Journal of the Operational Research Society

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjor20

An adaptive evidence combination method for decision analysis under uncertainty

Xingli Wu & Huchang Liao

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2021.1993759

ORIGINAL ARTICLE

An adaptive evidence combination method for decision analysis under uncertainty

Xingli Wu and Huchang Liao

Sichuan University, Chengdu, China

ABSTRACT
Due to the imperfection of devices and the individuation of human cognition, the process of data fusion often involves uncertainty. Dempster–Shafer theory defines the basic probability assignments of possible hypotheses and is effective in combining uncertain information from multiple sources. However, the existing evidence combination methods lack the flexibility to achieve compensation between conflicting pieces of evidence. This study aims to propose an adaptive evidence combination method that takes into account the personalized compensation requirements of decision makers in solving problems of conflicting evidence. To achieve this, an adjustment coefficient is added to the basic probability assignment of each hypothesis to control the compensation degrees between conflicting pieces of evidence in a flexible manner. The parameters of information reliability and importance are further incorporated into the model. The algebraic properties of the proposed evidence combination method are described. In addition, we conduct two case studies, one on vehicle recognition based on multiple sensors and one on purchasing decisions based on online reviews. Through the sensitivity analysis of the adjustment coefficient and the comparative analysis with other evidence combination methods, the advantages of the proposed method in dealing with high levels of conflicting evidence are verified.

ARTICLE HISTORY
Received 22 February 2021
Accepted 8 October 2021

KEYWORDS
Data fusion; Dempster–Shafer theory; adaptive evidence combination; target recognition; purchasing decision

1. Introduction

Data fusion refers to an information processing method in which data or information obtained from single or multiple sources (e.g. devices, persons or timelines) is integrated to accomplish a decision-making task. Data fusion techniques have been widely used in the fields of defence, medicine, finance, geo-science and economy (Dezert & Smarandache, 2009) and proven to be useful for reducing uncertainties, improving reliability and enhancing recognition accuracy (Qi et al., 2019). There are four main categories of data fusion techniques, specifically probabilistic methods, statistical methods, knowledge-based theory and evidence theory (Qi et al., 2019). The probabilistic and statistical methods are based on large amounts of prior information and have a high degree of accuracy. By contrast, the knowledge-based theory requires the intervention of human expertise regardless of the amount of data and is easy to implement because of its simplicity. Due to the unreliability and instability of devices, limited human cognition, disagreement among human groups and inexplicit objects, practical data fusion problems usually involve uncertainties (e.g. conflicting, unreliable or incomplete information). Compared with the above three data fusion methods, the advantage of evidence theory lies in its ability to model and process uncertainties without prior information (Xiao, 2020).

Evidence theory, also known as Dempster–Shafer (D-S) theory (Dempster, 1967; Shafer, 1976), provides a mathematical framework for quantifying uncertainties through basic probability assignments (BPAs), which represent the probability distribution of hypotheses (Abellan & Bosse, 2020). As a generalization of Bayesian probability theory, D-S theory allocates a BPA to any subset of a frame of discernment or even to the universal set representing ignorance. In addition, the remaining support for a hypothesis does not always imply its negation. The D-S theory not only conforms to the associative and commutative law but also can generate fault-tolerant results (Xiao, 2020). Because of these advantages, D-S theory arises in many fields of application, such as target recognition and diagnosis (Agreh & Ghaffari-Hadigheh, 2019; Medjkoune et al., 2017), classifiers (Roy et al., 2021) and multi-criteria decision making (Fang et al., 2021).

Managing conflicts, that is, contradictory evidence from different sources, is difficult in the application of D-S theory (Sentz & Ferson, 2002). Dempster's (1967) rule is a traditional evidence combination method. It aggerates the BPAs of a hypothesis based on geometric operators. Because zero multiplied by any number is still zero, in the

CONTACT Huchang Liao liaohuchang@163.com Business School, Sichuan University, Chengdu 610064, China

case of highly conflicting evidence, the inference obtained using Dempster's rule is often counter-intuitive (Dezert & Smarandache, 2009). It does not support a hypothesis in the case of complete conflict even if there is a piece of evidence that strongly supports this hypothesis. To manage conflicts, a variety of evidence combination methods have been developed based on Dempster's rule. For example, Zhang (1994) proposed a centre combination rule by adding a weight factor to Dempster's rule to describe the intersection degrees between hypotheses. Dubois and Prade (1986) suggested combining the union of BPAs rather than their intersections and proposed a disjunctive pooling operation. Although this method can manage conflicts, the results are imprecise (Sentz & Ferson, 2002). As an extension of Dubois and Prade (1986) rule, Dezert and Smarandache (2009) theory offers an interesting fusion rule, namely the proportional conflict redistribution rule # 5 (PCR5), which proportionally allocates conflicting BPAs to non-empty sets involved in the conflicts based on the BPAs of non-empty sets assigned by sources. Yager (1987) removed the normalization process of Dempster's rule and allocated the BPAs associated with conflicts to the empty set. Yager's (1987) rule provided an intuitive explanation for the combination of evidence but intensified the uncertainty of results, especially for a large number of sources of evidence. To reduce the uncertainty, Inagaki (1991) defined a parameter to reassign the BPAs of the empty set obtained by Yager's rule to existing hypotheses. This parameter has a considerable influence on the combination results, but it is difficult to determine reasonable values. In addition, the averaging method based on arithmetic operators proposed by Murphy (2000) can be used for evidence combination. Although this method does not produce counter-intuitive results, especially in conflict situations, it does not conform to Bayesian conditioning and cannot reinforce the concordant hypothesis (Sentz & Ferson, 2002). The discounting method (Shafer, 1976) provides the interesting idea that conflicts are generated from unreliable sources and hence the BPAs of each hypothesis should be discounted to the universal set according to the reliability of the sources. The lower the reliability is, the higher the discounted degree is. However, discounting unreliable evidence cannot successfully manage conflicts because conflicts between different pieces of evidence do not always mean unreliable evidence. Especially when the information is generated from persons, for instance in group decision-making problems, disagreements due to personal preferences are allowed and have nothing to do with reliability (Khorshidi & Aickelin, 2021).

Overall, the PCR5 rule performs well in combining evidence in high-conflict situations. However, neither this method nor any of the above-mentioned evidence combination methods captures the following issues simultaneously: (a) the attitudinal character of decision makers, (b) the reliability of sources (c) and the importance of sources. Attitudinal characters in data fusion are reflected in decision makers' tolerance of hypotheses that are not supported by at least one piece of evidence or to a small degree. Different decision makers have different degrees of tolerance (Aggarwal & Tehrani, 2019). Some decision makers may not accept hypotheses that are not supported by at least one piece of evidence (e.g. a group of state leaders choosing national development strategies, for which the opinions of all leaders should be followed), while some decision makers are tolerant of these hypotheses (e.g. a customer making a purchase decision based on online reviews in which the reviewers' opinions can be averaged to obtain a collective one). This tolerance attitude can be modelled with the compensation mechanism between data from different sources (Yager, 1980). More specifically, when using D-S theory for data fusion, we need to determine how much a "larger" BPA of a hypothesis from one source compensates for a "smaller" BPA of this hypothesis from another source. To solve this problem, we introduce an adjustment coefficient to the BPA of evidence. The ratio between the BPAs of different hypotheses for each piece of evidence is adjusted without changing the monotonicity. In this way, the compensation degrees between conflicting pieces of evidence can be achieved in a flexible way. The focus on the personalized compensation requirements of decision makers ensures good interpretability of the management of conflicts. The counter-intuitive problem in Dempster's rule that fails to satisfy the decision structures of decision makers with low tolerance can be resolved.

In this study, reliability refers to the quality of information, and importance refers to the decision maker's subjective preference for information sources in the fusion process (Smarandache et al., 2010). We can only find a handful of studies (Fu et al., 2015; Yang & Xu, 2013) that have distinguished the concept of importance from that of reliability in the data fusion process. In most of these studies, the importance of sources was addressed through the discounting method, which discounts the BPAs of each hypothesis to the universal set according to the importance of sources. Such a process is the same as the treatment of reliability and distorts the actual meaning of the importance of sources. Based on the discounting method and a geometric average operator, this study incorporates the reliability and importance of sources into the model of evidence combination in different ways.

The contributions of this study are summarized below:

1. An adaptive evidence combination (AEC) method is proposed for data fusion, which considers the attitudinal characters of decision makers and avoids the counter-intuitive problem of Dempster's rule.
2. The reliability and importance of sources of evidence are incorporated into the fusion process.
3. The algebraic properties of the proposed evidence combination method are discussed, including the idempotence, commutativity, associativity and continuity.
4. Two case studies, including vehicle recognition based on multiple sensors and purchasing decisions based on online reviews, are conducted to illustrated the proposed method.

The rest of the study is organized as follows. Section 2 introduces the conflict problems in D-S theory. Section 3 proposes the AEC method. Section 4 and Section 5 explain the proposed method through case studies on vehicle recognition and online purchasing decisions, respectively. Section 6 ends the paper with conclusions.

2. The conflict problem in D-S theory

This section introduces D-S theory and analyses the reasons for the counter-intuitive results provided by Dempster's rule. D-S theory consists of two main parts, specifically the establishment of the frame of discernment and the combination of evidence, which correspond to two challenges: 1) expressing evidence and 2) tackling conflicting evidence (Sentz & Ferson, 2002). In D-S theory, "probabilities" (BPAs) are allocated to the power set of a finite discrete space rather than to mutually exclusive singletons as in pure probability theory. In this way, one piece of evidence is not limited to one possible hypothesis but can be related to multiple possible hypotheses. This is consistent with uncertain environments, in which information appears to be subjective, hesitant, ignorant and inaccurate.

