TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 02/14 pp12–19
DOI: 10.26599/TST.2018.9010138
Volume 25, Number 1, February 2020

Data Fusion Algorithm Based on Fuzzy Sets and D-S Theory of Evidence

Guangzhe Zhao, Aiguo Chen*, Guangxi Lu, and Wei Liu

Abstract: In cyber-physical systems, multidimensional data fusion is an important method to achieve comprehensive evaluation decisions and reduce data redundancy. In this paper, a data fusion algorithm based on fuzzy set theory and Dempster-Shafer (D-S) evidence theory is proposed to overcome the shortcomings of the existing decision-layer multidimensional data fusion algorithms. The basic probability distribution of evidence is determined based on fuzzy set theory and attribute weights, and the data fusion of attribute evidence is combined with the credibility of sensor nodes in a cyber-physical systems network. Experimental analysis shows that the proposed method has obvious advantages in the degree of the differentiation of the results.

Key words: data fusion; fuzzy sets; Dempster-Shafer (D-S) theory

1 Introduction

Cyber-Physical Systems (CPSs) are receiving increased attention because of their wide applications, although the theories and technologies of CPSs still face considerable challenges, including energy management, privacy and security[1–3], data transmission and processing, and control technique. Data processing is one of the most important factor in CPSs. In a typical scenario, a large number of heterogeneous nodes are deployed to monitor the surrounding environment comprehensively and in a timely manner. Various types of data are collected at a high frequency[4]. The quality of decision making of CPSs is related to the accuracy of high-dimensional big data and the transmission rate of this accurate data[5]. In Refs. [6, 7], researchers used a small data set to represent the vast information carried by big sensory data. In Refs. [8–10], approximate aggregation algorithms for different applications are proposed. All these methods aim to improve transmission and processing performance by reducing data size.

From the point of view of improving accuracy, data fusion is used to combine data from multiple sources to achieve inferences, correlations, and associations that are more efficient and potentially more accurate than if they were achieved by means of a single source[11]. In the process of data fusion, different services have different sensitivities to different attribute of records[12]. For example, service A in the control process could particularly be influenced by temperature. Therefore, when judging the influence degree of the environmental status on service A, temperature attribute data will be given a greater weight in data fusion process. Meanwhile, another service B is affected by wind speed in the control process, but the impact of temperature changes is negligible. Then, the attribute weight of temperature should be reduced, and the attribute weight of wind speed should be increased. Under the same environmental conditions, with the adjustments of weights, the degrees of the impact of environmental changes on things will be different in the final fusion results. Therefore, how to scientifically and reasonably optimize the weight of each attribute is one of the key issues in data fusion.

• Guangzhe Zhao is with Beijing University of Civil Engineering and Architecture, Beijing 102616, China. E-mail: zhaoguangzhe@bucea.edu.cn.
• Aiguo Chen, Guangxi Lu, and Wei Liu are with the University of Electronic Science and Technology of China, Chengdu 611731, China. E-mail: agchen@uestc.edu.cn; m18328401268@163.com; uestcliuwei@163.com.
∗ To whom correspondence should be addressed.
Manuscript received: 2018-06-18; accepted: 2018-11-15

2 Related Work

The concept of data fusion was first proposed by the Joint Directors of Laboratories of the United States Department of Defense. It was defined as a process dealing with the association, correlation, and combination of data and information from single and multiple sources to achieve precise position and identity estimates and complete and timely assessments of situations, threats, and their significance[13]. Data fusion was originally applied to the military field to comprehensively evaluate a battlefield through various data. Later, it was also widely used in civilian fields. In addition to having decision-making ability, data fusion can also reduce data redundancy and improve the accuracy and efficiency of data acquisition. The main research of data fusion is focused on the algorithm[14], which could be divided into two categories.

One is the single-dimensional attribute data fusion method. All data belong to the same kind of attribute, while the locations of the data sources are different. The data of multiple identical attributes are finally synthesized into a representative feature data. For example, in Ref. [15], a tree-structure-based fusion method is proposed to construct a dynamic network transmission tree. Data are transmitted along the branches to the root. During the transmission process, data are collected at each convergence point, and the data transmission volume is gradually reduced. Reference [16] proposes a fusion algorithm from the perspective of the time correlation of data acquisition, which is used to quantify the data fusion frequency and reduce the load. In addition, prediction-based and compression-based data fusion methods have been proposed in Ref. [17].

