Information Sciences xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Information Sciences

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

An improved method to construct basic probability assignment based on the confusion matrix for classification problem

Q1 Xinyang Deng[a,b], Qi Liu[b,c], Yong Deng[a,d,*], Sankaran Mahadevan[d]

[a] School of Computer and Information Science, Southwest University, Chongqing 400715, China
[b] Center for Quantitative Sciences, Vanderbilt University School of Medicine, Nashville, TN 37232, USA
[c] Department of Biomedical Informatics, Vanderbilt University School of Medicine,Nashville, TN 37232, USA
[d] School of Engineering, Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Article history:
Received 26 May 2015
Revised 15 September 2015
Accepted 11 January 2016
Available online xxx

Keywords:
Dempster–Shafer evidence theory
Belief function
Basic probability assignment
Confusion matrix
Classification
Transmembrane proteins

ABSTRACT

The determination of basic probability assignment (BPA) is a crucial issue in the application of Dempster–Shafer evidence theory. Classification is a process of determining the class label that a sample belongs to. In classification problem, the construction of BPA based on the confusion matrix has been studied. However, the existing methods do not make full use of the available information provided by the confusion matrix. In this paper, an improved method to construct the BPA is proposed based on the confusion matrix. The proposed method takes into account both the precision rate and the recall rate of each class. An illustrative case regarding the prediction of transmembrane protein topology is given to demonstrate the effectiveness of the proposed method.

1. Introduction

Dempster–Shafer evidence theory [9,45], also called Dempster–Shafer theory, has been widely applied in many fields, such as information fusion [50], classification [5,37,38], and others [4,10–13,16,20,25,42,49,57,68]. Tabassian et al. [51,52] used Dempster–Shafer theory to handle data with imperfect labels in ensemble learning, and addressed the situation that the class memberships of the training data are subject to ambiguity. Deng [19] proposed a generalized evidence theory (GET) to address conflict management in an open world environment. Thanks to its flexibility, Dempster–Shafer theory has been combined with other theories like fuzzy set theory [8,31,69] and genetic algorithm [22], and many useful tools have been developed to handle various types of uncertainty, which further extends the application of the Dempster–Shafer theory. For instance, in [32], Kang et al. proposed an uncertain-graph structure, called evidential cognitive map (ECM), to represent causal reasoning by combining the cognitive maps and Dempster–Shafer theory. Recently, in the fields of evolutionary game theory [6,7,15,58–64] and game theory [53], Dempster–Shafer evidence theory has also attracted some interests [17,18,35].

The determination of basic probability assignment (BPA) is one of the most important problems in evidential systems. The construction of BPA based on the confusion matrix is a practical and effective method [1,2,24,40,43,54,65]. In a previous related study, Xu et al. [65] presented an elegant method for the construction of BPA based on recognition rate, substitution

Q2 * Corresponding author at: School of Computer and Information Science, Southwest University, Chongqing 400715, China. Tel.: +86 23 68254555.
E-mail address: prof.deng@hotmail.com (Y. Deng).

http://dx.doi.org/10.1016/j.ins.2016.01.033
0020-0255/

rate, and rejection rate of the confusion matrix. However, Xu et al.'s method does not consider the difference of the classifier's recognition ability for different classes. To overcome the shortcoming, Parikh et al. [40] proposed a modified method, which is more effective and has been successfully used in condition monitoring. The improvement proposed by Parhikh et al. is on the basis of the prior knowledge provided by the confusion matrix. Specifically, that method utilized the precision rate of each actual class according to the confusion matrix. However, the prior knowledge contained in the confusion matrix is not only the precision rate, but also the recall rate of each class which is another important aspect to reflect the classifier's recognition ability for each class.

Based on this idea, an improved BPA construction method is proposed based on the confusion matrix in this paper. Section 2 introduces some basic concepts and related previous work. Section 3 presents the proposed method. Section 4 gives an illustrative case to demonstrate the effectiveness of the proposed method. Section 5 concludes the paper.

2. Preliminaries

2.1. Basic concepts

The Dempster–Shafer evidence theory [9,45], first proposed by Dempster and further developed by Shafer, is widely used to handle uncertain information. In this theory, basic probability assignment (BPA) is used to represent the uncertain information, and Dempster's rule of combination is used to combine multiple BPAs.

In Dempster–Shafer theory, a problem domain denoted by a finite nonempty set Ω of mutually exclusive and exhaustive hypotheses is called the frame of discernment. Let 2^{Ω} denote the power set of Ω. A BPA is a mapping $m: 2^{\Omega} \to [0, 1]$, satisfying

$$m(\emptyset) = 0 \quad and \quad \sum_{A \in 2^{\Omega}} m(A) = 1 \tag{1}$$

Dempster's rule of combination, also called orthogonal sum, is defined as follows

$$m(A) = \begin{cases} \dfrac{1}{1-K} \displaystyle\sum_{B \cap C = A} m_1(B) m_2(C), & A \neq \emptyset; \\ 0, & A = \emptyset. \end{cases} \tag{2}$$

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{3}$$

where K is a normalization constant, called conflict coefficient of two BPAs. The Dempster's rule satisfies commutative and associative properties. Thus if there exist multiple BPAs, the combination of them can be carried out in a pairwise way with any order. However, this rule may give counterintuitive results when there is conflict among BPAs. For this issue, many other combination rules have been proposed, for example Yager's conflict redistribution rule [66], Dubois and Prade's disjunctive rule [26], Murphy's simple averaging rule [39], Deng's weighted averaging rule [23], proportional conflict redistribution PCR5 and PCR6 rules [29,46,47].

2.2. Confusion matrix

Confusion matrix [44] is a concept from machine learning, which contains information about actual and predicted classifications done by a classification system. A confusion matrix has two-dimensions, one dimension is indexed by the actual class of an object, the other is indexed by the class that the classifier predicts. Fig. 1 presents the basic form of confusion matrix for a multi-class classification task, with the classes A_1, A_2, and A_n. In the confusion matrix, N_{ij} represents the number of samples actually belonging to class A_i but classified as class A_j.

