Measure divergence degree of basic probability assignment based on Deng relative entropy

Liguo Fei[a], Yong Deng[a,b,c,*]

[a] School of Computer and Information Science, Southwest University, Chongqing, 400715, China
[b] School of Electronics and Information, Northwestern Polytechnical University, Xian, Shaanxi 710072, China
[c] School of Engineering, Vanderbilt University, Nashville, TN, 37235, USA

Abstract

Dempster Shafer evidence theory (D-S theory) is more and more extensively applied to information fusion for the advantage dealing with uncertain information. However, the results opposite to common sense are often obtained when combining the different evidence using the Dempster's combination rules. How to measure the divergence between different evidence is still an open issue. In this paper, a new relative entropy named as Deng relative entropy is proposed in order to measure the divergence between different basic probability assignments (BPAs). The Deng relative entropy is the generalization of Kullback-Leibler Divergence because when the BPA is degenerated as probability, Deng relative entropy is equal to Kullback-Leibler Divergence. Numerical examples are used to illustrate the effectiveness of the proposed Deng relative entropy.

*Corresponding author: Yong Deng, School of Computer and Information Science, Southwest University, Chongqing, 400715, China.

Email address: ydeng@swu.edu.cn, ydeng@nwpu.edu.cn, prof.deng@hotmail.com (Yong Deng)

Preprint submitted to Elsevier *October 19, 2015*

Keywords: Dempster-Shafer evidence theory, relative entropy, Deng entropy, Deng relative entropy, divergence

---

1. Introduction

Dempster-Shafer evidence theory (D-S theory) [1, 2] has attracted the extensive attention of researchers with its great advantage to handle and combine uncertain information. This theory is widely used in object classification [3, 4], decision making [3, 5, 6, 7, 8, 9, 10, 11, 12], risk assessment [13], information fusion [14, 15]. However, the counter-intuitive conditions often occur when fusing the high conflicting evidence using the Dempster's combination rules [16, 17]. This kind of counter-intuitive results have a serious influence for the accuracy of evidence fusion.

It is so necessary and significant for researchers to remedy this weakness of Dempster's [1] combination rules. To improve this shortcoming, a series of alternative combination rules are presented [15, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31] currently. Generally speaking, there exist two categories of methods to deal with this problem. One is to improve the Dempster's combination rules and to reallocate the conflict. For example, in [18, 19], Lefevre used the part of the conflicting evidence and distributed the conflict into the focal element sets of all the evidence proportionally. In [24], the conflict of evidence is abandoned to utilize because Yager believe it is useless and distribute them into the universal set. However, sometimes it enlarges the uncertainty of evidence and gets the unreasonable fusion results. And the other one is to modify the conflicting evidences before the fusion. Schubert [15] and Han [28] proposed the modified algorithms to obtain the weights of evidence. In [20], Deng proposed a method about the evidence support based on the Jousselme distance function and determine a weighted average of all the evidence. In [26], Murphy presented a problem, the failure to balance multiple evidence, then illustrated the proposed solutions and described their limitations.

All of these methods can improve the fusion results in part and make up some weakness of D-S theory from a different perspective. However, some essence is ignored to figure out this problem for a long time. To resolve the problem in essence, in this paper, a new relative entropy is proposed named Deng relative entropy which is a generalized relative entropy to measure divergence between BPAs.

The remainder of this paper is constituted as follows. Section 2 introduces the D-S theory and its basic rules and some necessary related concepts about relative entropy. The proposed method of Deng relative entropy is presented in Section 3. Section 4 presented and analyzed the experimental results. Conclusion is given in Section 5.

2. Preliminaries

2.1. Dempster-Shafer evidence theory

Dempster-Shafer evidence theory (D-S theory) is proposed by Dempster [1] and developed later by Shafer[2]. This theory extends the elementary event space in probability theory to its power set named as frame of discernment and constructs the basic probability assignment(BPA) on it.

In addition, there is a combination rule presented by Dempster to fuse different BPAs. In particular, D-S theory can definitely degenerate to the probability theory if the belief is only assigned to single elements. The basic definitions about D-S theory is shown as follows:

2.1.1. *Frame of discernment*

D-S theory supposes the definition of a set of elementary hypotheses called the frame of discernment, defined as:

$$\theta = \{H_1, H_2, ..., H_N\} \tag{1}$$

That is, θ is a set of mutually exclusive and collectively exhaustive events. Let us denote 2^θ the power set of θ.

2.1.2. *Mass functions*

When the frame of discernment is determined, a mass function m is defined as follows.

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

which satisfies the following conditions:

$$m(\phi) = 0 \tag{3}$$

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

In D-S theory, a mass function is also called a basic probability assignment (BPA).