The other category is the method of multidimensional attribute data fusion method, which is often used in decision-layer fusion to fuse the recognition results of multiple classification models for measurement targets, and it then executes the final reasoning and decision-making process. Dempster-Shafer (D-S) evidence theory is commonly used in decision-layer data fusion. However, evidence conflict problems are prone to occur, such as the one-vote denial situation. That is, when some evidence is completely denied, no matter how high the degree of the other evidence is, it will be integrated to obtain a negative result. Two main solutions to this problem are presented.

(1) Rule of modifying evidence synthesis. The domain of the synthesis rule is believed to be incomplete, which leads to evidence conflicts. Reference [13] redistributes the conflicting information. For the evidence without a conflict, the basic D-S evidence theory synthesis rule is still adopted, the conflicting evidence is completely denied, and the conflicting evidence is assigned to unknown domain terms. Reference [18] further improved this process by stating that the conflicting evidence also carries support for the original domain and should not be placed into unknown items. Through the introduction of the concept of evidence credibility, the conflicting evidence is weighted and redistributed.

(2) Rule of modifying evidence data. It modifies the basic probabilities of the evidences support to the resulting events and then uses D-S evidence theory synthesis rules to synthesize evidence and reduce or even eliminate the evidence conflict by changing support. The typical Murphy's average distribution method averages the basic probability distributions of n evidence groups and then uses the basic D-S evidence synthesis rules to synthesize[19]. The advantage of this method is that it can converge quickly, and it can also handle conflicting data. However, Murphy's method simply calculates the averages without considering the correlations between the evidence. It will provide evidence of the great impact of large deviations on the final synthesis results. On the basis of Murphy's method, some scholars use the distance between evidence to measure the similarity between evidence to resolve the correlations.

The modification of the synthesis rule changes some basic properties of the synthesis rule, and the modified synthesis rule changes the commutation law and the binding law of the original synthesis rule. When more than one piece of evidence exists, the final result is related to the order of the synthetic evidence, and an unknown domain term that undermines the closure of the domain set is introduced. To modify the evidence data, starting with the evidence data themselves, the reliability of the evidence is derived from the mathematical data provided by all the evidence without considering the actual application. For example, in CPSs, the evidence data are obtained from various sensors, while the sensor itself has a reliability problem in providing data.

3 Proposed Data Fusion Algorithm

A data fusion algorithm based on fuzzy set theory and D-S evidence theory is proposed here, which combines the weights of environmental attributes and the credibility of CPSs nodes.

First, in accordance with fuzzy set theory, we can determine the degree of membership of an environmental factor to an abnormal level by changing the weight of the environmental attributes. Then, we use fuzzy set theory to determine the initial basic support of a numerical value of environmental attributes and adjust the support based on the weights of the attributes. Next, the membership degree is normalized to obtain the basic probability distribution in D-S evidence theory.

Then, according to D-S evidence theory, data fusion is combined with the credibility of attributes. The credibility of the attribute is the credibility of the sensor node that provides this attribute. At present, D-S evidence theory is mainly based on the evidence distance to calculate the similarity between the evidence. Without combining specific applications, the credibility of evidence is judged based on the source of the evidence.

In this process, the degree of membership is adjusted by the weight of the attribute. Therefore, the value of same abnormal data to different abnormal levels is different, and the final synthesis result will also be different. Considering the credibility of the source node of evidence, the accuracy of the synthesis result could be improved.

3.1 Basic probability determination method based on fuzzy sets and attribute weights

Assume that we need to determine the level of an abnormal environment according to the values of abnormal attributes. We define the set of attributes as $E = \{E_1, E_2, \ldots, E_n\}$ and set the abnormal level as $A = \{A_1, A_2, \ldots, A_m\}$. A_i represents the i-th abnormal level, and it is an integer. First, we need to use fuzzy set theory to determine the attributes membership degree to all abnormal levels based on the attribute value. Various attributes are the universe of the fuzzy set theory. The abnormal level A_i in CPSs is a fuzzy set. When the value of an attribute is normal, the degree of support for all exception levels is 0. When the value of an attribute is abnormal, the degree of support for all abnormal levels needs to be determined, that is, we need to determine a certain abnormal degree of membership to which the anomaly attribute belongs.