		Predicted			
		A_1	⋯ A_j ⋯		A_n
Actual	A_1	N_{11}	N_{1j}		N_{1n}
	⋮	⋮	⋮		⋮
	A_i	N_{i1}	⋯ N_{ij} ⋯		N_{in}
	⋮	⋮	⋮		⋮
	A_n	N_{n1}	N_{nj}		N_{nn}

Fig. 1. Confusion matrix.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

A number of measures of classification performance can be defined based on the confusion matrix. Some common measures are given as follows.

Accuracy is the proportion of the total number of predictions that were correct:

$$Accuracy = \sum_{i=1}^{n} N_{ii} \Big/ \sum_{i=1}^{n}\sum_{j=1}^{n} N_{ij} \tag{4}$$

Precision is a measure of the accuracy provided that a specific class has been predicted. It is defined by:

$$Precision_i = N_{ii} \Big/ \sum_{k=1}^{n} N_{ki} \tag{5}$$

Recall is a measure of the ability of a prediction model to select instances of a certain class from a data set, it is defined by the formula:

$$Recall_i = N_{ii} \Big/ \sum_{k=1}^{n} N_{ik} \tag{6}$$

The traditional F-score (F_1 score) is the harmonic mean of precision and recall:

$$F-score_i = \frac{2 \times Precision_i \times Recall_i}{Precision_i + Recall_i} \tag{7}$$

2.3. Related previous work

In supervised machine learning, let $\Omega = \{c_1, c_2, \ldots, c_N\}$ be a set of N classes, and $D = \{d_1, d_2, \ldots, d_M\}$ be a set of M samples. A sample whose class label is known can be denoted as a 2-tuple $< d_i,\ c_{[d_i]} >$, where $d_i \in D$ and $c_{[d_i]} \in \Omega$. Given a sample, classification is a process of determining the class label that the sample belongs to. Suppose there exists a classifier φ. A group of samples are classified by φ in an uncertain environment. The classification result C_φ can be shown in the form of a confusion matrix as follows.

$$C_\varphi = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1N} & n_{1(N+1)} \\ n_{21} & n_{22} & \cdots & n_{2N} & n_{2(N+1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n_{N1} & n_{N2} & \cdots & n_{NN} & n_{N(N+1)} \end{bmatrix} \tag{8}$$

Here, n_{ij} is the number of samples belonging to class c_i but classified as class c_j. The right-most column represents the samples of rejection class that refuse to be classified (i.e., cannot be classified as any actual classes).

According to Xu et al.'s method [65], a BPA is constructed for each actual class c_i using C_φ as follows.

$$m^\varphi(\{c_i\}) = \epsilon_c^\varphi, \quad \forall c_i \in \Omega \tag{9}$$

$$m^\varphi(\overline{\{c_i\}}) = \epsilon_e^\varphi, \quad \forall c_i \in \Omega,\ \overline{\{c_i\}} = \Omega \setminus \{c_i\} \tag{10}$$

$$m^\varphi(\Omega) = \epsilon_r^\varphi, \quad \forall c_i \in \Omega \tag{11}$$

where $\epsilon_c^\varphi = \sum_{i=1}^{N} n_{ii} / \sum_{i=1}^{N}\sum_{j=1}^{N+1} n_{ij}$, $\epsilon_r^\varphi = \sum_{i=1}^{N} n_{i(N+1)} / \sum_{i=1}^{N}\sum_{j=1}^{N+1} n_{ij}$, and $\epsilon_e^\varphi = 1 - \epsilon_c^\varphi - \epsilon_r^\varphi$. ϵ_c^φ, ϵ_e^φ ϵ_r^φ are the recognition rate, substitution rate, and rejection rate of classier φ, respectively.

In Xu et al.'s method, a same BPA is shared by each class, so that the difference of the classifier's recognition ability to different classes is not considered. To overcome this shortcoming, Parikh et al. [40] proposed a modified method to improve Xu et al.'s method. Based on C_φ, the constructed BPA using Parikh et al.'s method is

$$m^\varphi(\{c_i\}) = \varepsilon_{p[i]}^\varphi, \quad c_i \in \Omega \tag{12}$$

$$m^\varphi(\overline{\{c_i\}}) = 1 - \varepsilon_{p[i]}^\varphi, \quad c_i \in \Omega \tag{13}$$

$$m^\varphi(\Omega) = 0 \tag{14}$$

where $\varepsilon_{p[i]}^\varphi = n_{ii} / \sum_{j=1}^{N} n_{ji}$, $\varepsilon_{p[i]}^\varphi$ is the recognition rate of an actual class c_i. When a sample has been classified as class c_{N+1}(rejection class), it means the sample cannot be classified as any actual classes and hence the BPA is $m^\varphi(\Omega) = 1$.

The construction of BPA based on the confusion matrix is a very practical method, it has attracted many researchers' attention. The methods proposed by Xu et al. [65] and Parikh et al. [40] are two typical methods and have been successfully applied to condition monitoring and fault diagnosis (CMFD) [27,41]. Other investigations have been reported in [1,2,24,43,54]. For example, Denoeux [24] addressed the building of k-nearest neighbor classifier based on the Dempster–Shafer theory and derived a BPA construction method based on the normalized confusion matrix.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

The output information that a classification system supplies can be divided into three levels [65], namely abstract level, rank level, and measurement level, respectively. The confusion matrix provides available information in the abstract level. Summarizing previous studies, the focus is on the exploitation of available knowledge contained in the confusion matrix. Therefore, the utilization of available information provided by the confusion matrix is a key issue in supervised machine learning.