2.1.3. *Dempster's rule of combination*

In a real system, there may be many evidence originating from different sensors, so we can get different BPAs. Dempster [1] proposed orthogonal sum to combine these BPAs. Suppose m_1 and m_2 are two mass functions. The Dempster's rule of combination denoted by $m = m_1 \bigoplus m_2$ is defined as follows:

$$m(A) = \frac{\sum_{B \bigcap C = A} m_1(B) m_2(C)}{1 - K} \tag{5}$$

with

$$K = \sum_{B \bigcap C = \phi} m_1(B) m_2(C) \tag{6}$$

Note that the Dempster's rule of combination is only applicable to such two BPAs which satisfy the condition $K < 1$.

2.2. *Kullback-Leibler Divergence*

In probability theory and information theory, the Kullback-Leibler divergence [32, 33] (also information divergence, information gain, relative entropy, or KLIC) is a non-symmetric measure of the difference between two probability distributions P and Q. KL measures the expected number of extra bits required to code samples from P when using a code based on Q, rather than using a code based on P. Typically P represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution. The measure Q typically represents a theory, model, description, or approximation of P. And its definition is shown as follows:

$$D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \tag{7}$$

where P and Q are the probability distributions and usually have the same type.

In words, it is the average of the logarithmic difference between the probabilities P and Q, where the average is taken using the probabilities P. The K-L divergence is only defined if P and Q both sum to 1 and if $Q(i) > 0$ for any i such that $P(i) > 0$. If the quantity $0 \times log(0)$ appears in the formula, it is interpreted as zero.

There are some properties for the K-L divergence.

1. The KullbackCLeibler divergence is always non-negative,

$$D_{KL}(P||Q) \geq 0$$

a result known as Gibbs' [34] inequality, with $D_{KL}(P||Q)$ zero if and only if $P = Q$.

2. The Kullback-Leibler divergence is additive for independent distributions in much the same way as Shannon entropy. If $P1$, $P2$ are independent distributions, with the joint distribution $P(x, y) = P1(x)P2(y)$, and Q, $Q1$, $Q2$ likewise, then

$$D_{KL}(P||Q) = D_{KL}(P_1||Q_1) + D_{KL}(P_2||Q_2)$$

3. Deng relative entropy

There are some open issues in D-S theory still unresolved, and how to measure the discrepancy and conflict of two evidence is the key step which is very important for evidence fusion. It is obvious that D-S theory is the generalization of probability theory, and combined with the Kullback-Leibler Divergence mentioned above, A new relative entropy named Deng relative entropy are defined as follows:

$$D_d(m_1||m_2) = \sum_i m_1(F_i) \log \frac{m_1(F_i)}{m_2(F_i)} \tag{8}$$

where F_i is a proposition in mass function m_1 and m_2, respectively. The Deng relative entropy is similar with Kullback-Leibler Divergence in form, but it uses mass functions instead of probability distribution functions. Specially, the BPA will turn into probability if it's only assigned to single elements, and Deng relative entropy will also degenerate to Kullback-Leibler at the same time. The same property can be reasoned out shown as follows.

The Deng relative entropy is always non-negative,

$$D_d(m_1||m_2) \geq 0$$

The $D_d(m_1||m_2)$ zero if and only if $m_1 = m_2$, similarly.

4. Numerical examples and discussions

Example 4.1. *Let us suppose a frame of discernment X=*$\{\theta_1, \theta_2, \theta_3\}$*,two evidence's mass function in three differen conditions as follows:*

Condition 1*:* $m_1(\theta_1)$ = *0.4,* $m_1(\theta_2)$ = *0.3,* $m_1(\theta_3)$ = *0.3;* $m_2(\theta_1)$ = *0.4,* $m_2(\theta_2)$ = *0.3,* $m_2(\theta_3)$ = *0.3*

Condition 2*:* $m_1(\theta_1, \theta_2)$ = *0.4,* $m_1(\theta_1, \theta_3)$ = *0.6 ;* $m_2(\theta_1, \theta_2)$ = *0.4,* $m_2(\theta_1, \theta_3)$ = *0.6*

Condition 3*:* $m_1(\theta_1, \theta_2, \theta_3)$ = *1 ;* $m_2(\theta_1, \theta_2, \theta_3)$ = *1*

the calculation process of condition 1, 2 and 3 of Deng cross entropy as follows:

condition 1. $D_d = 0.4 \times \log \frac{0.4}{0.4} + 0.3 \times \log \frac{0.3}{0.3} + 0.3 \times \log \frac{0.3}{0.3} = 0$

condition 2. $D_d = 0.6 \times \log \frac{0.6}{0.6} + 0.4 \times \log \frac{0.4}{0.4} = 0$

condition 3. $D_d = 1 \times \log \frac{1}{1} = 0$

From the example 4.1, it can be seen that the Deng relative entropy is zero for the same BPA.