For any exception level A_i, each attribute E_j has a value E_{ji}. When the value of attribute E_j is not in the range $[E_{ji}^{\min}, E_{ji}^{\max}]$, the support probability of E_j to the exception class A_i is 0. Otherwise, a certain probability that it belongs to the abnormal-level A_i exists. The support degree for determining a property for a fuzzy set is determined by the membership function. When the value of attribute E_j is closer to the middle of the range, a high support probability of attribute E_j to the abnormal level A_i corresponds to a high value of the membership degree function. When the value of attribute E_j is closer to the two sides of the range, a low support probability of attribute E_j to the abnormal level A_i corresponds to a low value of the membership degree function. So, the membership curve, which has the characteristics of high in the middle and low on both sides, can be represented by Gaussian membership functions in Eq. (1),

$$u_{A_i}(E_j) = \begin{cases} \mathrm{e}^{\frac{-(E_{ji}-\mu_{ji})^2}{2{\sigma_{ji}}^2}}, & E_{ji} \in [E_{ji}^{\min}, E_{ji}^{\max}]; \\ 0, & E_{ji} \notin [E_{ji}^{\min}, E_{ji}^{\max}] \end{cases} \quad (1)$$

where the value of expectation μ_{ji} is the average of $E_{ji}^{\min}$ and $E_{ji}^{\max}$. σ_{ji} is the standard deviation of Gaussian function, whose value is $(E_{ji}^{\max} - \mu_{ji})/3$. According to Eq. (1), the attribute corresponding exception can be calculated. According to the degree of membership, a membership matrix U can be proposed in Eq. (2),

$$U = \begin{bmatrix} u_{A_1}(E_1) & u_{A_2}(E_1) & \cdots & u_{A_m}(E_1) \\ u_{A_1}(E_2) & u_{A_2}(E_2) & \cdots & u_{A_m}(E_2) \\ u_{A_1}(E_3) & u_{A_2}(E_3) & \cdots & u_{A_m}(E_3) \\ \vdots & \vdots & \ddots & \vdots \\ u_{A_1}(E_m) & u_{A_2}(E_m) & \cdots & u_{A_m}(E_m) \end{bmatrix} \quad (2)$$

According to the detection attribute value, the degree of membership of each attribute to abnormal level can be initially determined by the membership degree function. The weight of each attribute is considered the same. No distinction exists between the weights of attributes. However, an attribute will belong to multiple abnormal levels with different degrees of membership. When the weight of this attribute is high, then the result of the abnormal level is affected by this attribute. When this abnormal property has an abnormality, then a high probability exists that the abnormal level is high. Therefore, the degree of membership needs to be adjusted according to the weight.

According to this analysis, the basic adjustment criteria is based on the weight of attributes. When the weight of an attribute is higher than the average weight, the degree of membership with a high level of exceptional support for this attribute needs to be increased, and the degree of membership with a low level of support for the anomaly needs to be reduced. When the weight of an attribute is lower than the average weight, the weight is adjusted in the opposite direction. The abnormal level is divided according to the set of abnormal levels with a degree of membership greater than 0. The rank of the abnormality whose attribute degree is greater than 0 is divided into two parts by the middle average, namely, the part with the high abnormal level and the part with the lower abnormal level. Taking the middle as the benchmark, the farther the abnormal level is from the benchmark, the smaller the proportion of adjustment is, and the closer it is to the elevation, the greater the adjustment ratio. When the weight of an attribute is greater than the average weight, the degree of membership with a high level of attribute support anomaly increases, and the degree of membership with a low level of anomaly support decreases.

When the weight of an attribute is less than the average weight, the degree of membership with a low level of attribute support anomaly increases, and the degree of membership with a low level of abnormality is reduced.