3. Proposed method

Based on the abundant prior knowledge contained in the confusion matrix, an improved method to construct the BPA is proposed in this section. Generally, a confusion matrix contains the information about actual and predicted classifications given by a classification system. In order to evaluate the performance of such systems, the data in the matrix is usually used. Based on the data coming from the confusion matrix, some indices, for instances accuracy, sensitivity (also called recall), specificity, precision, have been developed to evaluate the performance of a classification system. In the task of multi-class classification, there inevitably exists differences among the classification abilities of a classification system to different classes. Such differences are hard to be reflected by any single performance index. However, by analyzing these performance indices, it is found that the combination of recall and precision contains all the available information included in a confusion matrix. Based on the recall matrix and precision matrix, the original confusion matrix can be reconstructed completely. The combination of recall and precision provides the ability to reflect the differences of classifier's classification abilities to different classes. Therefore, the proposed method utilizes the recall matrix and precision matrix to construct the BPA for each class. These two performance indices, recall and precision, are treated as two different sources of evidence. By synthesizing the evidences generated from recall and precision, the final BPA makes full use of the information contained in the confusion matrix. specifically, the proposed BPA construction method is given as below.

Assume the number of samples belonging to class c_i but classified as class c_j is indicated by n_{ij}. First, let r^r_{ij} be the ratio of n_{ij} to the total number of samples belonging to class c_i, namely $r^r_{ij} = \frac{n_{ij}}{\sum_{j=1}^{N+1} n_{ij}}$; And let r^p_{ij} be the ratio of n_{ij} to the total number of samples classified as class c_j, namely $r^p_{ij} = \frac{n_{ij}}{\sum_{i=1}^{N} n_{ij}}$. Using r^r_{ij}, r^p_{ij} and C_φ, two matrices, namely the recall matrix indicted by C^r_φ and the precision matrix indicated by C^p_φ, are obtained as

$$C^r_\varphi = \begin{bmatrix} r^r_{11} & r^r_{12} & \cdots & r^r_{1N} & r^r_{1(N+1)} \\ r^r_{21} & r^r_{22} & \cdots & r^r_{2N} & r^r_{2(N+1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r^r_{N1} & r^r_{N2} & \cdots & r^r_{NN} & r^r_{N(N+1)} \end{bmatrix} \quad and \quad C^p_\varphi = \begin{bmatrix} r^p_{11} & r^p_{12} & \cdots & r^p_{1N} & r^p_{1(N+1)} \\ r^p_{21} & r^p_{22} & \cdots & r^p_{2N} & r^p_{2(N+1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r^p_{N1} & r^p_{N2} & \cdots & r^p_{NN} & r^p_{N(N+1)} \end{bmatrix}$$

Next, using C^r_φ, for each actual class c_i, a BPA m^r_i is derived as

$$m^r_i(\{c_i\}) = r^r_{ii} \Big/ \sum_{j=1}^{N} r^r_{ji}, \quad c_i \in \Omega \tag{15}$$

$$m^r_i(\Omega) = 1 - m^r_i(\{c_i\}), \quad c_i \in \Omega \tag{16}$$

$$m^r_i(A) = 0, \quad \forall A \in 2^\Omega \setminus \{\{c_i\}, \Omega\} \tag{17}$$

Using C^p_φ, for each actual class c_i, a BPA m^p_i is derived as

$$m^p_i(\{c_i\}) = r^p_{ii} \Big/ \sum_{j=1}^{N+1} r^p_{ij}, \quad c_i \in \Omega \tag{18}$$

$$m^p_i(\Omega) = 1 - m^p_i(\{c_i\}), \quad c_i \in \Omega \tag{19}$$

$$m^p_i(A) = 0, \quad \forall A \in 2^\Omega \setminus \{\{c_i\}, \Omega\} \tag{20}$$

By synthesizing the two BPA m^r_i and m^p_i, the final BPA m^φ_i is derived as

$$m^\varphi_i = m^p_i \oplus m^r_i \tag{21}$$

where m^φ_i is the constructed BPA for each class c_i based on the classifier φ. If a target has been classified as rejection class, the BPA is $m^\varphi(\Omega) = 1$.

The proposed BPA construction method can be used to the combination of multiple classifiers, as shown in Fig. 2. Assume there are l base classifiers, indicated by $\varphi_1, \ldots, \varphi_l$. The samples are composed by two parts, namely training data set $D_{training}$ and test data set D_{test}. For the evidential multiple classifiers system, it consists of five main steps to classify the unclassified targets in D_{test}.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

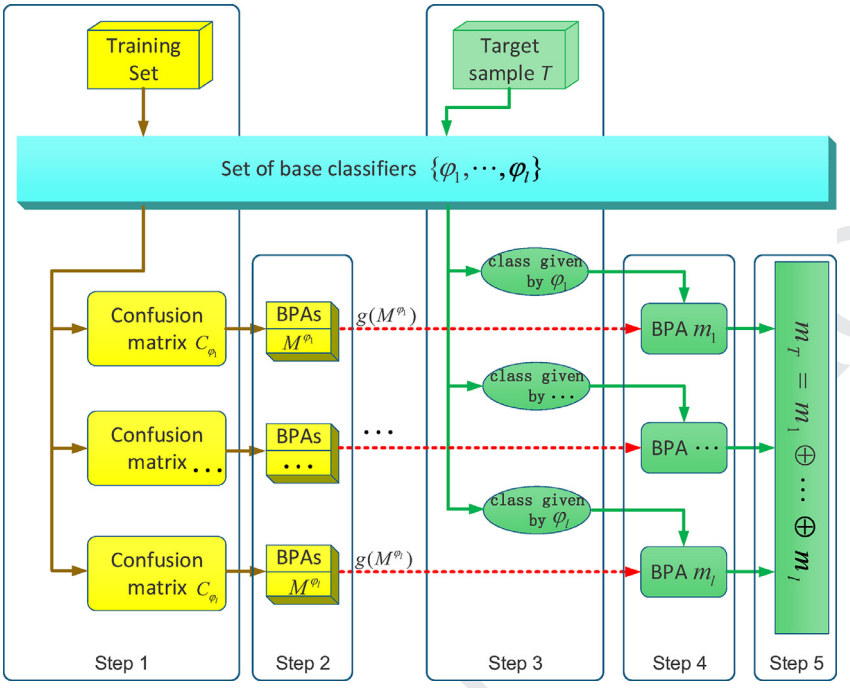

Fig. 2. Diagrammatic sketch for classification problem by using Dempster–Shafer theory.