Example 4.2. *Let us suppose a frame of discernment X=*$\{\theta_1, \theta_2\}$*, two evidence's mass function shown as follows:*

$m_1(\theta_1) = a$, $m_1(\theta_2) = 1-a$, $a \in [0, 1]$; $m_2(\theta_1) = 0.5$, $m_2(\theta_2) = 0.5$

the result of Example 4.2 of Deng cross entropy is:$D_d = a \times \log \frac{a}{0.5} + (1-a) \times \log \frac{1-a}{0.5}$.

The Deng relative entropy between m_1 and m_2 is shown in Figure 1 with the parameter a has changed. It is obvious that when the value of a changes in the interval [0, 0.5], the divergence between m_1 and m_2 getting smaller and smaller, and their Deng relative entropy decreases correspondingly. Then, the value of Deng relative entropy turns into 0 when m_1 and m_2 are exactly the same. Moreover, equally obvious is that when the parameter a changes in the interval [0.5, 1], the divergence between m_1 and m_2 is growing, and their Deng relative entropy increases correspondingly.

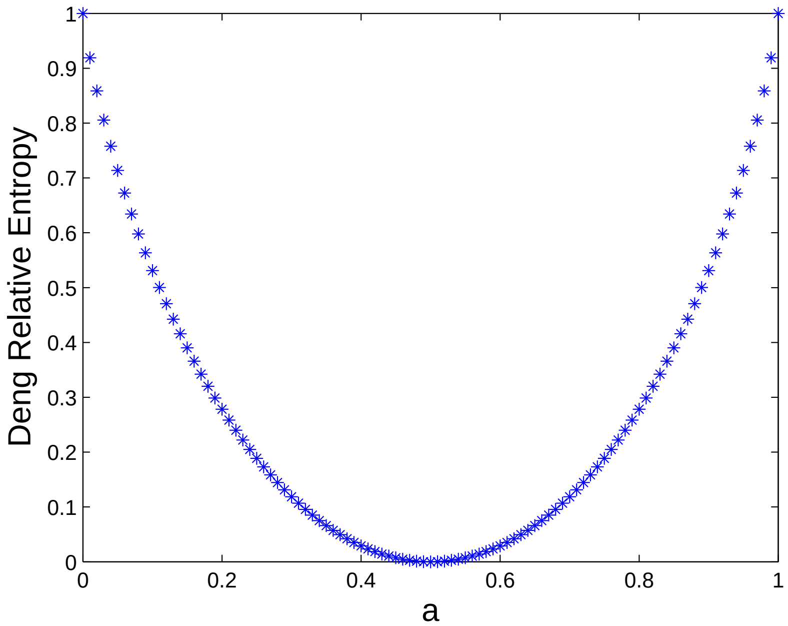

Figure 1: Deng relative entropy with changing parameter a

5. Conclusion

Dempster Shafer evidence theory is very important in the field of information fusion and applied widely in many processes because of its powerful features to handle the uncertainty. However, the counter-intuitive results are often obtained if there exist the larger conflict between different evidences. To resolve this serious problem effectively, in this paper, a new relative entropy is proposed named Deng relative entropy which is the generalization of Kullback-Leibler divergence, and the Deng relative entropy will degenerate to the K-L divergence when all the belief are assigned to single elements. Numerical examples are used to illustrate the efficiency of Deng relative entropy. The new relative entropy presents a method to measure the divergence between BPAs. There are some properties about Deng relative entropy are discussed, but there still exist some shortcomings have to be improved in the following works.

Acknowledgments

The work is partially supported by National High Technology Research and Development Program of China (863 Program) (Grant No. 2013AA013801), National Natural Science Foundation of China (Grant Nos. 61174022,61573290), China State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No.BUAA-VR-14KF-02).

References

[1] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematics and Statistics 38 (2) (1967) 325–339.

[2] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.

[3] V.-N. Huynh, T. T. Nguyen, C. A. Le, Adaptively entropy-based weighting classifiers in combination using dempster–shafer theory for word sense disambiguation, Computer Speech & Language 24 (3) (2010) 461–473.

[4] Y. Deng, Deng entropy: a generalized shannon entropy to measure uncertainty, viXra (2015) 1502.0222.
URL http://vixra.org/abs/1502.0222

[5] L. V. Utkin, A new ranking procedure by incomplete pairwise comparisons using preference subsets, Intelligent Data Analysis 13 (2) (2009) 229–241.

