

The Combining Rule of Dempster-Shafer Theory for Correlative Evidence

Ping Wang

Electric and Information College, Xihua University, Chengdu 610039, P. R. China

Abstract

This paper puts forward an improvement method for the combining rule of Dempster-Shafer evidence theory based on the correlation between evidences. Different from D-S theory, the correlation coefficient between evidences is introduced. By decreasing the probability evaluation of certainty and increasing the probability evaluation of uncertainty, the effect of correlative evidence to fusion result is reduced, and then we can get the fusion result which is closer to the fact. Finally, an example is provided to show the advantage of this method.

Keywords: Data fusion, Correlation coefficient, D-S evidence theory

1. Introduction

Dempster-Shafer (D-S) theory is a kind of reasoning algorithm based on evidence theory. It was put forward by Dempster at first, and developed by Shafer, so it was called D-S evidence theory [1][2]. The algorithm solved two difficult problems of probability: it can give clear expression to unknown case; and the support to negative supposition of evidence can express by some value if the evidence supports a part of supposition. D-S evidence theory is used to deal with some inexact information or inexact information sometimes

Suppose Ω is the aggregate of all probable elements of variable X , and the elements of Ω are mutually exclusive, so we call Ω sample space. If the number of elements in Ω is N , then the elements of power aggregate of Ω (2^Ω) are 2^N . Each element of power aggregate is corresponding to a proposition (subset) of value of X .

For each subset A (proposition) in Ω , let it corresponding to a number $M \in [0, 1]$, thus meeting:

(1) $M(\emptyset)=0$, \emptyset is called empty aggregate or impossible case

(2) $\sum_{A \subseteq \Omega} M(A)=1$

So we call function M basic probability assignment function on 2^Ω , $M(A)$ is the basic probability evaluation of A .

If $\bar{A}=\Omega-A$, It is necessary to point out that we can't infer the belief degree of \bar{A} from A . Consequently, for probability, $P(\bar{A})+P(A)=1$; but for D-S evidence theory, $M(A)+M(\bar{A})\leq 1$. It is obviously that $M(A)$ is not probability.

If A is a subset of Ω , and $M(A)>0$, A is called focal Element of belief function BEL, the union of all focal elements is called core.

Define the belief function of proposition as:

$$\text{BEL}: 2^\Omega \rightarrow [0, 1]$$
$$\text{BEL}(A) = \sum_{B \subseteq A} M(B), \quad \forall A \subseteq \Omega$$

Where $\bar{A}=\Omega-A$

$\text{BEL}(A)$ represents the sum of probability of all subsets of A ; showing the whole belief degree or reliability of A .

Then

$$\text{BEL}(\emptyset)=0$$
$$\text{BEL}(\Omega)=1$$

Define the Plausibility Function of proposition: $2^\Omega \rightarrow [0, 1]$

$$\text{PL}(A)=1-\text{BEL}(\bar{A}) = \sum_{B \cap A \neq \emptyset} M(B), \quad \forall A \subseteq \Omega$$

Where:

$\bar{A}=\Omega-A$

$\text{PL}(A)$ represents the belief degree that we don't refuse A , It is the sum of all basic probability evaluation of sets which intersect with A .

$\text{BEL}(A)$ and $\text{PL}(A)$ are the lower limit and upper limit function of proposition A , separately. They meet $\text{PL}(A)\geq \text{BEL}(A)$, if $\text{PL}(A)=\text{BEL}(A)$, D-S theory turns into Bayes theory.

In fact, among all evidences about proposition A , there are some support evidences which stand by A , and some are negative evidences which refuse A or stand by \bar{A} , there are other uncertain evidences which don't support or refuse A . $\text{BEL}(A)$ is maximum reliability which can be computed from the support evidences of A , $\text{PL}(A)$ is maximum plausibility which can be computed from the evidences which don't refuse A (support evidences and uncertain evidences). we can use $\text{BEL}(A)$ and $\text{PL}(A)$ to measure the uncertainty of proposition A . $\text{BEL}(A)$ and $\text{PL}(A)$ are the lower limit and upper limit of $P(A)$ separately, i.e.

$$\text{BEL}(A)\leq P(A)\leq \text{PL}(A)$$

$PL(A)$ - $BEL(A)$ shows the uncertainty of proposition A. The smaller the value is, the clearer the evidence is to support proposition A.