At step 1, for each sample s, whose real class is c_s^{real}, in $D_{training}$, every base classifier among $\{\varphi_1,\ldots,\varphi_l\}$, indicated by φ_k, can give a predicted class c_s^{pred}. By comparing c_s^{real} and c_s^{pred}, we can repeatedly update the confusion matrix C_{φ_k} until every sample in $D_{training}$ has been taken into consideration. As a result, C_{φ_1}, ..., C_{φ_l} are obtained.

At step 2, based on the obtained confusion matrix C_{φ_k}, $k=1,2,\ldots,l$, the recall matrix $C_{\varphi_k^r}$ and precision matrix $C_{\varphi_k^p}$ are calculated first. Then, for each actual class i in Ω, its recall BPA m_i^r and precision BPA m_i^p can be constructed by means of Eqs. (15)–(20). At last, combine m_i^r and m_i^p to obtain $m_i^{\varphi_k}$ which represents the recognition ability of φ_k to class i. For convenience, let $m_i^{\varphi_k}$ be kept in M^{φ_k} in the form of key-value pairs: $M^{\varphi_k}[\text{“}i\text{”}]=m_i^{\varphi_k}$. Due to there are l base classifiers, it will generate M^{φ_1}, ..., M^{φ_l}.

Steps 1 and 2 are the training phase. After these two steps have been finished, real classification process for unclassified samples in D_{test} can be executed. At step 3, for each unclassified target T in D_{test}, each base classifier gives a prediction: $c_T^{\varphi_k}=\varphi_k(T),\ k=1,\ldots,l$.

At step 4, in terms of the predicted classes given by different base classifiers, a combination result can be obtained. Given $c_T^{\varphi_1}$, $c_T^{\varphi_2}$, ..., $c_T^{\varphi_l}$, the predicted result of multiple classifiers is given by:

$$m_T = g\left(M^{\varphi_1}, c_T^{\varphi_1}\right) \oplus g\left(M^{\varphi_2}, c_T^{\varphi_2}\right) \oplus \cdots \oplus g\left(M^{\varphi_l}, c_T^{\varphi_l}\right) \tag{22}$$

where $g(M^{\varphi_k}, c_T^{\varphi_k}) = M^{\varphi_k}[\text{“}c_T^{\varphi_k}\text{”}]$, $k=1,2,\ldots,l$, and m_T is the final BPA to determine which class target T belongs to. It's worth noting that the evidence combination rule “$\oplus$” plays an important role in the evidential multiple classifiers system. Different rules may lead to significantly different results. The impact of combination rules on the results will be discussed later.

At step 5, the predicted class for T can be determined based on the pignistic transformation [48]:

$$c_T^{pred} = \arg\max_x \left\{ \sum_{A\subseteq\Omega, x\in A} \frac{1}{|A|} \frac{m_T(A)}{1-m_T(\emptyset)} \right\}, x\in\Omega. \tag{23}$$

The above steps are mainly concluded in Algorithm 1. The evidential multi-classifier system has some important differences to other algorithms. First, in this system the idea of multi-information fusion has been imported to combine the results derived from various base classifiers. The evidential multiple classifiers system is independent from base classifiers so that diverse algorithms can be used to construct base classifiers. Second, different from other classification systems which are on the basis of probabilities, this system is based on BPAs which have inherent advantages to express and deal with stochastic and epistemic uncertainty. Third, a new BPA construction method is proposed as an important component of the classification system. This BPA construction method provides the ability to reflect the differences of classifier's classification

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

Algorithm 1: The evidential multi-classifier system based on the proposed BPA construction method.

Input: Base classifiers $\varphi_1, \ldots, \varphi_l$, set of classes Ω, training data set $D_{training}$ and test data set D_{test}.
Output: Predicted classes of unclassified targets in D_{test}, indicated by S_{pred}.

`/* Construct the confusion matrix for each base classifier. */`
Create initial confusion matrices $C_{\varphi_1}, \ldots, C_{\varphi_l}$;
foreach *sample s whose real class is* c_s^{real} *in* $D_{training}$ **do**
  foreach *base classifier* φ_k *in* $\{\varphi_1, \ldots, \varphi_l\}$ **do**
    Predict the class of s: $c_s^{pred} \leftarrow \varphi_k(s)$;
    Compare c_s^{real} and c_s^{pred}, then update the confusion matrix C_{φ_k};
  end
end
`/* For each base classifier` φ_k`, generate a set of BPAs` M^{φ_k} `to represent the recognition ability of` φ_k `to different classes in` Ω`. */`
foreach C_{φ_k} *in* $\{C_{\varphi_1}, \ldots, C_{\varphi_l}\}$ **do**
  Calculate recall matrix $C_{\varphi_k^r}$ and precision matrix $C_{\varphi_k^p}$;
  foreach *class i in* Ω **do**
    Construct recall BPA m_i^r and precision BPA m_i^p by means of Eqs. (15) - (20);
    Combine m_i^r and m_i^p: $m_i^{\varphi_k} \leftarrow m_i^r \oplus m_i^p$;
    Set $M^{\varphi_k}[\text{“}i\text{”}] \leftarrow m_i^{\varphi_k}$;
  end
end
`/* Predict the class of each unclassified target in` D_{test}`. */`
foreach *unclassified target T in* D_{test} **do**
  foreach *base classifier* φ_k *in* $\{\varphi_1, \ldots, \varphi_l\}$ **do**
    Predict the class of T: $c_T^{\varphi_k} \leftarrow \varphi_k(T)$;
  end
  $m_T \leftarrow g(M^{\varphi_1}, c_T^{\varphi_1}) \oplus g(M^{\varphi_2}, c_T^{\varphi_2}) \oplus \cdots \oplus g(M^{\varphi_l}, c_T^{\varphi_l})$, where $g(M^{\varphi_k}, c_T^{\varphi_k}) \leftarrow M^{\varphi_k}[\text{“}c_T^{\varphi_k}\text{”}]$, $k \in \{1, \ldots, l\}$; `/* Combine the prediction results of all base classifiers. */`
  $c_T^{pred} \leftarrow \arg\max_x \left\{ \sum_{A \subseteq \Omega, x \in A} \frac{1}{|A|} \frac{m_T(A)}{1 - m_T(\emptyset)} \right\}$, $x \in \Omega$; `/* Obtain the predicted class of` T `based on the pignistic transformation. */`
  Set $S_{pred}[\text{“}T\text{”}] \leftarrow c_T^{pred}$;
end
Return S_{pred}.