Therefore, the adjustment of an attribute relative to an anomaly grade membership degree needs to be determined based on the relationship between attributes weight and average weight, and the distance between anomaly grade and the anomaly rank classification benchmarking. The weight of attribute E_j is represented by w_j. Then, the total weight w and the average weight $\overline{w}$ could be calculated by Eqs. (3) and (4), respectively,

$$w = w_1|w_2|w_3|\cdots|w_n = \sum_{j=1}^{n} w_j \tag{3}$$

$$\overline{w} = \frac{(w_1|w_2|w_3|\cdots|w_n)}{n} = \frac{\sum_{i=1}^{n} w_i}{n} \tag{4}$$

Assume that attribute E_j calculates the abnormality membership degree greater than zero set according to the previous membership function $\{A_a, A_{a+1}, \ldots, A_b\}$. The minimum value of the abnormal level is A_a, and the maximum value of the abnormal level is A_b. The level of abnormality is divided into the middle of the benchmark A_{mid} as follows:

$$A_{\text{mid}} = \frac{A_a + A_b}{2} \tag{5}$$

Therefore, the distance from an exception level to the benchmark can be calculated in Eq. (6),

$$D_{A_i} = |A_i - A_{\text{mid}}| \tag{6}$$

The total distance from the anomaly level collection to the benchmark equals the total distance from the low anomaly level collection to the benchmark. Therefore, we have

$$D = D_{A_i} + D_{A_{i+1}} + \cdots + D_{A_b}, \ \ A_i \geqslant A_{\text{mid}} \tag{7}$$

For a membership value $u_{A_i}(E_j)$, the adjustment equation is shown in Eq. (8),

$$u'_{A_i}(E_j) = \begin{cases} u_{A_i}(E_j) - (\frac{w_j - \overline{w}}{w})(1 - \frac{D_{A_i}}{D})u_{A_i}(E_j), & A_i < A_{\text{mid}}; \\ u_{A_i}(E_j), & A_i = A_{\text{mid}}; \\ u_{A_i}(E_j) + (\frac{w_j - \overline{w}}{w})(1 - \frac{D_{A_i}}{D})u_{A_i}(E_j), & A_i > A_{\text{mid}} \end{cases} \tag{8}$$

The first expression in brackets is the attribute weights. This expression indicates whether the attribute weight E_j is higher or lower than the average weight value. A value of positive indicates that the weight is higher than the average weight, and a value of negative indicates that the weight is lower than the average weight; And its absolute value represents a proportion that exceeds or falls below the weight. The second parenthesis is a function of abnormal grade distance, which indicates the proportion of an abnormal grade distance to the total distance. A low proportion corresponds to a high percentage of adjustment. A low ratio corresponds to a great proportion of adjustment. Therefore, we should use 1 to reduce the distance ratio by using Eq. (8).

The membership degree is adjusted by Eq. (8). When the attribute weight is higher than the average, the degree of membership of the low abnormal level is reduced, and the degree of membership of the high abnormal level is increased. According to the adjusted degree of membership, a new membership matrix U_{new} is computed by Eq. (9),

$$U_{\text{new}} = \begin{bmatrix} u'_{A_1}(E_1) & u'_{A_2}(E_1) & \cdots & u'_{A_m}(E_1) \\ u'_{A_1}(E_2) & u'_{A_2}(E_2) & \cdots & u'_{A_m}(E_2) \\ u'_{A_1}(E_3) & u'_{A_2}(E_3) & \cdots & u'_{A_m}(E_3) \\ \vdots & \vdots & \ddots & \vdots \\ u'_{A_1}(E_m) & u'_{A_2}(E_m) & \cdots & u'_{A_m}(E_m) \end{bmatrix} \tag{9}$$

To obtain the basic probability distribution in D-S evidence theory, the corresponding degree of membership of each evidence needs to be normalized according to Eq. (10),

$$P_{A_i}(E_j) = \frac{u'_{A_i}(E_j)}{\sum\limits_{i=1}^{m} u'_{A_i}(E_j)} \tag{10}$$

