Knowledge-Based Systems xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A new method to determine basic probability assignment using core samples

Chenwei Zhang [a], Yong Hu [b], Felix T.S. Chan [c], Rehan Sadiq [d], Yong Deng [a,*]

[a] School of Computer and Information Science, Southwest University, Chongqing 400715, China
[b] Institute of Business Intelligence & Knowledge Discovery, Department of E-commerce, Guangdong University of Foreign Studies, Sun Yat-Sen University, Guangzhou 510006, China
[c] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
[d] School of Engineering, University of British Columbia Okanagan, 3333 University Way, Kelowna, BC, Canada

ARTICLE INFO

Article history:
Received 13 September 2013
Received in revised form 2 March 2014
Accepted 11 June 2014
Available online xxxx

Keywords:
Data fusion
Dempster–Shafer theory of evidence
Basic probability assignment
Core sample
Convex hull

ABSTRACT

The Dempster–Shafer theory of evidence (D–S theory) has been widely used in many information fusion systems. However, the determination of basic probability assignment (BPA) remains an open problem which can considerably influence final results. In this paper, a new method to determine BPA using core samples is proposed. Unlike most of existing methods that determining BPA in a heuristic way, the proposed method is data-driven. It uses training data to generate core samples for each attribute model. Then, helpful core samples in generating BPAs are selected. Calculation of the relevance ratio based on convex hulls is integrated into the core sample selection as a new feature of the proposed method. BPAs are assigned based on the distance between the test data and the selected core samples. Finally, BPAs are combined to get a final BPA using the Dempster's combination rule. In this paper, compound hypotheses are taken into consideration. BPA generated by the proposed method can be combined with some other sources of information to reduce the uncertainty. Empirical trials on benchmark database shows the efficiency of the proposed method.

1. Introduction

Data fusion is an evolving technology used by human to integrate data from sensors continually and make inferences about the real world [1]. The information provided by single sensor may be lack of accuracy or limited. Thus, the use of multiple sensor is aimed to improve the accuracy and can provide the user with increased information about the environment [2]. Applications of data fusion ranges from environment analysis [3,4], clinical diagnosis [5,6], transportation management [7,8] and so on. In order to integrate information from multisensor efficiently, different strategies have been developed for data fusion. Classical Bayesian theory and Dempster–Shafer theory of evidence (D–S theory) [9,10] are two mainstream frameworks in data fusion. When comparing with the classical Bayesian theory, the merit of D–S theory is outweigh. As a useful tool to handle uncertainty, D–S theory has been a research hotspot in the field of data fusion [11–20].

Dempster–Shafer theory of evidence, introduced by Dempster [9] and then developed by Shafer [10] since the early 1980's, has now developed into a mature stage. But at the same time, it also has some problems. Although Dempster–Shafer theory of evidence is widely used and has good performance in practical applications [21–24], some basic problems are still not clarified. Within the framework of D–S theory, the construction of basic probability assignment (BPA) remains an important problem which can considerably influence final results. However, the determination of BPA remains an open issue. Till now, there is no general method to determine BPA. Many authors have addressed this problem through different approaches [25–28], but most of the existing approaches determine BPA heuristically. Yager [26] associated the D–S belief structure with a whole class of fuzzy measures, and discussed the entropy of a fuzzy measure. Xu [28] put forward a method to obtain BPA based on the normal distribution of data in each attribute. Denoeux [25] proposed a neural network classifier based on D–S theory. In Denoeux's work [25], the determination of BPA is implemented in a multilayer neural network with the architecture of one input layer and two hidden layers. The weight vector in the neural network is used to determine BPA based on the distance of pattern to its *k*-nearest neighbor (*k*-NN) prototypes.

* Corresponding author. Tel./fax: +86 023 68254555.
E-mail addresses: ydeng@swu.edu.cn, professordeng@hotmail.com (Y. Deng).

http://dx.doi.org/10.1016/j.knosys.2014.06.015
0950-7051/© 2014 Elsevier B.V. All rights reserved.

Prototype is considered as a representative pattern for each class. This method shows excellent performance as compared to existing statistical and neural network techniques. However, prototype in Denoeux's work is generated on every single class. Although each prototype is assumed to possess a degree of membership to each class, Denoeux's methods do not consider the BPA on compound hypotheses and do not exploit all the strengths of the theory. Compound hypotheses are useful and important to represent the uncertainty and imprecise situation and hypothesis on compound set is reallocated proportionally among singleton during the combination process.