abilities to different classes. At last, different BPA combination rules can be used to synthesize the results of multiple base classifiers in terms of different application scenarios, which enhances the flexibility of the classification system.

4. Case study

In this section, the prediction of transmembrane protein topology is used to illustrate the our proposed method, partial results are from our previous study [14]. The topology of transmembrane proteins, i.e. the number and position of the transmembrane helixes and the in/out location of the N and C terminals of the protein sequence, is an important issue in the study of transmembrane proteins [21,36]. For a protein sequence, if both the transmembrane helixes and location of the N and C terminals have been correctly predicted, the topology of the protein sequence is said to be correctly predicted. As shown in Fig. 3, given a protein sequence, the task is to determine the class label for each residue among three classes: “i”(intracellular), “M”(transmembrane), “o”(extracellular).

In this case, a data set of 125 transmembrane protein sequences with known topology is collected from the data set MPtopo [30]. MPtopo (Membrane Proteins topology) is a growing database of membrane proteins whose topologies have been verified experimentally by means of crystallography, gene fusion, and other methods. In MPtopo, these membrane proteins are divided into three subsets, namely 3D_Helix, 1D_Helix and 3D_Other. When we conducted this experiment, the subset 3D_Helix includes 125 protein sequences of known 3D structure, where the topologies of these proteins are highly accurate so that they are usually used for evaluating existing prediction algorithms and creating new ones. In the paper, the 125 transmembrane protein sequences with known topologies are selected to verify the proposed method. These protein sequences totally contain 515 transmembrane helixes and 29902 residues. Five individual algorithms, OCTOPUS [56],

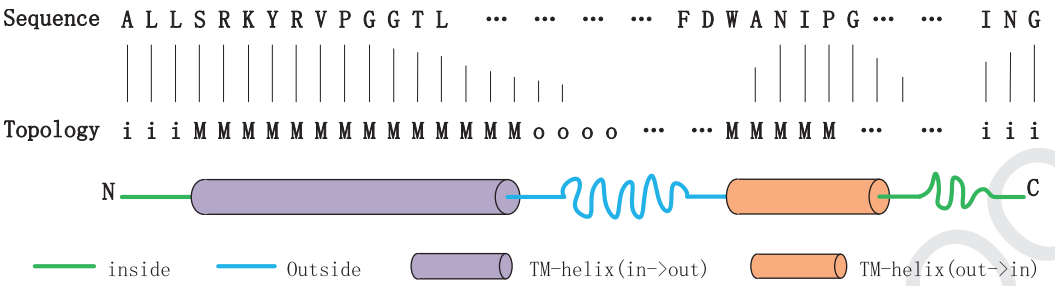

Fig. 3. Prediction of transmembrane proteins topology.

Table 1
The performance of various algorithms in the form of a confusion matrix.

Truth (Actual class)	Algorithm	Prediction i	M	o
i	OCTOPUS	7655	389	839
	PRO	7574	450	859
	PRODIV	7323	442	1118
	SCAMPI-msa	7655	389	839
	SCAMPI-seq	7359	455	1069
	Parikh	7247	540	1096
	Proposed	7636	358	889
M	OCTOPUS	1877	9785	1458
	PRO	1922	9588	1610
	PRODIV	1819	9884	1417
	SCAMPI-msa	1877	9785	1458
	SCAMPI-seq	1907	9628	1585
	Parikh	1550	10374	1196
	Proposed	1802	9811	1507
o	OCTOPUS	1230	578	6091
	PRO	1051	714	6134
	PRODIV	1117	775	6007
	SCAMPI-msa	1230	578	6091
	SCAMPI-seq	1101	564	6234
	Parikh	1631	810	5458
	Proposed	916	518	6465

PRO-TMHMM and PRODIV-TMHMM [55], SCAMPI-msa and SCAMPI-seq [3], have been used for the prediction. For these algorithms, the prediction of transmembrane helix is considered successful when the overlapping region of predicted and observed transmembrane segment contains at least 9 amino acids. If all transmembrane helixes and the orientation of the transmembrane protein sequence have been correctly predicted, the topology of transmembrane protein is said to be correctly predicted.

A combination predictor is constructed based on these five individual algorithms. In order to reflect the performance of the combination predictor faithfully and to avoid overfitting, the experiment is performed and verified using a ten-fold cross validation. Each fold roughly contains 12–13 transmembrane proteins and their total homology has been reduced by 30% by using cd-hit program [34] to avoid the influence of homology. The combination of five algorithms has been implemented to determine the class of each residue. Parikh et al.'s method and the proposed method in this paper are adopted, respectively. In addition, unless otherwise specified, the combination rule used in this paper is the Dempster's rule of combination. The impact of combination rules on the results will be specially studied later. By using ten-fold cross validation, the final classification result can be derived. For the sake of comparison, the performance of various algorithms are shown in Table 1 in the form of a confusion matrix. In Table 1, three actual classes i, M, and o, are listed in the left-most column, the predicted classes are listed in the three right-most columns, and the related algorithms are given in the middle column.