$P_{A_i}(E_j)$ represents the basic probability that E_j is allocated to the abnormal A_i. After the membership degree is normalized, the basic probability distribution of the evidence to the abnormal level can be obtained and extracted to obtain a basic probability distribution matrix P in Eq. (11),

$$P = \begin{bmatrix} p_{A_1}(E_1) & p_{A_2}(E_1) & \cdots & p_{A_m}(E_1) \\ p_{A_1}(E_2) & p_{A_2}(E_2) & \cdots & p_{A_m}(E_2) \\ p_{A_1}(E_3) & p_{A_2}(E_3) & \cdots & p_{A_m}(E_3) \\ \vdots & \vdots & \ddots & \vdots \\ p_{A_1}(E_m) & p_{A_2}(E_m) & \cdots & p_{A_m}(E_m) \end{bmatrix} \tag{11}$$

3.2 D-S evidence synthesis method based on node data credibility

The existing methods mostly establish a reliability model of a sensor based on the characteristics of the sensor itself, such as the length of time the sensor is used, the remaining energy, and the like. This method is flawed in CPSs because malicious attacks may occur. Regardless of the sensor's own characteristics or external interference, the results caused by them are only reflected in the value of the sensor data. So it's necessary to establish the reliability of a sensor's data from the correct angle of the historical data provided by the sensor. D-S evidence synthesis is performed by using the credibility as the weight of the data.

3.2.1 Node reliability based on sliding window

The data collected for each sensor will be based on the anomaly detection process of the data's space-time attributes. During this process, the process will initially detect whether the data provided by this sensor are correct. From the historical data provided by the sensor and the correctness of the data collected at the current moment, a sensor reliability model is established in the data sliding window mode. The model includes a historical data queue, the correct data, the error data, and a sliding window, as shown in Fig. 1.

The historical data queue is a sensor that has collected and transmitted data in chronological order, and the data passed the anomaly detection process and can distinguish between the correct data and wrong data. The sliding window is a virtual window, and the window size can be adjusted according to the current window. The ratio of correct data to the window size is used to calculate the credibility of the data. The length of the sliding window is denoted as L, and the number of correct data in the sliding window is denoted as N. The credibility of the current data C, is calculated by Eq. (12),

$$C = \frac{N}{L} \tag{12}$$

When a sensor transmits the wrong data, the reliability of this sensor should be reduced, and multiple consecutive correct values are needed to gradually increase the reliability of the sensor. The sliding window is used to describe this process, which is reflected in the data filling. If erroneous data are again generated in the process of increasing the credibility, then the credibility will be reduced again, and the process of increasing credibility will be prolonged. That is, wrong data appear during the filling process, and the size of the sliding window needs to be increased to lengthen the filling process. When the continuous acquisition of correct data reaches a certain standard, the influence of eliminating historical error data can be achieved. Then the sliding window should be canceled, and the reliability of the sensor becomes a stable value.

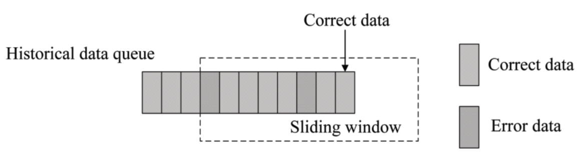

Fig. 1 Sliding window with correct data and error data.

3.2.2 D-S evidence synthesis process

The value of an attribute evidence is the data after the fusion of the results of similar sensors. The average of the credibility of the fusion source sensor is taken as the credibility of this fusion data. If there are n E_j type sensors, acording to the sensor credibility model, these sensors have a certain degree of confidence in the collected values, which are $C_{E_j}^1, C_{E_j}^2, C_{E_j}^3, \ldots, C_{E_j}^n$, respectively. Then, the credibility C_{E_j} of the attribute E_j evidence can be calculated by Eq. (13),

$$C_{E_j} = \frac{C_{E_j}^1 + C_{E_j}^2 + \cdots + C_{E_j}^n}{n} \tag{13}$$

According to the above method, each attribute evidence can be calculated a credibility, which can be regarded as the weight of each evidence. After the weights of each attribute are obtained, the weighted sum of the basic distribution probabilities of all attribute evidences is calculated according to the weights of attributes for each exception level.

$$p_{A_i}(E) = \sum_{j=1}^{n}(p_{A_i}(E_j) \times C_{E_j}) \tag{14}$$

After weighting, the weighted sum obtained above becomes a one-dimensional matrix P in Eq. (15),

$$P = [p_{A_1}(E), p_{A_2}(E), \ldots, p_{A_m}(E)] \tag{15}$$

The element $p_{A_i}(E)$ represents a support degree of evidence set E to abnormal level A_i. Then, it is normalized according to Eq. (16),

$$p'_{A_i}(E) = \frac{p_{A_i}(E)}{\sum\limits_{i=1}^{m} p_{A_i}(E)} \tag{16}$$

the basic probability distribution that is in accordance with the requirement of D-S evidence theory is obtained.