Consider a finite set of N possible hypotheses, which are mutually exclusive and exhaustive, denoted as $\Theta = \{H_1, H_2, ..., H_N\}$ and named as the frame of discernment. There are 2^N power sets, that is, $P(\Theta) = \{\emptyset, H_1, ..., H_N, \{H_1, H_2\}, ..., \{H_1, H_N\}, ..., \{H_1, ..., H_{N-1}\}, \Theta\}$. The probability distribution of the power sets can be defined using a BPA function (also named a mass function) $m : P(\Theta) \rightarrow [0, 1]$, which satisfies: $m(\emptyset) = 0$, $\sum_{H \subseteq \Theta} m(H) = 1$, where $m(H)$, named the BPA, indicates the strength of the evidence supporting the arbitrary hypothesis H in $P(\Theta)$, and it is independent of any subset of H. A piece of evidence can be represented by the possible power sets with their BPAs, that is,

$$M = \left\{ \{H, m(H)\} | \forall H \subseteq \Theta, m(H) \geq 0, \sum_{H \subseteq \Theta} m(H) = 1 \right\} \quad (1)$$

where H is a focal element of M if $m(H) > 0$.

Following Dempster's rule, two pieces of evidence are combined by:

$$m(H) = [m_1 \oplus m_2](H) = \begin{cases} 0, & H = \emptyset \\ \dfrac{\sum_{B \cap C = H} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C)}, & H \neq \emptyset \end{cases} \quad (2)$$

where $H, B, C \subseteq \Theta$, m_1 and m_2 are the BPA functions of two pieces of independent evidence M_1 and M_2, respectively, and "$\oplus$" denotes an orthogonal sum operator.

According to Dempster's rule, the BPAs of a hypothesis for two pieces of evidence are directly multiplied. In this way, if one piece of evidence negates a hypothesis, the result is negation of this hypothesis, no matter how much the other evidence supports it. Zadeh's (1986) example is a good illustration of this problem. Consider the possible diseases of a patient as $\Theta = \{H_1 = \text{meningitis}, H_2 = \text{concussion}, H_3 = \text{tumour}\}$. Two doctors gave this patient a diagnosis of $M_1 = \{\{H_1, 0.9\}, \{H_2, 0\}, \{H_3, 0.1\}\}$ and $M_2 = \{\{H_1, 0\}, \{H_2, 0.9\}, \{H_3, 0.1\}\}$, respectively. Combining the two pieces of evidence following Dempster's rule, we obtain $m(H_3) = 1$, although both doctors held that the patient had a low probability of a tumour. Such a fusion result is not acceptable to the decision makers in this medical diagnosis problem (Dezert et al., 2004).

Intuitively, in the fusion process, if both pieces of evidence strongly (or weakly) support a hypothesis, then their combined form should also strongly (or weakly) support that hypothesis. In addition, there are two high-conflict situations. 1) If two pieces of evidence are completely conflicting, that is, one hypothesis is supported by one piece of evidence but opposed by another piece of evidence, then should this hypothesis be supported by the combination result? 2) If two pieces of evidence are incompletely conflicting, that is, one piece of evidence supports a hypothesis strongly and the other piece of evidence supports it weakly, does the combination result support this hypothesis strongly or weakly? The answers vary for decision makers because of their personalized preference structures. The AEC method presented in the next section addresses strongly conflicting evidence by taking into account the attitudinal characters of the decision makers.

3. An adaptive evidence combination method based on D-S theory

This section provides a different angle for the management of conflicting evidence in the fusion process by introducing an adjustment coefficient to achieve flexible compensation between BPAs. Two parameters are incorporated into the fusion process, describing the reliability of information and the importance of information. The algebraic properties of the proposed evidence combination method are further discussed.

3.1. The non-iterative algorithm of the AEC method

The requirements for compensation between information in the fusion process can be divided into three categories: 1) the "smaller" BPAs are not allowed to be compensated equally by the "larger" BPAs; 2) the "smaller" BPAs are compensated by the "larger" BPAs to the same degree; and 3) the "smaller" BPAs are compensated by the "larger" BPAs to a certain degree. The arithmetic aggregation operator is in fully compensated form and only meets the second compensation requirement. The geometric aggregation operator has the property of incomplete compensation, which is in line with the third compensation requirement. The challenge of combining evidence is to determine how to adjust the compensation degrees flexibly to meet different requirements. To handle this challenge, we introduce an adjustment coefficient into the BPA of each hypothesis. Different coefficients correspond to different degrees of compensation.

Let a non-negative number θ be an adjustment coefficient. Adding it to the BPAs of all the elements in $P(\Theta)$, we can obtain the adjusted form of evidence M as follows:

$$M^{at} = \left\{ \{H, m^{at}(H)\} \middle| \forall H \subseteq \Theta, m^{at}(H) = m(H) + \theta, m(H) \geq 0, \sum_{H \subseteq \Theta} m(H) = 1 \right\} \tag{3}$$

$m^{at}(H)$ is an extension value of the BPA of H, which is only used for the calculation and has no actual meaning. θ is used to adjust the ratio between the BPAs of two focal elements without changing their monotonicity. Given $B, C \subseteq \Theta$ and $m(B) < m(C)$, we have

$$\frac{m(B)}{m(C)} < \frac{m(B) + \theta_1}{m(C) + \theta_1} < \frac{m(B) + \theta_2}{m(C) + \theta_2} < 1, \ \forall \theta_2 > \theta_1 > 0 \tag{4}$$

The larger θ is, the larger $m^{at}(B)/m^{at}(C)$ is.

Furthermore, suppose that $B, C, B', C' \subseteq \Theta$, where $m(B) + m(C) = m(B') + m(C')$, $m(B) < m(B')$ and $m(C) > m(C')$. There is

$$m^{at}(B)m^{at}(C) - \theta^2 = m(B)m(C) + \theta(m(B) + m(C)) \tag{5}$$

Let $\theta_1(m(B) + m(C)) = \theta_1(m(B') + m(C')) = k_1$ and $\theta_2(m(B) + m(C)) = \theta_2(m(B') + m(C')) = k_2$. There is $k_2 > k_1$ if $\theta_2 > \theta_1$. Then we have

$$\frac{m^{at}(B)m^{at}(C) - {\theta_1}^2}{m^{at}(B')m^{at}(C') - {\theta_1}^2} < \frac{m^{at}(B)m^{at}(C) - {\theta_2}^2}{m^{at}(B')m^{at}(C') - {\theta_2}^2} < 1 \tag{6}$$

since $\frac{m(B)m(C) + k_1}{m(B')m(C') + k_1} < \frac{m(B)m(C) + k_2}{m(B')m(C') + k_2} < 1$.

That is to say, the larger the adjustment coefficient θ added to the BPAs is, the greater the degree of compensation of large BPAs to small BPAs is when the geometric operator is used for fusion. If a decision maker cannot tolerate a hypothesis that is supported for some pieces of evidence but opposed for others, then we set $\theta = 0$. As the tolerance degree of the decision maker for such a hypothesis increases, the value of θ should increase.

Suppose that the BPA functions of two pieces of evidence, M_1 and M_2, are m_1 and m_2 on $P(\Theta)$, respectively. We define the AEC method to combine M_1 and M_2 as follows:

$$m(H) = \left[m_1^{at} \oplus m_2^{at}\right](H) = \begin{cases} 0, \text{ if } m_1(H) = 0 \text{ and } m_2(H) = 0 \\ \dfrac{\sum_{B \cap C = H} \left(m_1^{at}(B) m_2^{at}(C) - \theta^2\right)}{\sum_{H \subseteq \Theta} \sum_{B \cap C = H} \left(m_1^{at}(B) m_2^{at}(C) - \theta^2\right)}, \text{ if } m_1(H) > 0 \text{ or } m_2(H) > 0 \end{cases} \tag{7}$$

where B and C are focal elements of M_1 or M_2, and $m_1^{at}(B)$ and $m_2^{at}(C)$ are the extension values of the BPAs $m_1(B)$ and $m_2(C)$, respectively. We set $m'(H) = \sum_{B \cap C = H} \left(m_1^{at}(B) m_2^{at}(C) - \theta^2\right)$ as the support degree of hypothesis H. The denominator of Eq. (7) is a normalization factor. There is $\sum_{H \subseteq \Theta} m(H) = 1$.

Formally, the feature of the AEC method to combine the BPA of a non-empty element $H \subseteq \Theta$ is that:

1. If $\theta = 0$, there is $m(H) > 0$ when at least one pair of $B, C \subseteq \Theta$ satisfies $m(B) > 0$ and $m(C) > 0$, where $B \cap C = H$; otherwise, $m(H) = 0$.
2. If $\theta > 0$, there is $m(H) > 0$ when at least one $B \subseteq \Theta$ satisfies $m(B) > 0$, where $H \subseteq B$; otherwise, $m(H) = 0$.

Consequently, 1) if $\theta = 0$, then the AEC method degenerates to Dempster's rule and 2) if $\theta > 0$, as long as one piece of evidence supports the hypothesis, it will appear in the combined evidence. This shows that the AEC method is more flexible than Dempster's rule in meeting personalized compensation requirements.

Table 1. The extension values of the BPAs of evidence and their degrees of support determined by the AEC method.

				Doctor 1			
				Disease	Meningitis	Concussion	Tumour
				m_1	0.9	0	0.1
				m_1^{at}	1.0	0.1	0.2
Doctor 2	Disease	m_2	m_2^{at}		$m_1^{at}m_2^{at}-\theta^2$		
	Meningitis	0	0.1		0.099	–	–
	Concussion	0.9	1.0		–	0.099	–
	Tumour	0.1	0.2		–	–	0.03

Note. The hypotheses that are not mentioned by both doctors are eliminated here, and their BPAs are zero.

Furthermore, in the combination process, the BPA $m(\Theta)$ of the ignorance part is assigned to the hypothesis that appears in the combined evidence rather than all the hypotheses in $P(\Theta)$. Consequently, the AEC method conforms to the intuitive assumption that, if a hypothesis does not exist in the original evidence, then it should not appear in the combined evidence.

Example 1 *(Zadeh's example).* Let $\theta = 0.1$. Using Eq. (3), we obtain the extension values of the BPAs, as shown in Table 1.

The support degree for meningitis is $m'(H_1) = 1.0 \times 0.1 - 0.1^2 = 0.099$. Similarly, $m'(H_2) = 0.099$ and $m'(H_3) = 0.03$. After normalization, we can obtain the combined evidence as $\{\{H_1, 0.434\}, \{H_2, 0.434\}, \{H_3, 0.132\}\}$. This suggests that the patient is more likely to have meningitis or concussion than a tumour, which turns out to be intuitive.