According to Table 1, the recall rate, precision rate and F-score ($F-score = \frac{2\times recall\times precision}{recall+precision}$) for each algorithm and each class, as well as the overall accuracy of residue, can be calculated. They are shown in Table 2. As shown in the last column of Table 2, thanks to utilizing the various available information, not only the precision rate but also recall rate, contained in each base classifier's confusion matrix, the proposed method occupies the highest accuracy for the residue prediction.

As can be found from Table 2, as for other classification performance measures such as each class' recall rate, precision rate and F-score, the proposed method also shows better performance than other algorithms. Particularly, in order to compare the proposed BPA construction method with Parikh et al.'s method, Fig. 4 graphically shows the results of recall rate, precision rate, F-score, as well as accuracy rate, derived from different classifier. It is found that the proposed method

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

Table 2
Performance of various algorithms on different performance measures.

Algorithm	Class	Recall(%)	Precision(%)	F-score	Accuracy(%)
	i	86.18	71.13	0.7793	
OCTOPUS	M	74.58	91.01	0.8198	78.69
	o	77.11	72.62	0.7480	
	i	85.26	71.81	0.7796	
PRO	M	73.08	89.17	0.8033	77.91
	o	77.66	71.30	0.7434	
	i	82.44	71.38	0.7651	
PRODIV	M	75.34	89.04	0.8162	77.63
	o	76.05	70.32	0.7307	
	i	86.18	71.13	0.7793	
SCAMPI-msa	M	74.58	91.01	0.8198	78.69
	o	77.11	72.62	0.7480	
	i	82.84	70.98	0.7646	
SCAMPI-seq	M	73.38	90.43	0.8102	77.66
	o	78.92	70.14	0.7427	
	i	81.58	69.50	0.7506	
Parikh	M	79.07	88.49	0.8351	77.18
	o	69.10	70.43	0.6976	
	i	85.96	73.75	0.7939	
Proposed	M	74.78	91.80	0.8242	79.97
	o	81.85	72.96	0.7715	

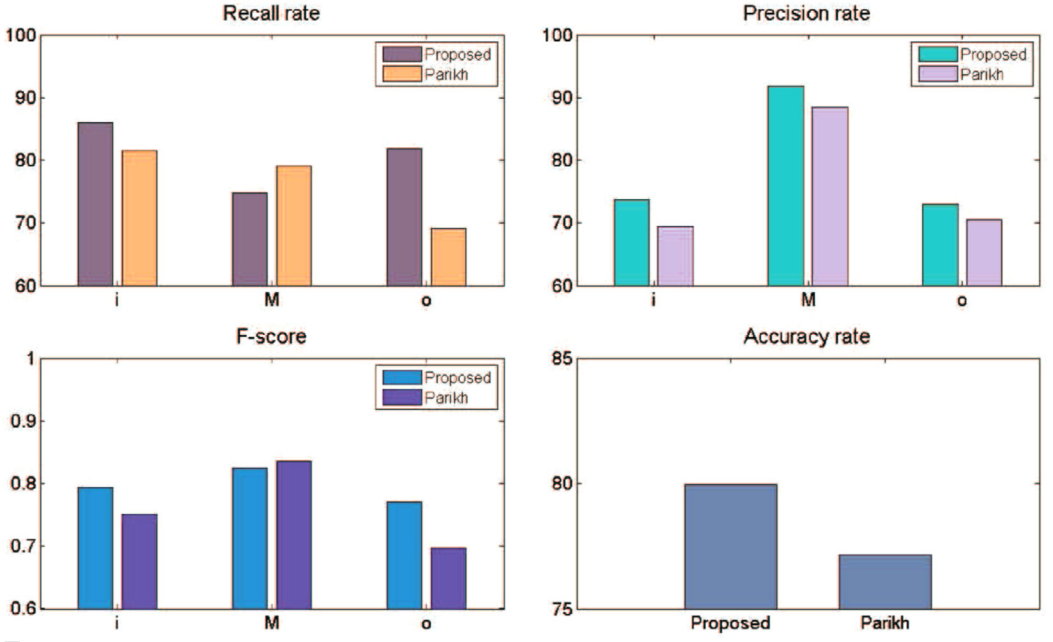

Fig. 4. Performance of combination based on different BPA construction methods.

outperforms Parikh et al's method in most performance indices. With regard to the precision rate and accuracy rate, the proposed method outperforms Parikh et al's method. With regard to the recall rate and F-score, although the results of proposed method are inferior to Parikh et al's with for class "M", but for classes "i" and "o", the proposed method outperforms Parikh et al's. The results shows that the proposed method is effective to a certain degree.

As for the whole topology prediction of protein sequence, Table 3 shows various algorithms' performance. It is clear that the proposed method has the highest prediction accuracy of transmembrane proteins topology, and it is 4.8% higher than the worst algorithm. The performance of Parikh et al's method is in the middle level, it outperforms PRODIV and SCAMPI-seq. In general, regarding the topology's prediction accuracy, the combination classifier based on the proposed BPA construction method is also better than other algorithms since abundant information contained in the confusion matrix has been considered in the process of BPA construction.

In this paper, apart from adopting different BPA construction method, the combination classifier is based on five individual algorithms including OCTOPUS, PRO-TMHMM, PRODIV-TMHMM, SCAMPI-msa and SCAMPI-seq. In order to study the

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

Table 3
Prediction accuracy of transmembrane proteins topology of various algorithms.

Algorithm	Prediction accuracy of transmembrane proteins topology(%)
OCTOPUS	71.2
PRO	70.4
PRODIV	69.6
SCAMPI-msa	71.2
SCAMPI-seq	69.6
Parikh	70.4
Proposed	74.4

Table 4
The performances of combination classifiers based on these two BPA construction methods (proposed and Parikh et al's [40]) under different combination schemes.