Therefore, we obtain the weighted basic probability distribution matrix in Eq. (17),

$$P' = [p'_{A_1}(E), p'_{A_2}(E), \ldots, p'_{A_m}(E)] \tag{17}$$

Then, on the basis of the D-S evidence, basic theoretical equations are synthesized in Eq. (18),

$$m(A) = \begin{cases} \frac{1}{1-k_{A_i,B_j}} \sum\limits_{A_i \cap B_j = A} m_1(A_i)m_2(B_j), & A \neq \varnothing; \\ 0, & A = \varnothing \end{cases} \tag{18}$$

Among them, m_1 and m_2 represent two pieces of evidence. k represents the degree of conflict between these two pieces of evidence and is calculated by Eq. (19),

$$k = \sum_{A_i \cap B_j = \varnothing} m_1(A_i)m_2(B_j) \tag{19}$$

According to Murphy's improved method for conflict evidence of D-S evidence theory, the weighted evidence is used to obtain the weighted basic probability distribution. The D-S evidence synthesis equation is used to assign the probability after the addition. If n pieces of evidence exists, then the addition after the probability of synthesis occurs $n - 1$ times. First, the weighted basic probability distribution is synthesized by itself according to the synthesis equation, and the result P_1 of the first synthesis can be obtained. Then, the result P_1 of the first synthesis and the synthesis equation are used to synthesize the result P_2 after the synthesis of two times and, in turn, recursively. n pieces of evidence exist on the synthesis that occurs $n - 1$ times.

3.3 Multidimensional attribute data fusion

After synthesizing the rule operation of D-S evidence theory, each attribute has a support probability for an abnormal grade. All attribute evidence is synthesized, and all attributes have a comprehensive support probability for the abnormal grade. Assuming that the results are expressed as $m(A_i)$, the meaning of $m(A_i)$ represents a comprehensive support probability for all attributes for the exception level A_i, as shown in Table 1. A new membership degree matrix can be obtained by the membership degree.

The synthesis result is that each abnormal level has a support degree, and an abnormal level needs to be selected from them as the final data fusion result according to judgment rule. For example, suppose that A_1 and A_2 are two elements that satisfy Eq. (20),

$$\begin{cases} m(A_1) = \max\{m(A_i), A_i \in A\}; \\ m(A_2) = \max\{m(A_i), A_i \in A, A_i \neq A_1\} \end{cases} \tag{20}$$

If $m(A_1) - m(A_2) > \varepsilon$, then the exception level A_1 is the final fusion result. ε is a preset distinguishing threshold.

The rule of judgment is to select the maximum probability and the second largest probability to compare in the abnormal result level. If the probability difference between them is greater than the preset value, the maximum probability corresponding to the grade is the final fusion result. If the probability difference between the two is less than the preset value, then the rank result corresponding to the maximum probability and the rank result corresponding to the second largest probability have no discrimination degree and no final fusion result.

4 Experimental

Assume that four kinds of environmental attributes, named E_1, E_2, E_3, and E_4, exist. Three kinds of environmental abnormal levels exist, named A_1, A_2, and A_3. D-S evidence synthesis is required for the values of the four environmental attributes. The initial basic probability distribution is shown in Table 2.

For the above basic probability distribution, the D-S synthesis rule, Murphy's average combination rule, and the proposed method were used for evidence synthesis. In our method, the reliability of nodes is

Table 1 Form of D-S evidence synthesis.

Level	Comprehensive support probability
A_1	$m(A_1)$
A_2	$m(A_2)$
$\vdots$	$\vdots$
A_m	$m(A_m)$

Table 2 Basic probability distribution.