Based on Zadeh's example, we further use different values of the adjustment coefficient in the combination process and describe the influence of compensation on the combination results. If $\theta = 0$, the combination of M_1 and M_2 is $M^{(1)} = \{H_1(0), H_2(0), H_3(1)\}$, which is the same as the result obtained by Dempster's rule. We further consider two representative adjustment coefficients, 0.01 and 1. The combination results are $M^{(2)} = \{H_1(0.3), H_2(0.3), H_3(0.4)\}$ and $M^{(3)} = \{H_1(0.448), H_2(0.448), H_3(0.104)\}$, respectively. This shows that, with the increase in θ, the combined BPAs of H_1 and H_2 gradually increase because the degree of compensation of the BPA 0.9 to the BPA 0 increases. In addition, when the adjustment coefficient is large enough (e.g. larger than 100), the BPA 0 can be compensated by the BPA 0.9 to the same extent, and the results ($M^{(4)} = \{H_1\ (0.45), H_2(0.45), H_3(0.1)\}$) are close to those obtained through the averaging method.

Due to the limitations of sources in providing completely accurate information, the reliability of various sources is often taken into account in the process of evidence combination (Fu et al., 2015). As defined by Smarandache et al. (2010), reliability is an objective attribute of an evidence source, indicating its ability to provide a correct measurement or assessment for the problem under consideration. The reliability of one piece of evidence is determined by its source, independent of the reliability of other pieces of evidence involved in the combination process. The reliability can be obtained according to the available properties, such as the measurement accuracy of a given sensor (Smarandache et al., 2010). In addition, objective estimation techniques have been used in the literature (Achroufene et al., 2019; Fu et al., 2015). In general, the greater the gap between the evidence from one source and the evidence from other sources is, the less reliable that source is. The discounting method is effective for combining evidence from sources that are not entirely reliable. Based on the discounting method, we incorporate the reliability factor into the AEC method.

Let $\sigma \in [0, 1]$ be a reliability factor of an evidence source. $\sigma = 0$ means a totally unreliable source, and $\sigma = 1$ means a totally reliable source. The discounted form of evidence M can be defined as:

$$M_\sigma^{at} = \{\{H, m_\sigma^{at}(H)\} | \forall H \subseteq \Theta, m_\sigma^{at}(H) = \left\{ \begin{array}{c} \sigma m(H) + \theta, \text{ if } H \subset \Theta \\ \sigma m(H) + (1-\sigma) + \theta, \text{ if } H = \Theta \end{array} \right., \sum_{H \subseteq \Theta} m(H) = 1 \right\} \quad (8)$$

The main idea of Eq. (8) is that, if a source of evidence is not completely reliable, then the BPAs of the focus elements of the evidence are reallocated to the universal set Θ proportionally. In this way, the impact of this evidence on the combination results can be reduced. We call $m_\sigma^{at}(H)$ the discounted extension value of the BPA. If $\sigma = 1$, then, Eq. (8) degenerates to Eq. (3).

Consider two pieces of evidence, $M_1 = \{H, m_1(H)\} |\ \forall H \subseteq \Theta\}$ and $M_2 = \{H, m_2(H)\} | \forall H \subseteq \Theta\}$, in the same frame of discernment. Let their degrees of reliability be σ_1 and σ_2 and their discounted forms be $M_{\sigma 1}^{at} = \{H,\ m_{\sigma 1}^{at}(H)\} | \forall H \subseteq \Theta\}$ and $M_{\sigma 2}^{at} = \{H,\ m_{\sigma 2}^{at}(H)\} | \forall H \subseteq\ \Theta\}$, respectively. Similar to Eq. (7), the AEC method with the reliability factor can be defined as:

$$m_\sigma(H) = [m_{\sigma 1}^{at} \oplus m_{\sigma 2}^{at}](H) = \left\{ \begin{array}{l} 0, \text{ if } m_1(H) = 0 \text{ and } m_2(H) = 0 \\ \dfrac{\sum_{B \cap C = H} \left(m_{\sigma 1}^{at}(B) m_{\sigma 2}^{at}(C) - \theta^2\right)}{\sum_{H \subseteq \Theta} \sum_{B \cap C = H} \left(m_{\sigma 1}^{at}(B) m_{\sigma 2}^{at}(C) - \theta^2\right)}, \text{ if } m_1(H) > 0 \text{ or } m_2(H) > 0 \end{array} \right. \quad (9)$$

If $\sigma_1 = 0$ (or $\sigma_2 = 0$) and $\theta = 0$, then $m_\sigma(H) = m_{\sigma 2}^{at}(H)$(or $m_{\sigma 1}^{at}(H)$). If $\sigma_1 = \sigma_2 = 1$, then Eq. (9) degenerates to Eq. (7).

In addition, the importance of the sources of evidence may be different. In data fusion, the importance (weight) is defined as a subjective attribute of an evidence source, indicating the decision maker's subjective preference for that source (Smarandache et al., 2010). The definition of importance belongs to a relative concept, that is, the importance of one source is generated by comparing it with other sources. For example, when choosing a national development strategy, the opinions of the top leader are more important than those of ordinary leaders, regardless of how different his/her opinions are from those offered by other leaders. In particular, in the process of multi-criteria decision making, decision makers often place different values on different criteria (Cinelli et al., 2020); for example, one customer may value the food more than the service when choosing a restaurant. In most cases, the weights of sources are provided by decision makers directly (Smarandache et al., 2010) or determined using weighting methods, such as the analytic hierarchy process (Saaty, 1977) and the best–worst method (Rezaei, 2016), based on decision makers' preference information. Arithmetic and geometric average operators are fundamental approaches to information averaging (Li et al., 2019) and have been widely used to aggregate criteria with different degrees of importance in multi-criteria decision making (Aggarwal & Tehrani, 2019). Considering that the AEC method is based on geometric operators, we incorporate the weight factor into the AEC method through geometric average operators.

Let w be the weight of a source of evidence M. The weighted form of the evidence M can be defined as:

$$M_w^{at} = \left\{ \{H, m_w^{at}(H)\} | \forall H \subseteq \Theta, m_w^{at}(H) = (m^{at}(H))^w, \sum_{H \subseteq \Theta} m(H) = 1 \right\} \tag{10}$$

where $m^{at}(H) = m(H) + \theta$[1].

Consider two pieces of evidence, $M_1 = \{H, m_1(H)\}| \ \forall H \in \Theta\}$ and $M_2 = \{H, m_2(H)\}| \ \forall H \in \Theta\}$, in the same frame of discernment. Let their weights be w_1 and w_2, with $w_1 + w_2 = 1$, and their weighted forms be $M_{w1}^{at} = \{H, m_{w1}^{at}(H)\}|\forall H \subseteq \Theta\}$ and $M_{w2}^{at} = \{H, m_{w2}^{at}(H)\}|\forall H \subseteq \Theta\}$, respectively. The AEC method with the weight factor can be defined as:

$$m_w(H) = [m_{w1}^{at} \oplus m_{w2}^{at}](H) = \begin{cases} 0, \text{ if } m_1(H) = 0 \text{ and } m_2(H) = 0 \\ \dfrac{\sum_{B \cap C = H} (m_1^{at}(B)^{w_1} m_2^{at}(C)^{w_2} - \theta)}{\sum_{H \subseteq \Theta} \sum_{B \cap C = H} (m_1^{at}(B)^{w_1} m_2^{at}(C)^{w_2} - \theta)}, \text{ if } m_1(H) > 0 \text{ or } m_2(H) > 0 \end{cases} \tag{11}$$

Furthermore, combining Eqs. (9) and (11), we define the AEC method that incorporates the parameters of both information reliability and importance as follows:

$$m_{\sigma w}(H) = [m_{\sigma w1}^{at} \oplus m_{\sigma w2}^{at}](H) = \begin{cases} 0, \text{ if } m_1(H) = 0 \text{ and } m_2(H) = 0 \\ \dfrac{\sum_{B \cap C = H} ((m_{\sigma 1}^{at}(B))^{w_1} (m_{\sigma 2}^{at}(C))^{w_2} - \theta)}{\sum_{H \subseteq \Theta} \sum_{B \cap C = H} ((m_{\sigma 1}^{at}(B))^{w_1} (m_{\sigma 2}^{at}(C))^{w_2} - \theta)}, \text{ if } m_1(H) > 0 \text{ or } m_2(H) > 0 \end{cases} \tag{12}$$

where $m_{\sigma 1}^{at}(B)$ and $m_{\sigma 2}^{at}(C)$ are the discounted extension values of $m(B)$ and $m(C)$, respectively.

Equation (12) is a non-iterative algorithm of the AEC method. If the source of evidence is completely reliable, then Eq. (12) degenerates to Eq. (11); if there are no relative weights between sources of evidence, then Eq. (12) degenerates to Eq. (9); and, if all the sources of evidence are completely reliable without relative importance, then Eq. (12) degenerates to Eq. (7).

3.2. The iterative algorithm of the AEC method

In practical terms, it is usually necessary to combine evidence from a large number of sources. Using Eq. (12) to deal with this problem is computationally complex. Thus, we further develop an iterative algorithm below.

Consider n pieces of independent evidence $M = \{M_1, M_2, ..., M_n\}^T$. Let their weight vector be $W = \{w_1, w_2, ..., w_n\}^T$ such that $\sum_{j=1}^n w_j = 1$, where $M_j = \{H_j, m(H_j)\}|\forall H \subseteq \Theta\} (j = 1, 2, ..., n)$, and their reliability vector be $\Psi = \{\sigma_1, \sigma_2, ..., \sigma_n\}^T$ with $\sigma_j \in [0, 1]$. The iterative form of the AEC method to combine these n pieces of evidence can be defined as follows:

$$m_{\sigma w(n)}(H) = [m_{\sigma w1}^{at} \oplus \cdots \oplus m_{\sigma wn}^{at}](H) = \begin{cases} 0, \text{ if } \forall j \in \{1, 2, ..., n\}, m(H_j) = 0 \\ \dfrac{\sum_{C_1 \cap \cdots \cap C_n = H} \left(\prod_j^n (m_\sigma^{at}(C_j))^{w_j} - \theta \right)}{\sum_{H \subseteq \Theta} \sum_{C_1 \cap \cdots \cap C_n = H} \left(\prod_j^n (m_\sigma^{at}(C_j))^{w_j} - \theta \right)}, \text{ if } \exists j \in \{1, 2, ..., n\}, m(H_j) > 0 \end{cases} \tag{13}$$

where C_j is a hypothesis in M_j, and its discounted extension value of BPA is

$$m_\sigma^{at}(C_j) = \begin{cases} \sigma_j m(C_j) + \theta, \text{ if } C_j \subset \Theta \\ \sigma_j m(C_j) + (1 - \sigma_j) + \theta, \text{ if } C_j = \Theta \end{cases} \tag{14}$$

Note that C_j, $\forall j \in \{1, 2, ..., n\}$, is a focal element of at least one piece of evidence in M. $\sum_{C_1 \cap \cdots \cap C_n = H} (\prod_j^n (m_\sigma^{at}(C_j))^{w_j} - \theta)$ denotes the degree of support for hypothesis H. If all the pieces of evidence are of the same importance, then $w_j = 1/n, j = 1, 2, ..., n$. This iterative algorithm is suitable for situations in which there is no or little interaction between focal elements; otherwise, the amount of calculation will increase exponentially.