Basic algorithms					Prediction accuracy of residue (%)		Prediction accuracy of topology (%)	
OCTOPUS	PRO	PRODIV	SCAMPI-msa	SCAMPI-seq	Parikh	Proposed	Parikh	Proposed
✓	✓	✓	✓	✓	77.18	79.97	70.4	74.4
	✓	✓	✓	✓	78.80	78.15	70.4	72.8
✓	✓	✓	✓		77.99	78.71	69.6	72.0
✓			✓	✓	77.81	78.69	70.4	71.2
✓	✓		✓		77.17	78.69	69.6	71.2
✓		✓			76.63	78.60	68.0	71.2
	✓	✓			77.24	78.18	69.9	70.4

Table 5
Significance level P value of matched-pair chi-square test on the results of residue prediction between different algorithms.

Algorithm	Parikh	PRO	SCAMPI-msa	PRODIV	SCAMPI-seq	OCTOPUS
Parikh	–	<0.001	<0.001	<0.001	<0.001	<0.001
Proposed	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Table 6
Significance level P value of matched-pair chi-square test on the results of topology prediction between different algorithms.

Algorithm	Parikh	PRO	SCAMPI-msa	PRODIV	SCAMPI-seq	OCTOPUS
Parikh	–	1.000	1.000	1.000	1.000	1.000
Proposed	0.180	0.267	0.388	0.210	0.180	0.388

impact of different combination schemes of individual algorithms on the classification performance, we conduct the classification task under different combination schemes. The performances of combination classifiers based on these two BPA construction methods (proposed and Parikh et al's) are shown in Table 4. As can be seen from Table 4, the proposed BPA construction method generally outperforms Parikh et al's method, whether for the index of topology's prediction accuracy or for the index of residue's prediction accuracy. When these four algorithms PRO-TMHMM, PRODIV-TMHMM, SCAMPI-msa and SCAMPI-seq, are combined, although Parikh et al's method has a little higher prediction accuracy of residue, but the topology prediction accuracy of the proposed method exceeds that of Parikh et al's method more than 2% . Therefore, as displayed in Table 4, the superiority of proposed BPA construction method has been demonstrated.

Aiming to examine whether these results mentioned above are statistically meaningful or not, the matched-pair chi-square test is used on the results of residue and topology prediction between different algorithms. "Proposed" and "Parikh" are still the combination of five individual algorithms based on different BPA construction methods. Tables 5 and 6 show the outcomes, respectively. Herein, the results are said statistically significant if $P \leq 0.05$. As can be found in Table 5, regarding residue prediction, the significance level P value is less than 0.001 in each pair of algorithms, which means the difference is significant among the results of residue prediction obtained from different algorithms. Therefore, it is statistically meaningful that the proposed method outperforms other algorithms in terms of the results of residue prediction. Conversely, according to Table 6, the results of topology prediction are not with significant difference in statistic. However, even that, the significance level P value of proposed method is also lower than that of Parikh et al's method in each case. The lower the value of P, the more probability the results are statistically significant.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

Table 7
Performance of different combination rules based on the proposed BPA construction method.

Combination rule	Class	Recall(%)	Precision(%)	F-score	Prediction accuracy of residue (%)	Prediction accuracy of topology (%)
Dempster's rule	i	85.96	73.75	0.7939		
	M	74.78	91.80	0.8242	79.97	74.40
	o	81.85	72.96	0.7715		
Yager's rule [66]	i	85.96	73.75	0.7939		
	M	74.78	91.80	0.8242	79.97	74.40
	o	81.85	72.96	0.7715		
Dubois and Prade's rule [26]	i	74.20	59.00	0.6573		
	M	72.41	88.09	0.7948	69.13	59.20
	o	57.99	57.66	0.5783		
Murphy's rule [39]	i	85.96	73.74	0.7938		
	M	74.76	91.81	0.8241	79.96	74.40
	o	81.86	72.95	0.7715		
Deng's rule [23]	i	85.96	73.75	0.7939		
	M	74.78	91.80	0.8242	79.97	74.40
	o	81.85	72.96	0.7715		
PCR6 [46]	i	85.96	73.75	0.7939		
	M	74.78	91.80	0.8242	79.97	74.40
	o	81.85	72.96	0.7715		

There are several reasons causing that the results are statistically significant in residue prediction and not statistically significant in topology prediction. For example, the amount of used transmembrane protein sequences is so small that the results are not statistically significant in topology prediction. Apart from that, the main reason is the prediction conducts on different scale or granularity. In this case of transmembrane protein topology prediction, the evidential multiple classifiers system directly works on the scale of residue, the class of each residue is immediately determined by the synthesis of five individual algorithms' outcomes. As a broader scale, the topology of each transmembrane protein sequence is not directly impacted by the evidential multiple classifiers system. The determination of transmembrane protein sequence topology is based on the combination prediction of residue. Therefore, the prediction accuracy of residue is the principal index to reflect the classification performance of combination classifiers. From this viewpoint, it is statistically significant that the evidential combination of multiple classifiers based on the propose BPA construction method outperforms other algorithms and the proposed BPA construction method outperforms Parikh et al's method.