Evidence	Level		
	A_1	A_2	A_3
E_1	0.60	0.12	0.28
E_2	0	0.70	0.30
E_3	0.65	0.25	0.10
E_4	0.55	0.32	0.13

taken into account. In the experiments, we specify that the confidence values of E_1, E_2, E_3, and E_4 are 0.4, 0.2, 0.9, and 0.7, respectively. The results shown in Table 3 were obtained.

As seen in Table 3, given that evidence E_2 does not support A_1 at all, in all the D-S synthesis rules, no matter what the other evidence is, the probability of $m(A_1)$ is always 0. The other three methods can handle this problem. The one-vote veto phenomenon leads to the following case. The Murphy average combination rule simply averages the evidence for weighted sums. The processing is relatively simple. No weight information is discriminated. The distance-based method degenerates into Murphy's weighted average when only two pieces of evidence are processed. If only two pieces of evidence exist, then the distance is not differentiated, and no connection with practical applications occurs. If more evidence of low credibility exists, then the results that are different from the actual results will be obtained. This paper starts from the practical application and uses the credibility of the sensor provider as the weight to process the evidence, which is more in line with the actual situation. The maximum probability and the second largest probability also have a certain distance when the number of evidence is small and can be judged according to the decision. Therefore, we come to the final fusion result.

Table 3 Comparison of the results of different methods.

Method	Evidence		
	E_1, E_2	E_1, E_2, E_3	E_1, E_2, E_3, E_4
D-S synthesis rule	$m(A_1) = 0.00$ $m(A_2) = 0.50$ $m(A_3) = 0.50$	$m(A_1) = 0.00$ $m(A_2) = 0.71$ $m(A_3) = 0.29$	$m(A_1) = 0.00$ $m(A_2) = 0.86$ $m(A_3) = 0.14$
Murphy's average combination rule	$m(A_1) = 0.26$ $m(A_2) = 0.49$ $m(A_3) = 0.25$	$m(A_1) = 0.56$ $m(A_2) = 0.35$ $m(A_3) = 0.09$	$m(A_1) = 0.72$ $m(A_2) = 0.25$ $m(A_3) = 0.03$
Our method	$m(A_1) = 0.47$ $m(A_2) = 0.29$ $m(A_3) = 0.24$	$m(A_1) = 0.86$ $m(A_2) = 0.11$ $m(A_3) = 0.03$	$m(A_1) = 0.92$ $m(A_2) = 0.07$ $m(A_3) = 0.01$

5 Conclusion

In this paper, we analyze some existing defects in data fusion algorithms of decision-making layer. Then, a data fusion method based on fuzzy sets and D-S evidence theory is proposed. The basic probability distribution of evidence is determined based on fuzzy set theory and attribute weights. Then, the data fusion of attribute evidence is combined with credibility of sensor node in CPSs. Furthermore, a method for modeling the reliability of sensors based on the correctness and error of historical data is presented. Finally, a simulation experiment is conducted on the credibility model to prove the validity of our model. Numerical analysis of the new synthesis method was performed. A comparison of the proposed method with other synthetic methods proved that the synthesis method based on node credibility is better. Thus, this method has an obvious advantage in terms of degree of differentiation of the results.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 61462089) and the Fundamental Research Funds for Beijing University of Civil Engineering and Architecture (No. X18002).

References

[1] Z. Cai and X. Zheng, A private and efficient mechanism for data uploading in smart cyber-physical systems, *IEEE Transactions on Network Science and Engineering,* doi:10.1109/TNSE.2018.2830307.

[2] X. Zheng, Z. Cai, J. Yu, C. Wang, and Y. Li, Follow but no track: Privacy preserved profile publishing in cyber-physical social systems, *IEEE Internet of Things Journal,* vol. 4, no. 6, pp. 1868–1878, 2017.

[3] G. Luo, K. Yan, X. Zheng, L. Tian, and Z. Cai, Preserving adjustable path privacy for task acquisition in mobile crowdsensing systems, *Information Sciences*, doi:10.1016/j.ins.2018.12.013.

[4] C. Aranzazu-Suescun and M. Cardei, Distributed algorithms for event reporting in mobile-sink WSNs for Internet of Things, *Tsinghua Science and Technology,* vol. 22, no. 4, pp. 413–426, 2017.

[5] Q. Meng, K. Wang, X. He, and M. Guo, QoE-driven big data management in pervasive edge computing environment, *Big Data Mining and Analytics,* vol. 1, no. 3, pp. 222–233, 2018.