Motivated by the properties of the classical evidence combination rules discussed by Sentz and Ferson (2002), we analyse four algebraic properties of the AEC method, namely idempotence, commutativity, associativity and continuity. Please refer to the Appendix for details.

Table 2. A vehicle classification for toll roads.

Type	Classification standard				Representative vehicle
	Axles (number)	Wheels (number)	Head height (m)	Wheelbase (m)	
H1	2	$2\sim4$	<1.3	<3.2	Cars, jeeps, taxis, vans, motorcycles
H2	2	4	≥1.3	≥3.2	Minivan, minivan, light truck, minibus
H3	2	6	≥1.3	≥3.2	Medium bus, large bus and medium truck
H4	3	$6\sim10$	≥1.3	≥3.2	Large bus, large bus, large trailer, 20-foot container car
H5	>3	>10	≥1.3	≥3.2	Heavy truck, heavy trailer, 40-foot container car

4. Case study 1: Target recognition for vehicle types

This section aims to illustrate the use of the AEC method for the fusion of small-size data through an example of vehicle recognition. The results are compared with those obtained using other evidence combination methods. A sensitivity analysis of the adjustment coefficient is conducted.

4.1. Description of vehicle recognition

Target recognition involves judging the size, shape, weight or other physical characteristic parameters of targets through sensors and then making decisions in classifiers (Dong & Kuang, 2015). Due to individual characteristics, the environment and other factors, different sensors in a multi-sensor target recognition system often have differing ability to recognize targets (Frikha & Moalla, 2015). Therefore, fusing the judgement information from multiple sensors is beneficial to improve the recognition ability (Xiao, 2020). In a recognition system, the types of targets are mutually exclusive and non-intersecting, forming a frame of discernment. The information detected by each sensor can constitute a piece of evidence. Therefore, D-S theory has generated successful applications in multi-sensor target recognition problems (Frikha & Moalla, 2015).

Intelligent vehicle recognition systems are essential for toll systems of roads, such as entrances, bayonets, bridges and crossings (Li & Lv, 2017). For the healthy development of highways, China reformulated the classification standards of toll road vehicles in 2019.[2] Vehicle profile parameters, such as the axle number, wheel number, head height, wheelbase and total mass, are used to classify vehicle types. Table 2 shows the classification standard of toll vehicles used in Guangdong Province, China,[3] consisting of four attributes: the number of axles (c_1), the number of wheels (c_2), the head height (c_3) and the wheelbase (c_4). There are five types of toll vehicles, denoted as H1, H2, H3, H4 and H5, respectively. Axles are integers, and the wheels of H2, H3, H4 and H5 are even in number.

Common vehicle recognition devices include magnetic field sensors, inductive-loop sensors and surveillance cameras (Li & Lv, 2017). Magnetic field sensors have the advantage of convenient installation and maintenance and are not affected by climate. However, they have blind spots for detection and low accuracy. Inductive-loop sensors are based on electromagnetic induction to obtain the vehicle speed, vehicle length, axle number, wheelbase and others. They have excellent performance under bad weather conditions, but their accuracy is easily affected by other factors, such as vehicle speed and vehicle parallel passage. The image-based vehicle recognition method identifies the vehicle type from the image acquired by surveillance cameras. Surveillance cameras are widely used, but their accuracy is easy influenced by the environment and the vehicle attributes. It is apparent that different vehicle recognition devices have their own advantages and disadvantages. To improve the accuracy of recognition, existing vehicle recognition systems usually use a variety of sensors (Li et al., 2021). In the following, we provide a numerical example showing how to recognize vehicle types based on the information provided by multiple sensors.

Suppose that a toll road is equipped with six sensors, consisting of two magnetic field sensors (M1, M2), two inductive-loop sensors (M3, M4) and two surveillance camera-based sensors (M5, M6). The vehicle classification standard listed in Table 2 is used for charging. The profile parameters of a vehicle detected by the six sensors are shown in Table 3. The spaces are the information that the sensors do not provide.

4.2. Methodology

Step 1. Determine the BPA of each piece of evidence

The five vehicle types form a frame of discernment $\Theta=\{H_1,H_2,\ldots,H_5\}$. Each sensor is a source of evidence. An intuitive method for determining the BPA of each vehicle type is based on the degree of matching between the attribute values detected and those of the vehicle type. Let $[x_i^{k-},x_i^{k+}]$ be the value range of type H_i with regard to attribute c_k, x_j^k be the value of attribute c_k measured by sensor S_j and $[x^{k-},x^{k+}]$ be the general value range of attribute c_k. With respect to attribute c_k, the possibility of a measured value x_j^k closing to the value of vehicle type H_i can be determined by the following:

Table 3. The profile parameters of a vehicle detected by six sensors.

	Measurement parameter			
Sensors	Axles (number)	Wheels (number)	Head height (m)	Wheelbase (m)
M1	2	6	2.8	5.1
M2	2	6	2.3	5.4
M3	2	6	2.5	5.3
M4	4	10	–	–
M5	3	6	2.6	5.6
M6	3	8	2.7	5.2

Table 4. The BPAs of five pieces of evidence for vehicle recognition.

	Vehicle type					
Sensors	H1	H2	H3	H4	H5	Ignorance
M1	0.16	0.23	0.26	0.23	0.12	0.00
M2	0.16	0.23	0.26	0.23	0.11	0.00
M3	0.16	0.23	0.26	0.23	0.11	0.00
M4	0.00	0.00	0.03	0.22	0.25	0.50
M5	0.12	0.21	0.24	0.28	0.16	0.00
M6	0.11	0.18	0.22	0.31	0.18	0.00

1. If x_j^k is not empty, then

$$m(H_{ij}^k) = \frac{\max\left(\frac{x^{k+}-x^{k-}}{2} - \min_{x_i^k\in[x_i^{k-},x_i^{k+}]}|x_i^k-x_j^k|,0\right)}{\sum_i \max\left(\frac{x^{k+}-x^{k-}}{2} - \min_{x_i^k\in[x_i^{k-},x_i^{k+}]}|x_i^k-x_j^k|,0\right)} \quad (15)$$

2. If x_j^k is empty, then

$$m(H_{ij}^k) = 0 \quad (16)$$

The collective possibility of a vehicle type H_i can be determined by:

$$m(H_{ij}) = \frac{1}{K}\sum_k m(H_{ij}^k) \quad (17)$$

where K is the number of attributes. In this case, $K = 4$. In evidence M_j, $m(H_{ij})$ is the BPA of H_i. The BPA of the ignorance part is:

$$m(\Theta_i) = 1-\sum_i m(H_{ij}) \quad (18)$$

Suppose that the general value ranges of the four attributes of vehicles are [2,6], [2-12], [0.5, 6] and [2, 7]. Using Eqs. (15)–(18), we obtain the BPAs of five pieces of evidence. The values in Table 4 express the extent to which each piece of evidence supports the corresponding vehicle types.

Step 2. Combine the evidence to determine the collective BPA

According to the historical recognition records, the classification accuracy of five sensors is obtained as 87%, 88%, 90%, 94%, 92% and 96%, respectively, which can be taken as the reliability of the sensors. Using Eqs. (8) and (9), we obtain the combination results, as shown in Figure 1(a). We assume that the weights of the six sensors are 0.2, 0.2, 0.15, 0.05, 0.2 and 0.2, respectively. Considering the weights of the evidence, we further apply the iterative algorithm of the AEC to make a combination. The results are shown in Figure 1(b).

In both situations, we consider 20 adjustment coefficients, including $\theta \in \{0, 0.05, 0.1, \cdots, 0.95, 1\}$. In total, the combined BPAs of the vehicle types satisfy $m(H_4)>m(H_3) >m(H_2)>m(H_5)>m(H_1)$. That is to say, the vehicle is most likely to be the fourth type (H4). The results obtained by the non-iterative algorithm provide us with a meaningful suggestion; that is, if we can accept the existence of evidence against H3, then the combined evidence supports H3 to a large extent. In addition, we find that the adjustment coefficient does not change the preference relationship of different vehicle types. As the adjustment coefficient increases, the overall differences between the BPAs of different vehicle types shrink. This is due to an increase in the compensation degree of large BPAs on small BPAs. The BPAs of the five vehicle types obtained through the iterative algorithm are closer than those obtained through the non-iterative algorithm due to the consideration of weights. H3 is more supportive when considering the importance of evidence than when not considering its importance. This is because the fourth sensor (M4) is largely opposed to H3, but this sensor has a small weight.