2. The combining rule of D-S evidence theory

The combining rule of evidence theory provides a principle to combine two evidences [3]-[5]. Suppose M_1 and M_2 are two independent basic probability evaluation on 2^Ω , BEL_1 and BEL_2 are two belief function on 2^Ω , M_1 and M_2 are corresponding basic probability evaluation separately, focal elements are A_1, \dots, A_k and B_1, \dots, B_r , further more:

$$K_1 = \sum_{i,j} M_1(A_i)M_2(B_j) < 1$$

So we define the combining rule of D-S evidence theory as:

$$M(C) = \begin{cases} \frac{\sum_{i,j} M_1(A_i)M_2(B_j)}{1 - K_1} & \forall C \subset U \quad C \neq \emptyset \\ 0 & C = \emptyset \end{cases}$$

Provided that $K_1 \neq 1$, M makes a certain basic probability evaluation; provided that $K_1=1$, M_1 and M_2 conflict, and the combination can't achieve.

Suppose $A, B \subseteq U$, the confidence interval of A and B are separately:

$$EI_1(A) = [BEL_1(A), PL_1(A)]$$

And

$$EI_2(B) = [BEL_2(B), PL_2(B)]$$

So the confidence interval after combination is

$$\begin{aligned} EI_1(A) \oplus EI_2(B) \\ = [1 - K_2(1 - BEL_1(A))(1 - BEL_2(B)), K_2 PL_1(A) PL_2(B)] \end{aligned}$$

Where

$$K_2 = \left\{ 1 - [BEL_1(A) \overline{BEL}_2(\bar{B}) \overline{BEL}_1(\bar{A}) \overline{BEL}_2(\bar{B})] \right\}^{-1}$$

D-S combining evidence theory makes an analysis based on two evidences. For the combination of multi-evidence, it infers result by the combination of any two evidences.

3. The problem of correlative evidence fusion

D-S combining evidence theory regards all sorts of source data as same importance, and neglects the correlation of data. D-S combining evidence can make correct decision if all evidences are independent, but when evidences are correlative, there may be false result.

For example, for fault diagnosis, suppose there are three persons: expert, old worker and student, they diagnose the same fault which the output of data acquisition system is 0, and their diagnosis results as follow:

Suppose event A is sensor error, event B is AD error, event C is normal.

$$\text{Expert(E): } M_E(A)=0.7, \quad M_E(\Omega)=0.3$$

Due to the authority of expert, the others have the same diagnosis with him.

$$\text{Old worker (W): } M_W(A)=0.7, \quad M_W(\Omega)=0.3$$

$$\text{Student(S): } M_S(A)=0.7, \quad M_S(\Omega)=0.3$$

The result of instrument test is

$$M_I(B)=0.7, \quad M_I(\Omega)=0.3$$

According to D-S combining rule, fuse their diagnosis results; at first we combine the diagnosis of expert and old worker.

	$M_E(A)=0.7$	$M_E(\Omega)=0.3$
$M_W(A)=0.7$	(A)0.49	(A)0.21
$M_W(\Omega)=0.3$	(A)0.21	(Ω)0.09

Table 1: Fusion of expert and worker.

The result after fusion is

$$M_I(A)=M_E(A) \oplus M_W(A)=0.91$$

$$M_I(\Omega)=M_E(\Omega) \oplus M_W(\Omega)=0.09$$

The confidence interval of event A (sensor error) is [0.91, 1.0].

Then combine the result with the diagnosis of student.

	$M_I(A)=0.91$	$M_I(\Omega)=0.09$
$M_S(A)=0.7$	(A)0.637	(A)0.063
$M_S(\Omega)=0.3$	(A)0.273	(Ω)0.027

Table 2: Fusion of first result and student.

Consequently, the final result after fusion is

$$M_I(A)=0.637+0.273+0.063=0.973$$

$$M_I(\Omega)=0.027$$

Then combine the result with the instrument test.

	$M_I(A)=0.91$	$M_I(\Omega)=0.09$
$M_I(B)=0.7$	(Ø) 0.637	(B) 0.063
$M_I(\Omega)=0.3$	(A) 0.273	(Ω) 0.027

Table 3: Fusion of persons and instrument.

Consequently, the final result after fusion is

$$M(A)=0.273$$

$$M(B)=0.063$$

$$M(\Omega)=0.027$$

$M(\emptyset)=0$ (according to the define of empty aggregation)

The disagreement coefficient of evidence is

$$K = 0.637$$

Standardize the focal element (divided by coefficient $(1 - K)$), the fusion result is

$$M(A)=0.273/0.363=0.76$$

$$M(B)=0.063/0.363=0.17$$

$$M(\Omega)=0.027/0.363=0.07$$

So the confidence interval of event A is [0.76, 0.83], and the confidence interval of event B is [0.17, 0.24].