[6] S. Cheng, Z. Cai, J. Li, and X. Fang, Drawing dominant dataset from big sensory data in wireless sensor networks, in *2015 IEEE Conference on Computer Communications (INFOCOM),* Hong Kong, China, 2015, pp.531–539.

[7] S. Cheng, Z. Cai, J. Li, and H. Gao, Extracting kernel dataset from big sensory data in wireless sensor networks, *IEEE Transactions on Knowledge and Data Engineering,* vol. 29, no. 4, pp. 813–827, 2017.

[8] J. Li, S. Cheng, Y. Li, and Z. Cai, Approximate holistic aggregation in wireless sensor networks, in *2015 IEEE 35th International Conference on Distributed Computing Systems,* Columbus, OH, USA, 2015, pp. 740–741.

[9] Z. He, Z. Cai, S. Cheng, and X. Wang, Approximate aggregation for tracking quantiles and range countings in wireless sensor networks, *Theoretical Computer Science,* vol. 607, pp. 381–390, 2015.

[10] S. Cheng, Z. Cai, and J. Li. Curve query processing in wireless sensor networks, *IEEE Transactions on Vehicular Technology,* vol. 64, no. 11, pp. 5198–5209, 2015.

[11] S. Chhabra and D. Singh, Data fusion and data aggregation/summarization techniques in WSNs: A review, *International Journal of Computer Applications,* vol. 121, no. 19, pp. 21–30, 2015.

[12] X. Ding and X. Guo, A survey of SNP data analysis, *Big Data Mining and Analytics,* vol. 1, no. 3, pp. 173–190, 2018.

[13] Handbook of Multisensor Data Fusion. Boca Raton, FL, USA: CRC Press, 2001.

[14] J. Xia, Y. Feng, L. Liu, and D. Liu, An information fusion model of innovation alliances based on the Bayesian network, *Tsinghua Science and Technology,* vol. 23, no. 3, pp. 347–356, 2018.

[15] W. Zhang and G. Cao, DCTC: Dynamic convoy tree-based collaboration for target tracking in sensor networks, *IEEE Transactions on Wireless Communications,* vol. 3, no. 5, pp. 1689–1701, 2004.

[16] M. A. Sharaf, J. Beaver, A. Labrinidis, and P. K. Chrysanthis, TiNA: A scheme for temporal coherency-aware in-network aggregation, in *Proceedings of the 3rd ACM International Workshop on Data Engineering for Wireless and Mobile Access — MobiDe 03,* San Diego, CA, USA, pp. 69–76, 2003.

[17] D. Petrovic, R. Shah, K. Ramchandran, and J. Rabaey, Data funneling: Routing with aggregation and compression for wireless sensor networks, in *Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications,* Anchorage, AK, USA, 2003, pp. 156–162.

[18] Q. Sun, Research on multi-sensor information fusion and multi-target tracking system, PhD dissertation, Zhejiang University, Hangzhou, China, 2000.

[19] C. K. Murphy, Combining belief functions when evidence conflicts, *Decision Support Systems,* vol. 29, no. 1, pp. 1–9, 2000.

Guangzhe Zhao received the PhD degree from Nagoya University, Japan in 2012. He is currently an associate professor with Beijing University of Civil Engineering and Architecture. His research interests include image processing and big data.

Aiguo Chen received the BS degree from the University of Electronic Science and Technology of China in 2004. He got the PhD degree in signal and information processing from Beijing University of Posts and Telecommunications, China in 2009. He was a visiting scholar in Arizona State University, USA from Jan. 2013 to Jan. 2014. He is currently an associate professor at the School of Computer Science and Engineering, University of Electronic Science and Technology of China, China. His main research interests include cloud computing, cyber-physical system, and big data.

Guangxi Lu received the BS degree from North China University of Technology in 2016. Currently, he is a PhD candidate majoring in software engineering at University of Electronic Science and Technology of China. His research interests include machine learning and deep neural networks.

Wei Liu received the BS and MS degrees in computer science from the University of Electronic Science and Technology of China in 2014 and 2017, respectively. His research interests include cyber-physical system and big data.