4.3. Comparison analysis

For comparison, we further apply seven representative evidence combination methods to solve Case 1, specifically Dempster's rule, Yager's rule, Inagaki's (1991) rule, the PCR5 rule, the discounting method, Zhang's (1994) rule and the averaging method. In this part, we do not consider the importance of evidence. The results obtained through these methods are shown in Table 5 and intuitively summarized in Figure 2. Note that only the discounting method and the AEC method can address the reliability of evidence. When applying Inagaki's rule, we consider five parameters (0, 0.25, 0.5, 0.75, 1), that is, the proportion of the BPAs of conflicts allocated to the universal set and the empty set. All the results show that the detected vehicle belongs to the fourth type, which indicates the effectiveness of the AEC method. Dempster's rule, PCR5 and Zhang's rule derive a BPA of the ignorant part of zero, which may not be accepted because there is incomplete

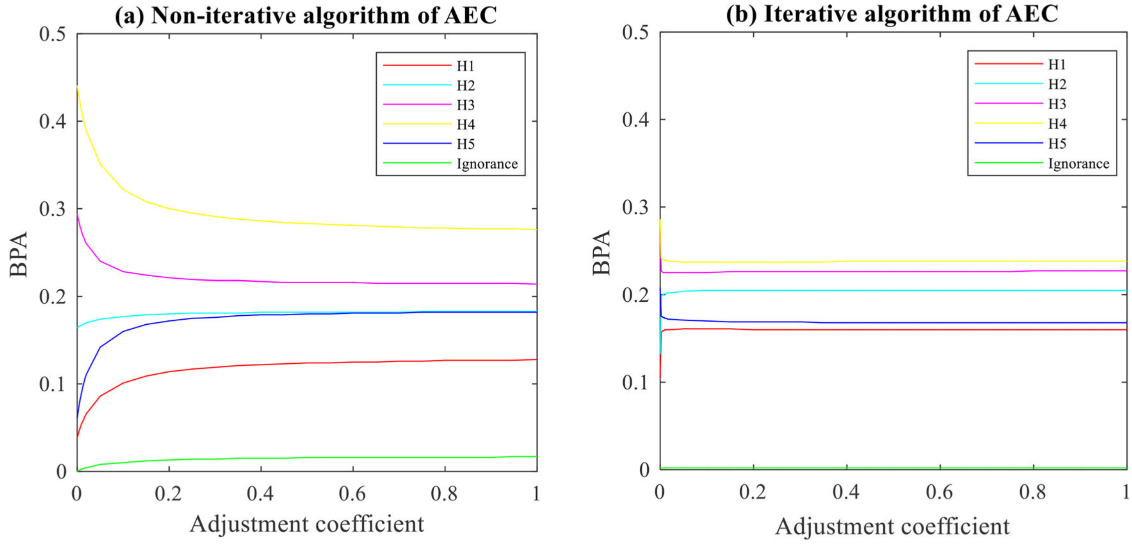

Figure 1. The BPAs of five vehicle types obtained through the proposed method.

Table 5. The BPAs of the combined evidence for vehicle recognition.

BPA		H1	H2	H3	H4	H5	Ignorance
Dempster's rule		0.016	0.150	0.333	0.481	0.021	0.000
Yager's rule		0.053	0.100	0.130	0.192	0.096	0.429
Inagaki's rule	0	0.053	0.100	0.130	0.192	0.096	0.429
	0.25	0.057	0.110	0.145	0.221	0.105	0.362
	0.50	0.060	0.121	0.165	0.261	0.114	0.278
	0.75	0.058	0.135	0.201	0.330	0.116	0.161
	1	0.016	0.150	0.333	0.481	0.021	0.000
PCR5 rule		0.037	0.151	0.236	0.381	0.195	0.000
Discounting method		0.039	0.164	0.294	0.441	0.061	0.001
Zhang's rule		0.010	0.089	0.238	0.635	0.029	0.000
Averaging method		0.133	0.205	0.239	0.255	0.143	0.025
AEC method ($\theta = 0.1$)		0.101	0.177	0.228	0.322	0.160	0.010

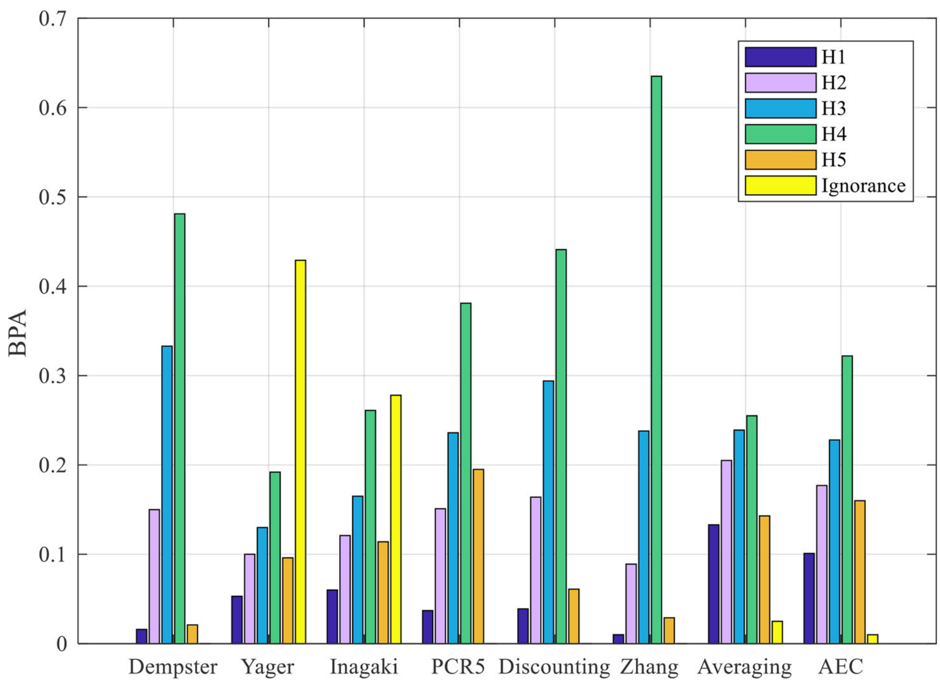

Figure 2. The BPAs of the five vehicle types obtained through different evidence combination methods.

evidence (M4). Conversely, the results obtained using Yager's rule and Inagaki's rule are quite uncertain because they assign the BPAs of conflicts to the empty set, which increases the difficulty of decision making. The results obtained through the discounting method are the same as the results obtained through the AEC method, with the adjustment coefficient being 0. The averaging method ignores the intersection between the focal elements of evidence in this case. This is why the BPA of the ignorant part obtained using the averaging method is greater than the BPA obtained using the discounting and AEC methods. In total, the AEC method is an advanced evidence combination method due to its flexibility in managing conflicts.

5. Case study 2: Purchasing decisions based on online reviews

In this section, we prove the effectiveness of the AEC method in dealing with large-size data through a case of purchasing decisions based on online reviews.

5.1. Description of the data

TripAdvisor.com (https://www.tripadvisor.com) is the world's leading travel website, with more than 7 million hotels, attractions and restaurants in more than 190 countries around the world. This platform provides travellers with online reviews and booking functions. Suppose that a traveller plans to book a hotel in Chengdu, Sichuan Province, China, from TripAdvisor.com. His/her basic requirements for the hotel are that it is located near Chunxi Road and that its price is between 300 and 500 RMB. Filtering the search terms provided by the site, we find five hotels that satisfy the basic requirements of this traveller. The basic information about these five hotels is shown in Table 6.

The traveller wants to select the hotel with the best online reviews. This is a group decision-making problem in that we need to make a selection from five alternative hotels ($i = 1, 2, \ldots, 5$), the performances of which are evaluated by groups of reviewers. As shown in Figure 3, text reviews describe in detail the experience and feelings of reviewers (customers) regarding hotels, involving various hotel attributes about which they are concerned. The (positive or negative) ratings of the same hotel vary from reviewer to reviewer. For example, the overall review of the hotel "Somerset Riverview Chengdu" by reviewer "ChaseHK" was negative, while the overall review by reviewer "andypscyan" was positive. That is to say, we need to manage conflicting reviews when fusing reviewers' opinions on a hotel. In addition, text reviews contain sentiment words with different orientations (positive or negative) and intensities (weak or strong) because the performances of the various attributes of a hotel are inconsistent. For example, "ChaseHK" held that the location of "Somerset Riverview Chengdu" was "good" but that the staff's English skills were "limited". Generally, the sentiment levels of an online review can be divided into five categories (Kang & Park, 2014), namely {*strongly negative (SN)*, *negative (N)*, *neutral (Ne)*, *positive (P)*, *strongly positive (SP)*}, which form a frame of discernment, $\Theta = \{H_1 = SN, H_2 = N, H_3 = Ne, H_4 = P, \ H_5 = SP\}$. Each text review can be deemed as a source of evidence for the hotel's performance. The frequency with which a sentiment level occurs in a text review can be considered as the BPA of that sentiment level. In terms of the sentiment, the mathematical form of a text review can be denoted as $M_{il} = \{\{H_\alpha, m_{il}(H_\alpha)\} | \alpha = 1, \ 2, \ldots, 5\}$ ($l \in \{1, 2, \ldots, \ L_i\}, \ i \in \{1, 2, \ldots, 5\}$), where $m_{il}(H_\alpha)$ is the frequency (BPA) of the α th sentiment level in the l th text review of the i th hotel and L_i is the number of all the text reviews of the i th hotel.

A Web crawler, Houyicaiji (http://www.houyicaiji.com), is applied to crawl the text reviews of the five alternative hotels from TripAdvisor.com. The data are collected in August 2020. Because of the limited space, we do not present the crawling results here. A natural language processing program, the Stanford CoreNLP (https://stanfordnlp.github.io/CoreNLP/) (Manning et al., 2014), is used for the sentiment analysis of text reviews. The Stanford CoreNLP can automatically identify sentiment words and their sentiment levels $\{SN, N, Ne, P, SP\}$.

Table 6. Basic information about the alternative hotels.

	Name	Number of text reviews	Star rating
Hotel 1	Sofitel Chengdu Taihe	531	4.0
Hotel 2	Somerset Riverview Chengdu	151	4.5
Hotel 3	Tianfu Sunshine Hotel	249	4.0
Hotel 4	Haiyatt Hotel	65	4.0
Hotel 5	Xin Liang Hotel	56	4.0

Reviewer	ChaseHK	andypscyan
Star rating	●●●○○	●●●●●
Overall text review	Average Chinese hotel	Good Experience
Detailed text review	I always try to stay at a somerset when I travel in Asia as they have good locations and quality staff. This hotel was a good location but the room we reserved was not the one we got upon check-in. Staff was ok, but limited English skills.	Very good room equipment and facility, locate in riverside and near the city central. But the breakfast quality need to be improving. Also is a nice place to stay with family because I sure children will like it very well.

Figure 3. Two online reviews of the hotel "Somerset Riverview Chengdu".