With the expert's diagnosis and the others' support, though the expert is uncertain to his diagnosis, the fusion result is certainty. Though the result of instrument test is opposition with the expert's diagnosis, the fusion result can't be changed. We know though the expert is correct in many cases, the diagnosis of instrument is obviously more credible. With the DS combining theory, the result of fusion is always decided by majority supporters. It is obviously that the fusion result is unreliable if we just fuse the evidence regardless of the correlation between them. By this reason, we need to eliminate the effect of correlation between evidences before fusion.

4. The combining rule of D-S theory for correlative evidence

In fact, there is very few occasional cases that multi-evidence is independent each other. Generally, the multi-source data is correlative, so we need to eliminate the effect of correlation between evidences before fusion. Therefore we put forward a new combination method that based on D-S theory, which eliminates the correlation between evidences according to their correlation coefficient.

For the correlative evidence, we should decrease the evidence belief and increase the probability evaluation of uncertain evidence. The new combination method that based on correlation coefficient may draw a more reliable conclusion than previous. So we can achieve fusion result that is closer to truth.

For this method, we suppose that:

- There are many evidences for focal element A, they are 1,2...N, the basic probability evaluation of each evidence is known, and it is $M_i(A)$, $i \in N$
- According to the experiment result or correlation predict, the correlation coefficient between evidences r_{ij} is known, and $r_{ij} \leq 1$.
- Eliminate the correlation according to the correlation coefficient of evidence for one of the two combining evidences. The basic probability evaluation after eliminated correlation is $(1 - |r_{ij}|)M_i$
- Because the probability evaluation of certainty is decreased, the probability evaluation of uncertainty is increased, the incremental part is $(|r_{ij}|)M_i$

The main difference between the correlation combine evidence method and D-S combining evidence is that the correlation combining method will amend the evidence according to its correlation coefficient, and then the sharp effect of the repeat evidence on the final result of evidence fusion can be avoided.

Suppose that the belief function and plausibility function are separately:

$$BEL_i(A) \text{ and } PL_i(A), \quad i \in N,$$

So the combining rule of D-S Theory for correlative evidence is:

$$M(C) = \begin{cases} \frac{\sum_{\substack{i,j \\ A_i \cap B_j = C}} (1 - |r_{ij}|) M_1(A_i) M_2(B_j)}{1 - K_1} & \forall C \subset U \quad C \neq \emptyset \\ 0 & C = \emptyset \end{cases}$$

Where

$$K_1 = \sum_{\substack{i,j \\ A_i \cap B_j \neq \emptyset}} (1 - |r_{ij}|) M_1(A_i) M_2(B_j) < 1$$

Suppose that $A, B \subseteq U$, the confidence interval of A and B is separately:

$$EI_1(A) = [BEL_1(A), PL_1(A)]$$

And

$$EI_2(B) = [BEL_2(B), PL_2(B)]$$

So the confidence interval after combination is

$$EI_1(A) \oplus EI_2(B)$$

$$= [1 - K_2(1 - BEL_1(A))(1 - BEL_2(B)), K_2 PL_1(A) PL_2(B)]$$

Where

$$K_2 = \left\{ 1 - [BEL_1(A) \overline{BEL}_2(\overline{B}) \overline{BEL}_1(\overline{A}) \overline{BEL}_2(\overline{B})] \right\}^{-1}$$

5. The application of DS combining rule for correlative evidence

We analyze the foretype with the D-S combining rule for correlative evidence, and then we can infer the result.

According to their experimental results of diagnosis, we infer the diagnosis correlation coefficients of 3 persons for Data Acquisition System fault are

Expert and old worker: $r_{EW}=0.8$.

Between the above result and Student: $r_{(EW)S}=0.9$.

The instrument and person: $r_{IP}=0$.

At first, we combine the diagnosis of expert and old worker.

By the D-S combining rule for correlative evidence, eliminate the correlation according to the correlation coefficient of evidence for one of the two combining evidences, then

Old worker:

$$M_W(A) = (1 - |r_{EW}|) \times 0.7 = 0.14$$

$$M_W(\Omega) = 0.3 + |r_{EW}| \times 0.7 = 0.86$$

	$M_E(A)=0.7$	$M_E(\Omega)=0.3$
$M_W(A)=0.14$	(A)0.098	(A)0.042
$M_W(\Omega)=0.86$	(A)0.602	(Ω)0.258

Table 4: Correlative fusion of expert and worker.

