Knowledge-Based Systems 54 (2013) 114–127

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Information-based dissimilarity assessment in Dempster–Shafer theory

Atiye Sarabi-Jamab [a], Babak N. Araabi [a,b,*], Thomas Augustin [c]

[a] Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran
[b] School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
[c] Department of Statistics, Ludwig Maximilians University, Munich, Germany

ARTICLE INFO

Article history:
Received 26 January 2013
Received in revised form 26 August 2013
Accepted 29 August 2013
Available online 12 September 2013

Keywords:
Dempster–Shafer theory of evidence
Dissimilarity measures
Uncertainty-based information
Conflict measures
Non-specificity measures

ABSTRACT

Dissimilarity assessment is a central problem in the Dempster–Shafer Theory (DST), where the difference in information content between two bodies of evidence (BoEs) should be quantified. Different dissimilarity measures (DMs) have been proposed; however, no single DM seems to be comprehensive enough to compare all aspects of information conveyed by BoEs. The information content of DMs are highly correlated as well. In this paper, DMs are categorized based on their interpretation of information content, emphasizing entropy-like DMs. A methodology is then proposed to select a set of more informative and less overlapping DMs called the "set of most discriminative dissimilarity measures" (smDDM). A forward selection procedure based on an appropriate criterion was utilized and the threshold for selection was derived naturally. To enhance the numerical evaluation, two experimental setups were designed and utilized with the existing setup to provide a sample of dissimilarity values. Comprehensive analysis supports the favorable properties of the proposed smDDM. The selected DMs came naturally from six different categories and subcategories of inner product-based and entropy-like DMs. Optimality analysis shows that the proposed selection procedure resulted in an appropriate near-optimal solution. Dissimilarity assessment is an integrated part of many applications of DST. The applicability and performance of the smDDM was examined and verified for two case studies: evidential clustering and sensor reliability evaluation.

1. Introduction

A main challenge in the Dempster–Shafer Theory (DST) is to quantify dissimilarities between two bodies of evidence (BoEs) [29,30,40]. If DST is assumed to be a natural extension of probability theory, such a dissimilarity assessment is like quantifying the distance between two probability distributions [6]; although, the uncertainty conveyed by a BoE is not restricted to randomness.

For a long time, Dempster's conflict factor has been the commonly used measure of dissimilarity between two BoEs [9]. Since then, several DMs were proposed to quantify different aspects of dissimilarity in DST, e.g. [20,14,54,27,42]. Indeed, More than 60 different dissimilarity measures (DMs) for BoEs have been utilized in different applications, including: belief function approximation [1,4,52], evidential clustering [11,44,55], uncertainty-based ranking [5,16,38,57], information fusion [7,28,41], evidential classification [17,21], and risk assessment [16,18,19,56]. A review and categorization of existing DMs in DST is provided in Section 3. A careful examination of these DMs reveals that none are comprehensive enough to compare all aspects of information conveyed by two BoEs. As a result, a set of DMs, instead of a single DM, promises to be a more appropriate tool in the comparison of BoEs.

A recent comprehensive survey by Jousselme and Maupin [32], classified 15 inner product-based DMs into 4 classes (metric, pseudo-metric, semi-pseudo-metric, non-metric) based on their formal properties. They highlighted basic limitations of the Dempster conflict [40] as a non-metric measure and proposed an alternative cosine function based on pseudo-metric measures. Using their analysis and results for applications like approximation algorithm evaluation, evidential risk assessment, and evidential clustering, the following difficulties develop:

- the DMs in each class are highly correlated and using more than one of them often causes redundant information content,
- the entropy-like DMs, e.g. [23,37,39] are not studied, even though they may be important to some applications,
- the results are sensitive to the size of the frame of discernment (FoD).

In this paper, we propose a framework for comprehensive assessment of dissimilarity between two BoEs. A set of DMs is selected to better quantify the difference in the amount of information

* Corresponding author. Address: University of Tehran, School of Electrical and Computer Engineering, Fanni Campus II, N. Kargar Ave., Tehran 14399, Iran. Tel.: +98 2188630024; fax: +98 2188633029.

E-mail addresses: asarabi@ut.ac.ir (A. Sarabi-Jamab), araabi@ut.ac.ir (B.N. Araabi), Thomas@stat.uni-muenchen.de (T. Augustin).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.08.030

conveyed by separate sources. We put emphasis on information content of BoEs; as a result, entropy-like DMs that address difference in conflict and non-specificity gain importance. A methodology to select a set of DMs with maximal discrimination power is proposed. In the proposed forward selection procedure, a max–min strategy on an appropriate criterion resulted in a set of DMs that is maximally uncorrelated while the standard deviation (a measure of discrimination) is kept as high as possible. The final outcome is the set of most discriminative dissimilarity measures (smDDM) that represent the minimal set of DMs needed for a thorough evaluation of differences between two BoEs, and consists of three inner product-based DMs along with five entropy-like DMs.

The experimental setup introduced by Jousselme and Maupin [32] for classification of DMs, considered only complete BoEs with different mass functions; however, in many applications, two BoEs with different numbers of focal elements (FEs) must be compared. To enhance the numerical evaluation, two new setups are designed and utilized along with the setup introduced by Jousselme et al.