Reviewer	Text review	Sensitive level and frequency				
		SN	N	Ne	P	SP
1	Very comfortable stay,Stayed for 2 nights and our stay was very comfortable and pleasant The staff are very friendly and they guided us ver	0.03	0.20	0.23	0.42	0.12
2	Beautiful Property. Great Amenities.,Always love a Sofitel and Taihe was no exception. Friendly, helpful, and knowledgeable staff. Wonderfu	0.04	0.09	0.12	0.53	0.22
3	Chengdu Sofitel Taihe Stay,Our family of 8 recently enjoyed a relaxing stay at the Sofitel Chengdu Taihe in 3 Prestige Suites with daily Club	0.22	0.39	0.22	0.11	0.05
4	A very pleasant experience,I almost had a perfect stay. The services we received were world class and the food was great. The only thing is	0.04	0.15	0.13	0.46	0.22
5	Great stay,We book Junior suites for 3 nights. It was a great stay. The hotel is five minutes walk from Jinjiang hotel station from exit c. The f	0.11	0.36	0.28	0.20	0.06
6	Great hotel and excellent staff,We are a group of women who stayed at the hotel during Dec 2019. We enjoy v much especially the good bre	0.02	0.08	0.30	0.47	0.13
7	Great Stay...Unbeatable Value,We couldn't pass this up as you can get a suite for such a great price. Stayed five nights and ended up being	0.03	0.09	0.28	0.41	0.18
8	Airy,We stayed at the Sofitel Chengdu Taihe as a group of 13 ladies. I liked this hotel better than the Sofitel Xian On Renmin Square, this on	0.07	0.32	0.33	0.21	0.06
9	Excellent,Rooms are well appointed and clean. Everything in the room worked. The service was outstanding, especially house-keeping. The	0.04	0.20	0.20	0.37	0.20
10	Nice surprise,Good hotel spacious room and well appointed pleasant staff. Had some good food in the hotel try the Dan Dan noodles. Breakfa	0.04	0.21	0.27	0.35	0.13
11	Only gets better,The Chengdu region is a very interesting area to visit. This is the third time to stay at this hotel in the past three weeks. It is	0.02	0.10	0.27	0.42	0.19
12	Nice Hotel, Nice People, Very good lounge,This is my second stay at this hotel and it was just as good as the first. The building and the rooms	0.04	0.14	0.14	0.45	0.22
13	Worst Sofitel I have ever experienced.,I had read the reviews and with an open mind I booked a river view suite (with lounge access) feeling	0.15	0.56	0.23	0.05	0.02
14	Probably the nicest hotel stay I have had,The hotel isn't only luxurious, well situated and clean but a great bargain too. The hotel is in the cent	0.13	0.39	0.12	0.25	0.11
15	Comfortable stay but we expected more,The hotel is dated and is in need of a refurbishment. The breakfast buffet spread was not up to expe	0.13	0.55	0.24	0.07	0.02
16	Not up to the Sofitel brand,This hotel was a disappointment to the Sofitel brand. Foyer was lovely & rooms were well appointed. Bathroom w	0.10	0.44	0.26	0.15	0.04
17	Good intentions average execution,At a Sofitel, you'd expect a premium service and product. The staff here were eager to help, good manner	0.22	0.52	0.22	0.03	0.01
18	Excellent Location,Chengdu was our first stop in our trip to China and hence, this hotel was the first of our pit stops. The hotel was a welcom	0.14	0.44	0.19	0.13	0.11
19	Glitzy Hotel,Hotel was well kept for the most part. Breakfast was good with a large selection of food. However, the swimming pool area had	0.07	0.33	0.33	0.21	0.06
20	Sofitel Taihe, Chengdu,this was a fantastic hotel, the room was stunning, staff, food, comfort, amazing, no complaints at all, great location with	0.00	0.00	0.01	0.49	0.50

Figure 4. Sentiment analysis results of text reviews.

Table 7. The BPAs of the combined evidence of the alternative hotels for the purchasing decision.

	Method		SN	N	Ne	P	SP	Utility
Hotel 1	Dempster's rule		0.000	0.000	0.00	0.000	1.000	1.000
	Averaging method		0.100	0.309	0.219	0.259	0.113	0.494
	AEC method	0	0.000	0.000	0.000	0.000	1.000	1.000
		0.01	0.096	0.316	0.238	0.250	0.100	0.485
		0.1	0.099	0.312	0.229	0.252	0.107	0.489
		1	0.100	0.309	0.222	0.257	0.112	0.493
Hotel 2	Dempster's rule		0.000	0.000	0.000	1.000	0.000	0.750
	Averaging method		0.088	0.277	0.211	0.289	0.134	0.526
	AEC method	0	0.000	0.000	0.345	0.464	0.191	0.711
		0.01	0.086	0.276	0.219	0.295	0.124	0.523
		0.1	0.089	0.277	0.215	0.290	0.129	0.523
		1	0.089	0.277	0.212	0.289	0.133	0.525
Hotel 3	Dempster's rule		0.000	0.000	0.000	1.000	0.000	0.750
	Averaging method		0.086	0.297	0.221	0.283	0.112	0.510
	AEC method	0	0.000	0.000	0.390	0.448	0.162	0.693
		0.01	0.082	0.299	0.239	0.277	0.103	0.505
		0.1	0.086	0.297	0.231	0.277	0.108	0.506
		1	0.086	0.297	0.224	0.282	0.111	0.509
Hotel 4	Dempster's rule		0.000	0.991	0.000	0.009	0.000	0.254
	Averaging method		0.086	0.298	0.237	0.274	0.106	0.504
	AEC method	0	0.082	0.295	0.252	0.274	0.098	0.503
		0.01	0.083	0.295	0.250	0.273	0.099	0.502
		0.1	0.086	0.296	0.244	0.271	0.103	0.502
		1	0.086	0.298	0.239	0.272	0.105	0.503
Hotel 5	Dempster's rule		0.000	1.000	0.000	0.000	0.000	0.250
	Averaging method		0.101	0.347	0.246	0.227	0.080	0.459
	AEC method	0	0.098	0.363	0.260	0.211	0.068	0.447
		0.01	0.099	0.361	0.258	0.212	0.070	0.449
		0.1	0.101	0.353	0.252	0.217	0.076	0.454
		1	0.101	0.348	0.247	0.224	0.079	0.458

It can also calculate the frequency of each sentiment level in a text review. Figure 4 provides a list of some of the results of the sentiment analysis undertaken using the Stanford CoreNLP.

5.2. Solving process

Considering the large amount of data, we use the iterative algorithm of the AEC method to fuse the text reviews of each alternative hotel. Two steps are conducted to facilitate this group decision-making problem.

Step 1. Combine text reviews into a collective one

In this case, we suppose that the text reviews provided by different reviewers are of the same importance and completely reliable. All the focal elements in the evidence are singleton terms, and each piece of evidence is complete. Thus, the calculation process is simple since we only need to integrate the BPAs of each sentiment level. Four adjustment coefficients are considered, 0, 0.01, 0.1 and 1. Through fusion, we can obtain the collective review of each hotel, denoted as $M_i = \{\{H_\alpha, m_i(H_\alpha)\} | \alpha = 1, 2, ..., 5\}$, where $m_i(H_\alpha)$ is the aggerated frequency (BPA) of the α th sentiment level for the i th hotel, $i \in \{1, 2, ..., 5\}$.

To deal with this case, the discounting method and Zhang's rule are the same as Dempster's rule since all the pieces of evidence are completely reliable and there is no intersection between the focal elements of the evidence. Yager's rule and Inagaki's rule are not applicable in this case because they do not have an iterative algorithm, which makes the calculations heavy. Therefore, we only use

Table 8. The utilities and rankings of the alternative hotels.

Method	Result	Hotel 1	Hotel 2	Hotel 3	Hotel 4	Hotel 5
Dempster's	Utility	1.00	0.750	0.750	0.254	0.250
	Rank	1	2	2	4	5
Averaging	Utility	0.494	0.526	0.510	0.504	0.459
	Rank	4	1	2	3	5
AEC	Utility	[0.485, 0.493]	[0.523,0.525]	[0.505, 0.509]	[0.502, 0.503]	[0.447, 0.458]
	Rank	4	1	2	3	5

Dempster's rule and the averaging method for the comparative analysis in this case study. Table 7 contains the combined BPAs of the sentiment levels for hotels.

Step 2. Calculate the utilities of hotels based on the combined evidence

The more positive the level of sentiment is, the better the hotel is. Without loss of generality, we can set the utility $u(H_\alpha)$, $\alpha = 1, 2, \ldots, 5$, of the five sentiment levels from the most negative to the most positive as 0, 0.25, 0.5, 0.75, and 1, respectively. The utility of each hotel can be determined by averaging the utilities of the five sentiment levels through their frequencies (Wu & Liao, 2021), that is,

$$U(M_i) = \sum_{\alpha=1}^{5} (u(H_\alpha) \times m_i(H_\alpha)) \tag{19}$$

Table 8 shows the utilities and rankings of the five hotels. The hotel rankings obtained through the AEC method and the averaging method are in line with the hotels' star ratings, which indicate that Hotel 2 is the best. Dempster's rule, however, gives a different result from the hotels' star ratings. The main reason for this difference is that Dempster's rule does not support completely conflicting evidence. For example, with only two reviewers not supporting the view that Hotel 1 is "negative", Dempster's rule has no support for this view of Hotel 1, although most other reviewers give strong support for this view. This does not satisfy the requirement of group decision-making problems in which disagreements are allowed. The results obtained through the averaging method are close to the results obtained through the AEC method with a large adjustment coefficient. Without knowing the decision maker's compensation requirements, we can solve the fusion problem by using a set of adjustment coefficients separately in the AEC method. The fusion results obtained in this way, though not precise, can reflect the possible combined BPAs of various hypotheses and ensure the reliability of the decision results under different conditions. As shown in Table 8, we obtain a range of utilities for each hotel, considering four adjustment coefficients. We can confirm that Hotel 2 is the best because its minimum utility value is greater than the maximum utility values of the other four hotels.

