under Grant No. 2012AA012704.

REFERENCES

[1] M. Shafiqul Islam, Amin Zargar, Roberta Dyck, Asish Mohapatra, Rehan Sadiq, "Data fusion-based risk assessment framework: an example of benzene," *Int J Syst Assur Eng Manag,* vol.3, pp.267-283, Oct-Dec 2012.

[2] Li Yan-na, Qiao Xiu-quan, Li Xiao-feng, "An Uncertain Context Ontology Modeling and Reasoning Approach Based on D-S Theory," *Journal of Electronics & Information Technology,* vol. 32, pp.1806-1811, 2010.

[3] Wang X, "Equivalence between Recursive and Analytical Evidential Reasoning Algorithms," *Journal of Software,* vol. 8, pp. 754-759, 2013.

[4] Jean Dezert, Pei Wang, Albena Tchamova. "On The Validity of Dempster-Shafer Theory," *15th International Conference on Information Fusion, Singapore,* pp.655-660, 2012.

[5] Feng Hai-shan, Xu Xiao-bin, Wen Cheng-lin, "A New Fusion Method of Conflicting Interval Evidence Based on the Similarity Measure of Evidence," J*ournal of Electronics & Information Technology,* vol. 34, pp.851-857, 2012.

[6] Li W, Liu P, Wang Y, "Early Flame Detection in Video Sequences based on DS Evidence Theory," *Journal of Computers,* vol. 8, pp.818-825, 2013.

[7] Andino Maseleno, Md. Mahmud Hasan, "The Dempster-Shafer Theory Algorithm and its Application to Insect Diseases Detection," *International Journal of Advanced Science and Technology,* vol. 50, pp.111-120, 2013.

[8] Faouzi SEBBAK, Abdelghani CHIBANI, Yacine AMIRAT, Farid BENHAMMADI, "An Evidential Fusion Approach for Activity Recognition under Uncertainty in Ambient Intelligence Environments," *UbiComp12, Pittsburgh, USA,* pp.834-840, 2012.

[9] E. Pashaa, H.R. Mostafaeib, M. Khalaj, F. Khalaj, "Fault Diagnosis of Engine Using Information Fusion Based on Dempster-Shafer Theory," *J. Basic. Appl. Sci. Res.,* vol. 2, pp.1078-1085, 2012.

[10] JIA R, SUN H, ZHANG C, "A New Safety Evaluation Model of Coal Mine Roof based on Multi-sensor Fusion in case of Information Confliction," *Journal of Computers,* vol. 7, pp. 499-506, 2012.

[11] Rashaad E. T. Jones, Erik S. Connors, and Mica R. Endsley, "Incorporating the Human Analyst into the Data Fusion Process by Modeling Situation Awareness Using Fuzzy Cognitive Maps," *12th International Conference on Information Fusion,* pp.1265-1271, 2012.

[12] Yager R R, "On the fusion of imprecise uncertainty measures using belief structures," *Information Sciences,* vol. 181, pp.3199-3209, 2011.

[13] LIANG Chang-yong, YE Chun-miao, ZHANG En-qiao, "An Evidence Combination Method based on Consistence of Conflict," *Chinese Journal of Management Science,* vol. 18, pp.152-156, 2010.

[14] Y. Deng, W. K. Shi, Z. F. Zhu and Q. Liu, "Combining belief functions based on distance of evidence," *Decision support systems,* vol. 38, pp.489-493, 2004.

[15] Tazid Ali, Palash Dutta, Hrishikesh Boruah, "A New Combination Rule for Conflict Problem of Dempster-Shafer Evidence Theory," *International Journal of Energy, Information and Communications,* vol. 3, pp.35-40, 2012.

[16] Yee Leung, Nan-Nan Ji, Jiang-Hong Ma, "An integrated information fusion approach based on the theory of evidence and group decision-making," *Information Fusion,* vol. 8, pp.1-13, 2012.

[17] HU Chang-hua, SI Xiao-sheng, ZHOU Zhi-jie, "An Improved D-S Algorithm Under the New Measure Criteria of Evidence Conflict," *Acta Electronica Sinica,* vol. 37, pp. 1578-1583, 2009.

[18] FUPING ZENG, MANYAN LU, DEMING ZHONG, "USING D-S EVIDENCE THEORY TO EVALUATION OF CONFIDENCE IN SAFETY CASE," *Journal of Theoretical and Applied Information Technology,* vol. 47, pp. 184-189, 2013.

[19] Chiara Foglietta, Andrea Gasparri, Stefano Panzieri, “A Networked Evidence Theory Framework for Critical Infrastructure Modeling,” *IFIP Advances in Information and Communication Technology*, vol. 390, pp. 205-215. 2012.
[20] Wen Jiang, Deqiang Han, Xin Fan, Dejie Duanmu, “Research on Threat Assessment Based on Dempster–Shafer Evidence Theory,” *Green Communications and Networks*, vol. 2, pp.975-984, 2012.
[21] Andino Maseleno, Md. Mahmud Hasan, “The Dempster-Shafer Theory Algorithm and its Application to Insect Diseases Detection,” *International Journal of Advanced Science and Technology*, vol. 50, pp. 111–120, 2013.
[22] YANG Feng-bao, WANG Xiao-xia, *Combination of Conflict for D-S Evidence Theory*. Beijing: National Defense Industry Press, 2010.

Yongwei Wang was born at Luoyang, China in 1977. He has studied at Zhengzhou Information Science and Technology Institute, Zhengzhou, China and received his master degree in 2004. He is working on his PhD thesis about information fusion. His interests of research are information security, computer network and situation awareness.

He is an associate professor at Zhengzhou Information Science and Technology Institute, Zhengzhou, China.

His hobbies are sports, photography and travel.

Kaiguo Yuan was born in Guizhou, China in 1982. He received the Ph.D degree in signal and information processing from Beijing University of Posts and Telecommunications in 2009.

Dr. Yuan is working in information security as a lecturer in Beijing University of Posts and Telecommunications.

Yunan Liu was born at Zhoukou, China in 1971. He has studied at the University of Zhengzhou, Zhengzhou, China and received his master degree in 2003. His research focus is on information security and information fusion.

He is a professor at Zhengzhou Information Science and Technology Institute, Zhengzhou, China.

His hobbies are travel, the game of Go, photography, gardening and Chinese chess.

Hongyong Jia was born at Zhumadian, China in 1975. He has studied at Beijing University of Posts and Telecommunication, Beijing, China and received his PhD degree in 2009. His research focus is on information security and information fusion.

He is an associate professor at Zhengzhou Information Science and Technology Institute, Zhengzhou, China.

Wei Qiu was born at Yichun, Zhengzhou, China in 1992. His research focus is on information fusion and situational awareness.

He is a student at Zhengzhou Information Science and Technology Institute, Zhengzhou, China.

..