The rest of the paper is organized as follows: Section 2 provides a short overview of basic concepts and notations in DST. Existing DMs in DST are reviewed and categorized based on their treatment of information content in Section 3. A methodology is then proposed for selecting the smDDM from among all inner product-based and entropy-like DMs in Section 4. The smDDM is selected, analyzed, and discussed comprehensively in Sections 5 and 6. The applicability and performance of smDDM is studied and compared with the other sets of DMs through two case studies in Section 7. The paper is concluded in Section 8.

2. Dempster Shafer theory

DST is a generalization of Bayesian theory on subjective probability. It is based on obtaining degrees of belief for one question from subjective probabilities for a related question where pieces of evidences appear as subsets of the FoD instead of elements of the FoD. This model conveys a higher level of uncertainty for modeling both ignorance and non-specificity [9,49].

Consider a discrete FoD with N elements, $\Theta = \{\omega_1, \ldots, \omega_N\}$. A mass function $m(\cdot)$ is defined as

$$m: P_\Theta \to [0,1], \quad \sum_{A \in P_\Theta} m(A) = 1, \quad m(A) \geqslant 0, \quad \forall A \in P_\Theta \tag{1}$$

where P_Θ is the power set of Θ. $m(\cdot)$ is also called a basic belief assignment. Those subsets of Θ (A_j's) with nonzero mass values ($m_j = m(A_j) > 0$) are called FEs. A BoE consists of all FEs and their mass values:

$$\begin{aligned} &\{A_1, A_2, \ldots, A_n\}, \quad \{m_1, m_2, \ldots, m_n\} \\ &\phi \neq A_j \subset \Theta, \quad m_j > 0, \quad \sum m_j = 1 \end{aligned} \tag{2}$$

where n is the number of FEs. This paper is confined to discrete FoDs.

In evidence theory [49], Shafer defines belief, plausibility, and commonality functions over a mass function as Eq. :

$$Bel: P_\Theta \to [0,1], \quad Bel(B) = \sum_{A \subseteq B, A \neq \phi} m(A) \quad \forall B \subseteq \Theta \tag{3}$$

$$Pl: P_\Theta \to [0,1], \quad Pl(B) = \sum_{A \subseteq \Theta, A \cap B \neq \phi} m(A) \quad \forall B \subseteq \Theta \tag{4}$$

$$Q: P_\Theta \to [0,1], \quad Q(B) = \sum_{A \subseteq \Theta, B \subseteq A} m(A) \quad \forall B \subseteq \Theta \tag{5}$$

The combination of two independent sources of evidence with their corresponding mass functions $m_1(\cdot)$ and $m_2(\cdot)$ were obtained using Dempster's rule of combination and denoted by the mass function $m_1 \oplus m_2(\cdot)$ [9]:

$$m_1 \oplus m_2(A) = (1/1-k) \sum_{B,C \subseteq \Theta, B \cap C = A} m_1(B) m_2(C) \quad \forall A \subseteq \Theta, \quad A \neq \phi \tag{6}$$

where

$$K = \sum_{B,C \subseteq \Theta, B \cap C = \phi} m_1(B) m_2(C) \tag{7}$$

3. Categorizing dissimilarity measures in DST

DMs in DST have been widely studied in different application areas like: Evidential classification [17], belief function approximation [14,15,52], and evidential clustering [48,55], to name but a few. In DST, two different types of indeterminacy are distinguishable; one for cases where the information focuses on sets with intersections, one for cases where the information focuses on sets where the cardinality is greater than one. These are called conflict and non-specificity, respectively [37]. This section provides a review and categorization of existing DMs based on the way they compare information content between two BoEs; they measure any of the conflict, non-specificity, or unified total uncertainty.

A distance on a set M is a function d: $M \times M \to R$ that satisfies the following conditions [12]:

(d1) non-negativity: $d(y,z) \geqslant 0$,
(d2) symmetry: $d(y, z) = d(z, y)$,
(d3) definiteness: $d(y,z) = 0 \iff y = z$,
(d4) triangle inequality/subadditivity: $d(y,z) \leqslant d(y,t) + d(z,t)$ for all t.

A distance function with these properties is known as a metric. Property (d3) can be split into (d3)′ and (d3)″:

(d3)′ reflexity: $d(y, y) = 0$,
(d3)″ separability: $d(y,z) = 0 \Rightarrow y = z$.

The idea of dissimilarity between two BoEs can be quantified by estimating the difference in their information content. A general formulation to quantify the DM for approximation algorithm evaluation was introduced in [10]:

$$d_U(m_1, m_2) = |U(m_1) - U(m_2)| \tag{8}$$

where U can be any uncertainty measure defined on a mass or belief function. This measure basically estimates the difference between two sources of information. In the sequel, DMs in Tables 1, 3 and 4 are introduced based on Eq. (8), by using various definitions of uncertainty. An uncertainty-based interpretation of information [34] is considered.

3.1. Measuring the degree of conflict between two BoEs

K in Eq. (7) is a commonly-used measure of conflict between two BoEs [9]. Several DMs have been proposed to measure the amount of conflict between two pieces of evidence [22–24,26,46]. These DMs can be categorized into two general types. In Type I, the DMs quantify a difference in information content; in Type II, the DMs quantify the difference in decision abilities.

3.1.1. Type I

Type I DMs utilize Eq. (8) to measure the degree of conflict through the difference in information content of two BoEs. Entropy-like measures are used to quantify the amount of uncertainty. For all of these DMs, $U(m)$ in Eq. (8) have a summation form $U(m) = \sum \varphi(m(A))$ for all $A \subseteq \Theta$, where the function φ usually has a logarithmic form similar to Shannon's entropy [34], except