6. Conclusions

This study focused on improving the interpretability of D-S theory in managing conflicting evidence in decision making under uncertainty by considering different characteristic parameters, including the attitudinal character of decision makers, the reliability of sources and the importance of sources. Firstly, by introducing an adjustment coefficient to the BPA of evidence, we proposed the AEC method, which can meet different compensation requirements of decision makers for combining the BPAs. With the increase in the value of adjustment coefficient, the compensation degree of the larger BPAs to the smaller BPAs increases. Such a setting can satisfy the personalized tolerance attitude of decision makers towards a hypothesis that is strongly supported by some pieces of evidence but weakly supported (or opposed) by others. Then, both the reliability and the importance of sources of evidence were incorporated into the AEC method. Furthermore, we conducted a case study of vehicle type recognition based on multiple sensors to demonstrate the non-iterative algorithm of the AEC method in fusing small-size data. Moreover, a case involving a purchasing decision based on text reviews proved the effectiveness of the iterative algorithm of the AEC method in fusing large-size data. Finally, through a comparative analysis, we showed that the proposed evidence combination method had the advantage of reliability in dealing with conflicting evidence.

The proposed method has implications for decision making under uncertainty, in which the sources of information, especially human subjective assessments, may be incompletely reliable. This study found that the reason for Dempster's rule possibly leading to counter-intuitive results in the aggregation process is that it ignores the attitude characteristics of decision makers. The proposed method modelled the personalized attitudinal characters, reliability and importance of information at the same time, thus having flexibility in reflecting personalized decision structures. As for traditional data fusion problems in which the sources of evidence are small scale, such as target recognition based on multiple sensors, the method proposed in this study can be regarded as an extension of Dempster's rule. This method is more applicable than Dempster's rule because it can solve the problem of decision makers being tolerant of hypotheses that are not supported by some pieces of evidence. Another advantage of the proposed method is that it can solve the problems of data fusion containing large-scale information sources. With the development of network technology, data are usually presented on a large scale. In recent years, large-scale data fusion, especially large-scale group decision making (Liu et al., 2019), has become a hot topic. Conflicts are more intense and complex in large-scale data fusion problems than in small-scale ones because of the wider range of sources of evidence. The proposed method is more suitable for large-scale data fusion problems than other evidence combination methods because it can handle high-conflict problems well and has relatively easy operations.

Some issues remain for further study. The AEC method is limited to relatively monotonous representations of uncertainties; that is, the BPA of any hypothesis is a precise number. In complex situations, the support for a hypothesis may be expressed in interval forms (Vijayabalaji & Rameshb, 2019). An interesting idea for solving such a problem is to integrate the AEC method with mathematical programming methods to estimate the possible evidence combination results. In addition, the proposed method is based on the assumption that hypotheses in a frame of discernment are exhaustive and exclusive. However, this assumption is not realistic for some practical data fusion problems. Concepts that are described in natural language, such as cheap/expensive, happiness/pain and old/young, are vague. There are interactive semantics between each pair of concepts. For example, we cannot give an exact age to distinguish the young from the old. Dezert and Smarandache (2009) theory provided interesting rules of combination to address fusion problems in which hypotheses potentially overlap. In the future, we will enhance the applicability of the AEC method by combining it with the Dezert–Smarandache theory.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes

1. If the weights are set as the coefficients of BPAs, then there is $(w_1 m_1^{at}(B))(w_2 m_2^{at}(C)) = (w_1 w_2)(m_1^{at}(B) m_2^{at}(C)) = (w_1(1-w_1))(m_1^{at}(B) m_2^{at}(C))$ where the weight is unfunctional to describe the importance of an evidence source. Therefore, it is sensible to set the weights as the index of the geometric operator.
2. http://www.gov.cn/fuwu/2019-06/04/content_5397254.htm
3. https://wenku.baidu.com/view/6470916248649b6648d7c1c708a1284ac95005f6.html

Funding

The work was supported by the National Natural Science Foundation of China (71971145, 71771156, 72171158).

ORCID

Xingli Wu http://orcid.org/0000-0003-2265-8754
Huchang Liao http://orcid.org/0000-0001-8278-3384

References

Abellan, J., & Bosse, E. (2020). Critique of recent uncertainty measures developed under the evidence theory and belief intervals. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *50*(3), 1186–1192.

Achroufene, A., Amirat, Y., & Chibani, A. (2019). RSS-based indoor localization using belief function theory. *IEEE Transactions on Automation Science and Engineering*, *16*(3), 1163–1180.

Aggarwal, M., & Tehrani, A. F. (2019). Modelling human decision behaviour with preference learning. *INFORMS Journal on Computing*, *31*(2), 318–334.

Agreh, O. Y., & Ghaffari-Hadigheh, A. (2019). Application of Dempster-Shafer theory in combining the experts' opinions in DEA. *Journal of the Operational Research Society*, *70*(6), 915–925.

Cinelli, M., Kadzinski, M., Gonzalez, M., & Słowinski, R. (2020). How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. *Omega*, *96*, 102261. https://doi.org/10.1016/j.omega.2020.102261

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, *38*(2), 325–339. https://doi.org/10.1214/aoms/1177698950

Dezert, J., & Smarandache, F. (2009). An introduction to DSmT. In F. Smarandache, & J. Dezert (Eds.), *Advances and applications of DSmT for information fusion (collected works)* (pp. 3–73).

Dezert, J., Smarandache, F., & Khoshnevisan, M. (2004). Counter-examples to Dempster's rule of combination. In F. Smarandache, & J. Dezert (Eds.), *Advances and applications of DSmT for information fusion* (pp. 105–121). American Research Press.

Dong, G. G., & Kuang, G. Y. (2015). Target recognition via information aggregation through Dempster–Shafer's evidence theory. *IEEE Geoscience and Remote Sensing Letters*, *12*(6), 1247–1251.

Dubois, D., & Prade, H. (1986). A set-theoretic view on belief functions: Logical operations and approximations by fuzzy sets. *International Journal of General Systems*, *12*, 193–226.

Fang, R., Liao, H. C., Yang, J. B., & Xu, D. L. (2021). Generalised probabilistic linguistic evidential reasoning approach for multi-criteria decision-making under uncertainty. *Journal of the Operational Research Society*, *72*(1), 130–144.

Frikha, A., & Moalla, H. (2015). Analytic hierarchy process for multi-sensor data fusion based on belief function theory. *European Journal of Operational Research*, *241*(1), 133–147.

Fu, C., Yang, J. B., & Yang, S. L. (2015). A group evidential reasoning approach based on expert reliability. *European Journal of Operational Research*, *246*(3), 886–893.

Inagaki, T. (1991). Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory. *IEEE Transactions on Reliability*, *40*(2), 182–188. https://doi.org/10.1109/24.87125
Kang, D., & Park, Y. (2014). Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach. *Expert Systems with Applications*, *41*(4), 1041–1050.
Khorshidi, H. A., & Aickelin, U. (2021). Multicriteria group decision making under uncertainty using interval data and cloud models. *Journal of the Operational Research Society*. https://doi.org/10.1080/01605682.2020.1796541
Li, D. M., Deng, L. B., & Cai, Z. M. (2021). Design of traffic object recognition system based on machine learning. *Neural Computing and Applications*. *33*, 8143–8156. https://doi.org/10.1007/s00521-020-04912-9
Li, T. C., Fan, H. Q., García, J., & Corchado, J. M. (2019). Second-order statistics analysis and comparison between arithmetic and geometric average fusion: Application to multi-sensor target tracking. *Information Fusion*, *51*, 233–243. https://doi.org/10.1016/j.inffus.2019.02.009
Li, F. L., & Lv, Z. H. (2017). Reliable vehicle type recognition based on information fusion in multiple sensor networks. *Computer Networks*, *117*, 76–84.
Liu, B. S., Zhou, Q., Ding, R. X., Palomares, I., & Herrerace, F. (2019). Large-scale group decision making model based on social network analysis: Trust relationship-based conflict detection and elimination. *European Journal of Operational Research*, *275*(2), 737–754.
Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 55–60).
Medjkoune, S., Mouchere, H., Petitrenaud, S., & Viard-Gaudin, C. (2017). Combining speech and handwriting modalities for mathematical expression recognition. *IEEE Transactions on Human-Machine Systems*, *47*(2), 259–272.
Murphy, C. K. (2000). Combining belief functions when evidence conflicts. *Decision Support Systems*, *29*(1), 1–9. https://doi.org/10.1016/S0167-9236(99)00084-6
Qi, J., Yang, P., Newcombe, L., Peng, X. Y., Yang, Y., & Zhao, Z. (2019). An overview of data fusion techniques for internet of things enabled physical activity recognition and measure. *Information Fusion*, *55*, 269–280.
Rezaei, J. (2016). Best-worst multi-criteria decision-making method: Some properties and a linear model. *Omega*, *64*, 126–130. https://doi.org/10.1016/j.omega.2015.12.001
Roy, P., Chowdhury, C., Kundu, M., Ghosh, D., & Bandyopadhyay, S. (2021). Novel weighted ensemble classifier for smartphone based indoor localization. *Expert Systems with Applications*, *164*, 113758. https://doi.org/10.1016/j.eswa.2020.113758
Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*, 234–281. https://doi.org/10.1016/0022-2496(77)90033-5
Sentz, K., & Ferson, S. (2002). *Combination of evidence in Dempster–Shafer theory. Technical report, SAND 2002-0835*. Sandia National Laboratories.
Shafer, G. (1976). *A mathematical theory of evidence* (Vol. 1). Princeton University Press.
Smarandache, F., Dezert, J., & Tacnet, J. M. (2010). Fusion of sources of evidence with different importances and reliabilities. In *BELIEF 2010: Workshop on the Theory of Belief Functions*.
Vijayabalaji, S., & Rameshb, A. (2019). Belief interval-valued soft set. *Expert Systems with Applications*, *119*, 262–271.
Wu, X. L., & Liao, H. C. (2021). Learning judgment benchmarks of customers from online reviews. *Or Spectrum*. https://doi.org/10.1007/s00291-021-00639-8
Xiao, F. (2020). A new divergence measure for belief functions in D–S evidence theory for multisensor data fusion. *Information Sciences*, *514*, 462–483. https://doi.org/10.1016/j.ins.2019.11.022
Yager, R. R. (1980). Competitiveness and compensation in decision making: A fuzzy set based interpretation. *Computers & Operations Research*, *7*(4), 285–300.
Yager, R. R. (1987). Quasi-associative operations in the combination of evidence. *Kybernetes*, *16*(1), 37–41. https://doi.org/10.1108/eb005755
Yang, J. B., & Xu, D. L. (2013). Evidential reasoning rule for evidence combination. *Artificial Intelligence*, *205*, 1–29.
Zadeh, L. (1986). A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *AI Magazine*, *7*(2), 85–90.
Zhang, L. (1994). Representation, independence, and combination of evidence in the Dempster-Shafer theory. In R. R. Yager, J. Kacprzyk, & M. Fedrizzi (Eds.), *Advances in the Dempster-Shafer theory of evidence* (pp. 51–69). John Wiley and Sons.