The result after fusion is

$$M_1(A)=M_W(A) \oplus M_E(A)=0.742$$

$$M_1(\Omega)=M_W(\Omega) \oplus M_E(\Omega)=0.258$$

And then, we fuse the diagnosis of student and the fusion result of the first step.

Student:

$$M_S(A)=(1-|r_{(EW)S}|) \times 0.7=0.07$$

$$M_S(\Omega)=0.3+|r_{(EW)S}| \times 0.7=0.93$$

	$M_1(A)=0.742$	$M_1(\Omega)=0.258$
$M_S(A)=0.07$	(A)0.052	(A)0.018
$M_S(\Omega)=0.93$	(A)0.69	(Ω)0.24

Table 5: Correlative fusion of persons and student.

The result after fusion is

$$M_2(A)=M_1(A) \oplus M_S(A)=0.76$$

$$M_2(\Omega)=M_1(\Omega) \oplus M_S(\Omega)=0.24$$

At last, we fuse the diagnosis of instrument and the fusion result of the second step. The correlation coefficient is 0.

	$M_2(A)=0.76$	$M_2(\Omega)=0.24$
$M_I(B)=0.7$	(Ø)0.532	(B)0.168
$M_I(\Omega)=0.3$	(A)0.228	(Ω)0.072

Table 6: Correlative fusion of second result and instrument.

The result after fusion is

$$M(A)=M_2(A) \oplus M_I(\Omega)=0.228$$

$$M(B)=M_2(\Omega) \oplus M_I(B)=0.168$$

$$M(\Omega)=M_2(\Omega) \oplus M_I(\Omega)=0.072$$

$M(\emptyset)=0$ (according to the define of empty aggregation)

The disagreement coefficient of evidence is:
 $K=0.532$

Standardize the focal element (divided by coefficient (1 - K))

$$M(A)=0.228/(1-K)=0.49$$

$$M(B)=0.36$$

$$M(\Omega)=0.15$$

So the confidence interval of event A is [0.49, 0.64], and the confidence interval of event A is [0.36, 0.51].

Though there are 3 persons support the event A, but they are correlative. With the conflict independent evidence from instrument, the confidence interval for event A is decreased from [0.76, 0.83] to [0.49, 0.64]. Compared with the D-S combining rule, it is obviously that the result of DS combining rule for correlative

evidence is more credible. The method is helpful to make a correct fusion decision.

6. Conclusions

For independent evidences, the fusion result of two methods makes no difference. But there are few cases which two combining evidences are independent at all, the difference makes D-S combining rule for correlative evidence very important. The improvement method which eliminates correlation is better to gain a correct result.

Because there are many evidences for each event, the correlation coefficient of between two evidences is differently. The sample in this paper is only a demonstration for the improved method, when we use D-S combining rule for correlative evidence, the correlation coefficient between two evidences needs to be decided by experiment or prediction.

Acknowledgement

The Project Supported by Scientific Research Fund of Sichuan Provincial Education Department (Grant No. 2006A088), and partially supported by Sichuan Province Key Nature Science & Technology Foundation (Grant No.05JY029-136).

References

- [1] P. Wang and G.Q. Yang, Improvement method for the combining rule of Dempster-Shafer evidence theory based on reliability, *Journal of Systems Engineering and Electronics*, 2:471-474, 2005.
- [2] Richards, A. John and X. P. Jia, A Dempster-Shafer relaxation approach to context classification, *IEEE Transactions on Geoscience and Remote Sensing*, 5:1422-1431, 2007.
- [3] Altincay and Hakan, On the independence requirement in Dempster-Shafer theory for combining classifiers providing statistical evidence, *Applied Intelligence*, 8:73-90, 2006.
- [4] J.W. Guan, A generalization of the dempster-shafer theory, *Proc. of Inter. Joint Conf. on Artificial Intelligence (IJCAI93)*, pp.529-579, 1993.
- [5] J.H. Lee, Combining the evidence of different relevance feedback methods for information retrieval, *Information Processing and Management*, 6:681-691, 1998.
- [6] P. Smets, The combination of evidence in the transferable belief model, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12:447-458, 1990.