[6] Y. Deng, F. Chan, A new fuzzy dempster MCDM method and its application in supplier selection, Expert Systems with Applications 38 (8) (2011) 9854–9861.

[7] X. Su, S. Mahadevan, P. Xu, Y. Deng, Dependence assessment in Human Reliability Analysis using evidence theory and AHP, Risk Analysis 35 (2015) 1296–1316.

[8] Y. Deng, Generalized evidence theory, Applied Intelligence 43 (3) (2015) 530–543.

[9] Y. Deng, W. Jiang, R. Sadiq, Modeling contaminant intrusion in water distribution networks: A new similarity-based DST method, Expert Systems with Applications 38 (1) (2011) 571–578.

[10] Y. Deng, R. Sadiq, W. Jiang, S. Tesfamariam, Risk analysis in a linguistic environment: A fuzzy evidential reasoning-based approach, Expert Systems with Applications 38 (12) (2011) 15438–15446.

[11] R. R. Yager, N. Alajlan, Decision making with ordinal payoffs under dempster–shafer type uncertainty, International Journal of Intelligent Systems 28 (11) (2013) 1039–1053.

[12] J.-B. Yang, D.-L. Xu, Evidential reasoning rule for evidence combination, Artificial Intelligence 205 (2013) 1–29.

[13] Y. Deng, A threat assessment model under uncertain environment, Mathematical Problems in Engineering 2015 (2015) 878024,http://dx.doi.org/10.1155/2015/878024.

[14] F. Cuzzolin, A geometric approach to the theory of evidence, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 38 (4) (2008) 522–534.

[15] J. Schubert, Conflict management in Dempster-Shafer theory using the degree of falsity, International Journal of Approximate Reasoning 52 (3) (2011) 449–460.

[16] L. Zadeh, A simple view of the dempster-shafer theory of evidence and its.

[17] W. Liu, Analyzing the degree of conflict among belief functions, Artificial Intelligence 170 (11) (2006) 909–924.

[18] E. Lefevre, O. Colot, P. Vannoorenberghe, Belief function combination and conflict management, Information fusion 3 (2) (2002) 149–162.

[19] E. Lefèvre, Z. Elouedi, How to preserve the conflict as an alarm in the combination of belief functions?, Decision Support Systems 56 (2013) 326–333.

[20] D. Yong, S. WenKang, Z. ZhenFu, L. Qi, Combining belief functions based on distance of evidence, Decision support systems 38 (3) (2004) 489–493.

[21] F. Smarandache, J. Dezert, Applications and advances of dsmt for information fusion (2004).

[22] P. Smets, The combination of evidence in the transferable belief model, Pattern Analysis and Machine Intelligence, IEEE Transactions on 12 (5) (1990) 447–458.

[23] Y.-M. Xiong, Z.-P. Yang, X.-F. Qu, Novel combination method of conflict evidence based on evidential model modification, Control and Decision 26 (6) (2011) 883–887.

[24] R. R. Yager, On the dempster-shafer framework and new combination rules, Information sciences 41 (2) (1987) 93–137.

[25] L. Xinde, W. Fengyu, X. Yefan, C. Zongxiong, A method of evidence of combination based on the improved similarity measure, in: Information and Automation (ICIA), 2014 IEEE International Conference on, IEEE, 2014, pp. 461–466.

[26] C. K. Murphy, Combining belief functions when evidence conflicts, Decision support systems 29 (1) (2000) 1–9.

[27] K. Yamada, A new combination of evidence based on compromise, Fuzzy sets and Systems 159 (13) (2008) 1689–1708.

[28] D. Han, Y. Deng, C. Han, Sequential weighted combination for unreliable evidence based on evidence variance, Decision Support Systems 56 (2013) 387–393.

[29] A. Martin, A.-L. Jousselme, C. Osswald, Conflict measure for the discounting operation on belief functions, in: Information Fusion, 2008 11th International Conference on, IEEE, 2008, pp. 1–8.

[30] L.-F. Wang, J.-S. Song, Y. Zhu, J.-P. Cao, Evidence combination based on fuzzy clustering analysis, Systems Engineering and Electronics 35 (1) (2013) 113–119.

[31] Q. Sun, X.-q. Ye, W.-k. Gu, A new combination rules of evidence theory, Acta Electronica Sinica 28 (8) (2000) 117–119.

[32] S. Kullback, R. A. Leibler, On information and sufficiency, The annals of mathematical statistics (1951) 79–86.

[33] S. Kullback, Information theory and statistics, Courier Corporation, 1968.

[34] J. W. Gibbs, A method of geometrical representation of the thermodynamic properties of substances by means of surfaces, Connecticut Academy of Arts and Sciences, 1873.