Appendix

Theorem 1. **(Idempotence)** *Let a complete piece of evidence under the frame of discernment,* $\Theta = \{H_1, H_2, \ldots, H_N\}$, *be* $M = \{H_l, m(H_l)\} | H_l \in \Theta, \sum_{l=1}^{N} m(H_l) = 1\}$, *and let* $\theta = 0$ *and* $\sigma = 1$. *Then,* $M \oplus M = M$.

Proof. In this case, $w = 0.5$, and there is a condition that $H_l \cap H_{l'} = \emptyset$. By Eq. (12), if $m(H) > 0$, then the support degree of the hypothesis H is $m'(H) = m(H)^{0.5} m(H)^{0.5} = m(H)$. After normalization, we obtain $m_{\sigma w(2)}(H) = m(H)$ if $m(H) > 0$, and $m_{\sigma w(2)}(H) = 0$ if $m(H) = 0$. The proof is completed.

It should be noted that if $\theta \to 0$, then $M \oplus M \approx M$. That is, when $\theta \to 0$, there is a larger tolerant degree to the hypothesis that is strongly supported in some evidence but weakly supported (or opposed) in others than $\theta = 0$ in the combination process. Then, the support degree of the hypothesis with small BPAs would increase after combination. In addition, if there is $B \cap C \neq \emptyset, \exists B, C \in M$, then Eq. (12) does not satisfy the idempotence of classical mathematics, *i.e.* $M \oplus M \neq M$. That is, when there is an intersection between focal elements, the intersected elements will gain more support after combination. Especially, if there are missing BPAs in evidence, with the increase of the number of sources, the support to the focal elements will increase, and the uncertainty (missing BPAs) will decrease. This feature is in line with the fact in decision making that as the number of experts increases, the missing decision information decreases. This is an important advantage of the D-S theory in combining uncertain information, and it is a critical difference from the pure probability theory.

Theorem 2. **(Commutativity)** *Let* $M_1 = \{H, m_1(H)\} | \forall H \subseteq \Theta\}$ *and* $M_2 = \{H, m_2(H)\} | \forall H \subseteq \Theta\}$ *be two pieces of evidence with the weight vector* $W = \{w_1, w_2\}^T, w_1 + w_2 = 1$, *and the reliability vector* $\sigma = \{\sigma_1, \sigma_2\}^T$. *There is* $M_1 \oplus M_2 = M_2 \oplus M_1$.

Proof. If $m_1(H) = 0$ and $m_2(H) = 0$, then $[m^{at}_{\sigma w1} \oplus m^{at}_{\sigma w2}](H) = [m^{at}_{\sigma w2} \oplus m^{at}_{\sigma w1}](H) = 0$. If $m_1(H)>0$ or $m_2(H)>0$, there is $(\sigma_1 m_1(H) + \theta)^{w_1} (\sigma_2 m_2(H) + \theta)^{w_2} - \theta = (\sigma_2 m_2(H) + \theta)^{w_2} (\sigma_1 m_1(H) + \theta)^{w_1} - \theta$ and $(\sigma_1 m_1(H) + \theta)^{w_1} (m_2(H) + 1 - \sigma_2 + \theta)^{w_2} - \theta = (m_2(H) + 1 - \sigma_2 + \theta)^{w_2} (\sigma_1 m_1(H) + \theta)^{w_1} - \theta$. Thus, we can obtain $\sum_{B \cap C = H} ((m^{at}_{\sigma 1}(B))^{w_1} (m^{at}_{\sigma 2}(C))^{w_2} - \theta) = \sum_{B \cap C = H} ((m^{at}_{\sigma 2}(C))^{w_2} (m^{at}_{\sigma 1}(B))^{w_1} - \theta)$. Therefore, $[m^{at}_{\sigma w1} \oplus m^{at}_{\sigma w2}](H) = [m^{at}_{\sigma w2} \oplus m^{at}_{\sigma w1}](H)$. The proof is completed.

Theorem 3. **(Associativity)** *Let* $M_1 = \{H, m_1(H)\} | \forall H \subseteq \Theta\}$, $M_2 = \{H, m_2(H)\} | \forall H \subseteq \Theta\}$ *and* $M_3 = \{H, m_3(H)\} | \forall H \subseteq \Theta\}$ *be three pieces of evidence with the weight vector* $W = \{w_1, w_2, w_3\}^T, w_1 + w_2 + w_3 = 1$, *and the reliability vector* $\sigma = \{\sigma_1, \sigma_2, \sigma_3\}^T$. *There is* $M_1 \oplus (M_2 \oplus M_3) = (M_1 \oplus M_2) \oplus M_3$.

Proof. If $m_1(H) = 0$, $m_2(H) = 0$ and $m_3(H) = 0$, then $[m^{at}_{\sigma w1} \oplus (m^{at}_{\sigma w2} \oplus m^{at}_{\sigma w3})](H) = [(m^{at}_{\sigma w1} \oplus m^{at}_{\sigma w2}) \oplus m^{at}_{\sigma w3}](H) = 0$. If $m_1(H)>0$, $m_2(H)>0$ or $m_3(H) >0$, then,

$$\begin{aligned}&(\sigma_1 m_1(H) + \theta)^{w_1} ((\sigma_2 m_2(H) + \theta)^{w_2} (\sigma_3 m_3(H) + \theta)^{w_3})\\ &\quad -(m_1(H) + 1 - \sigma_1 + \theta)^{w_1} ((\sigma_2 m_2(H) + \theta)^{w_2} (\sigma_3 m_3(H) + \theta)^{w_3})\\ &\quad -\theta = ((m_1(H) + 1 - \sigma_1 + \theta)^{w_1} (\sigma_2 m_2(H) + \theta)^{w_2}))(\sigma_3 m_3(H) + \theta)^{w_3}\\ &\quad -\theta = ((\sigma_1 m_1(H) + \theta)^{w_1} (\sigma_2 m_2(H) + \theta)^{w_2}))(\sigma_3 m_3(H) + \theta)^{w_3} - \theta;\end{aligned}$$

Thus, we have $\sum_{C_1 \cap (C_2 \cap C_3) = H} (\prod_j^3 (m^{at}_{\sigma j}(C_j))^{w_j} - \theta) = \sum_{(C_1 \cap C_2) \cap C_3 = H} (\prod_j^3 (m^{at}_{\sigma j}(C_j))^{w_j} - \theta)$.

Therefore, $[m^{at}_{\sigma w1} \oplus (m^{at}_{\sigma w2} \oplus m^{at}_{\sigma w3})](H) = [(m^{at}_{\sigma w1} \oplus m^{at}_{\sigma w2}) \oplus m^{at}_{\sigma w3}](H) = 0$. The proof is completed.

Theorem 4. **(Continuity)** *Let* $M_1 = \{H, m_1(H)\} | \forall H \subseteq \Theta\}$ *and* $M_2 = \{H, m_2(H)\} | \forall H \subseteq \Theta\}$ *be two pieces of evidence, and the approximate evidence of* M_1 *be* $M'_1 = \{H, m'_1(H)\} | \forall H \subseteq \Theta\}$ *such that 1)* $m'_1(H) \approx m_1(H)$ *for* $m_1(H)>0$, *and 2)* $m'_1(H) = m_1(H)$ *for* $m_1(H) = 0$. *There is* $M_1 \oplus M_2 \approx M'_1 \oplus M_2$.

Proof. For the simplicity of calculation, the weights of evidence are not considered here. Suppose that $m'_1(H) = m_1(H) + \varepsilon$ with $\varepsilon \to 0$.

1. If $m'_1(H) = 0$ and $m_2(H) = 0$, then, $[m'^{at}_{\sigma 1} \oplus m^{at}_{\sigma 2}](H) = [m^{at}_{\sigma 1} \oplus m^{at}_{\sigma 2}](H) = 0$.
2. If $m'_1(H)>0$ or $m_2(H)>0$, there is

$$\begin{aligned}&(\sigma_1 m'_1(H) + \theta)(\sigma_2 m_2(H) + \theta) - \theta^2\\ &= \sigma_1 \sigma_2 m'_1(H) m_2(H) + \theta(\sigma_1 m'_1(H) + \sigma_2 m_2(H))\\ &= \sigma_1 \sigma_2 (m_1(H) + \varepsilon) m_2(H) + \theta(\sigma_1 (m_1(H) + \varepsilon)\\ &\quad + \sigma_2 m_2(H)) = \sigma_1 \sigma_2 m_1(H) m_2(H)\\ &\quad + \theta(\sigma_1 m_1(H) + \sigma_2 m_2(H)) + \sigma_1 \sigma_2 \varepsilon m_2(H)\\ &\quad + \theta \sigma_1 \varepsilon \approx \sigma_1 \sigma_2 m_1(H) m_2(H) + \theta(\sigma_1 m_1(H) + \sigma_2 m_2(H))\\ &= (\sigma_1 m_1(H) + \theta)(\sigma_2 m_2(H) + \theta) - \theta^2.\end{aligned}$$

Thus, we can get $\sum_{B \cap C = H} ((\sigma_1 m'_1(B) + \theta)(\sigma_2 m_2(C) + \theta) - \theta^2) \approx \sum_{B \cap C = H} ((\sigma_1 m_1(B) + \theta)(\sigma_2 m_2(C) + \theta) - \theta^2)$.

Therefore, $[m'^{at}_{\sigma 1} \oplus m^{at}_{\sigma 2}](H) \approx [m^{at}_{\sigma 1} \oplus m^{at}_{\sigma 2}](H)$. When considering the weights of evidence, we can obtain the same conclusion since the exponential function does not change the continuity of the original function. The proof is completed.