Inspired by Denoeux's work [29,30,25], a new method to determine BPA is proposed in this paper. The main contribution of this paper is the development of a new way to determine the BPA, in which compound hypotheses are taken into consideration. BPA generated by the proposed method can be combined with some other sources of information to reduce the uncertainty. In this method, the basic probability of a pattern to a class is assigned by calculating distance with core samples representing each classification. Since not every core sample is helpful to generate a BPA, whether and which core samples should be selected is quantified by relevance ratio based on convex hulls. For each attribute in the pattern, BPAs are generated and then combined using Dempster's combination rule to get a final BPA.

The present work extends the use of the Denoeux's model where BPA can be assigned not only on single hypothesis but also on compound hypotheses. When evidence is not adequate, unlike classical probability theory in which possibilities should be determined forcedly, D–S theory provides a flexible way to determine BPA on both single hypotheses and compound hypotheses. Therefore, when existing information is insufficient to make a convincing conclusion solely on single hypotheses, is it reasonable to introduce compound hypotheses to modeling the uncertainty. Also, we develop some new features into the proposed method that uncertainty of subjectivity is reduced by core sample selection.

The rest of this paper is organized as follows. Section 2 starts with some concepts on D–S theory and some necessary related concepts. The proposed method determining basic probability assignment using core samples and its detailed procedures are presented Section 3. In Section 4, experiments on pattern classification task are conducted using the proposed method. Conclusion is presented in Section 5.

2. Preliminaries

2.1. Dempster–Shafer theory of evidence

In this section, the main concepts underlying the Dempster–Shafer theory of evidence are recalled. The Dempster–Shafer theory of evidence, as introduced by Dempster [9] and then developed by Shafer [10], has emerged from their works on statistical inference and uncertain reasoning. Compared with the Bayesian probability model, the merits of D–S theory have already been recognized in various fields. First, the D–S theory can handle more uncertainty in real world. In contrast to the Bayesian probability model in which probability masses can be only assigned to singleton subsets, in D–S theory probability masses can be assigned to both singletons and compound sets. Thus, more evidence will be provided to illustrate the hypotheses or the distribution between the singletons. Second, in D–S theory, no prior distribution is needed before the combination of evidence from individual information sources. Third, the D–S theory allows one to specify a degree of ignorance in some situations instead of being forced to be assigned for probabilities. Some notations in D–S theory are introduced.

2.1.1. Frame of discernment

In D–S theory, Let

$$\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\} \tag{1}$$

be the finite set of mutually exclusive and exhaustive events. The set of every subset of Θ, which has the cardinality of $2^{|\Theta|}$, is called the frame of discernment, denotes as

$$\Omega = \{\emptyset, \{\theta_1\}, \{\theta_2\}, \ldots, \{\theta_n\}, \{\theta_1, \theta_2\}, \ldots, \{\theta_1, \theta_2, \ldots, \theta_n\}\} \tag{2}$$

2.1.2. Basic probability assignment (BPA)

When some pieces of evidence assigns probability masses to the subsets of Ω, the resulting function is called a basic probability assignment or a mass function. Mathematically, a basic probability assignment is a function m mapping from the power set of Θ to $[0, 1]$, satisfying

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

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{4}$$

where $\emptyset$ is an empty set and A is any subsets of Θ and the mass function $m(A)$ represents how strongly the evidence supports A.

2.1.3. Dempster's combination rule

The BPA obtained by two individual information sources are combined according to Dempster's combination rule, defined as

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

with

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{6}$$

where A, B and C are subsets of $2^{|\Theta|}$, and K is a normalization constant, called the conflict coefficient of two BPAs.

2.1.4. Pignistic probability

After combination, mass function can be transformed into pignistic probability [31] for decision making. Pignistic probability for A is denoted as:

$$P_{pig}(A) = \sum_{B \subseteq \theta} \frac{card(A \cap B) \times m(B)}{card(B) \times (1 - m(\emptyset))} \tag{7}$$

Since $m(\emptyset) = 0$, this equation can be simplified as:

$$P_{pig}(A) = \sum_{B \subseteq \theta} \frac{card(A \cap B) \times m(B)}{card(B)} \tag{8}$$

where $card(X)$ stands for the cardinality of set X.

2.2. Classical convexity

In the field of computational geometry, convexity [32] is a widely studied problem. The determination of the convex hull [33] is useful in many analysis methods and has successfully been applied to many fields, such as pattern recognition and image processing. If not special specified, the following concepts are based on 2-dimensional space. However, convexity properties also can be seen in one dimension and higher dimensions. In this paper, convex hull is analogously used to describe the shape of a set of data points.

Please cite this article in press as: C. Zhang et al., A new method to determine basic probability assignment using core samples, Knowl. Based Syst. (2014), http://dx.doi.org/10.1016/j.knosys.2014.06.015