Essentially, the evidential multiple classifiers system is also a kind of ensemble classifiers. Within such system, the idea of multi-information fusion has been imported to combine the results derived from various base classifiers. Specifically, the base classifiers are treated as various sources of evidences, the uncertainty of information has been represented and handled in the forms of BPAs. Within this framework, the combination rule of evidences has played an important role, especially when evidences are in conflict. In order to study the impact of combination rules on the results, Table 7 shows the performance of different combination rules in this case study based on the proposed BPA construction method. Here, we have presented the results of several typical combination rules including Dempster's rule (normalized conjunctive rule), Yager's global conflict redistribution rule, Dubois and Prade's disjunctive rule, Murphy's simple averaging rule, Deng's weighted averaging rule, and proportional conflict redistribution PCR6 rule. From Table 7, it is found that these combination rules have produced the same results basically, except Dubois and Prade's rule. When using Dubois and Prade's rule, there are only 69.13% and 59.20% prediction accuracy for residue and topology, respectively. The reason of resulting in the unsatisfying results is that the use of disjunctive rule is just proper when the conflict is caused by poor reliability of some of the information sources. However, in this case study, the five base classifiers are reliable information sources. So the class of conjunctive rules, such as Dempster's rule, is more appropriate here. Through studying the impact of combination rules on the results, we have confirmed the important influence of combination rules. Therefore, we should pay great attention on the selection of the appropriate combination rule. Several key factors must be taken into consideration, for example the independence and reliability of evidences, degree of conflict among evidences, completeness of the frame of discernment, and so on. A lot of works have been proposed to address this issue, for recent advances please refer to [19,28,33,67].

5. Conclusion

As frequently stressed in previous studies, the determination of BPA is one of the most key problem in the application of Dempster–Shafer evidence theory. And in an evidential multiple classifiers system, the construction of BPA has a great influence on the classification performance. In this paper, an improved BPA construction method is proposed based on the confusion matrix for classification problem. The proposed method makes full use of the available information contained in the confusion matrix. The case study of transmembrane protein topology prediction demonstrates the effectiveness and superiority of the proposed BPA construction method, compared with other BPA construction method. The proposed confusion matrix based method of BPA construction is also useful for other applications such as ensemble learning, cluster analysis, etc.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

Acknowledgments

The authors thank the anonymous reviewers for their valuable comments and suggestions to improve this paper. 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), the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No. BUAA-VR-14KF-02), Fundamental Research Funds for the Central Universities (Grant No. XDJK2014D034), and China Scholarship Council.

References

[1] A. Al-Ani, M. Deriche, A new technique for combining multiple classifiers using the Dempster–Shafer theory of evidence, J. Artif. Intell. Res. 17 (2002) 333–361.
[2] H. Altincay, On the independence requirement in Dempster-Shafer theory for combining classifiers providing statistical evidence, Appl. Intell. 25 (1) (2006) 73–90.
[3] A. Bernsel, H. Viklund, J. Falk, E. Lindahl, G. von Heijne, A. Elofsson, Prediction of membrane-protein topology from first principles, Proc. Nat. Acad. Sci. U.S.A 105 (20) (2008) 7177–7181.
[4] M.J. Beynon, R. Andrews, G.A. Boyne, Evidence-based modelling of strategic fit: an introduction to RCaRBS, Eur. J. Oper. Res. 207 (2) (2010) 886–896.
[5] Y. Bi, J. Guan, D. Bell, The combination of multiple classifiers using an evidential reasoning approach, Artif. Intell. 172 (15) (2008) 1731–1751.
[6] S. Boccaletti, G. Bianconi, R. Criado, C. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendina-Nadal, Z. Wang, M. Zanin, The structure and dynamics of multilayer networks, Phys. Rep. 544 (1) (2014) 1–122.
[7] A. Burkov, B. Chaib-draa, Computing equilibria in discounted dynamic games, Appl. Math. Comput. 269 (2015) 863–884.
[8] C.-C. Chou, L.-J. Liu, S.-F. Huang, J.-M. Yih, T.-C. Han, An evaluation of airline service quality using the fuzzy weighted SERVQUAL method, Appl. Soft Comput. 11 (2) (2011) 2117–2128.
[9] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Stat. 38 (2) (1967) 325–339.
[10] X. Deng, Y. Deng, F.T. Chan, An improved operator of combination with adapted conflict, Ann. Oper. Res. 223 (1) (2014) 451–459.
[11] X. Deng, Y. Hu, F.T. Chan, S. Mahadevan, Y. Deng, Parameter estimation based on interval-valued belief structures, Eur. J. Oper. Res. 241 (2) (2015) 579–582.
[12] X. Deng, Y. Hu, Y. Deng, S. Mahadevan, Environmental impact assessment based on D numbers, Expert Syst. Appl. 41 (2) (2014) 635–643.
[13] X. Deng, Y. Hu, Y. Deng, S. Mahadevan, Supplier selection using AHP methodology extended by D numbers, Expert Syst. Appl. 41 (1) (2014) 156–167.
[14] X. Deng, Q. Liu, Y. Hu, Y. Deng, TOPPER: topology prediction of transmembrane protein based on evidential reasoning, Scientific World J. 2013 (2013) 123731.
[15] X. Deng, Q. Liu, R. Sadiq, Y. Deng, Impact of roles assignation on heterogeneous populations in evolutionary dictator game, Scientific Rep. 4 (2014) 6937.
[16] X. Deng, X. Lu, F.T. Chan, R. Sadiq, S. Mahadevan, Y. Deng, D-CFPR: D numbers extended consistent fuzzy preference relations, Knowl. Based Syst. 73 (2015) 61–68.
[17] X. Deng, Z. Wang, Q. Liu, Y. Deng, S. Mahadevan, A belief-based evolutionarily stable strategy, J. Theor. Biol. 361 (2014) 81–86.
[18] X. Deng, X. Zheng, X. Su, F.T. Chan, Y. Hu, R. Sadiq, Y. Deng, An evidential game theory framework in multi-criteria decision making process, Appl. Math. Comput. 244 (2014) 783–793.
[19] Y. Deng, Generalized evidence theory, Appl. Intell. 43 (3) (2015) 530–543.
[20] Y. Deng, A threat assessment model under uncertain environment, Math. Probl. Eng. 2015 (2015), doi:10.1155/2015/878024.
[21] Y. Deng, Q. Liu, Y.X. Li, Scoring hidden Markov models to discriminate β-barrel membrane proteins, Comput. Biol. Chem. 28 (3) (2004) 189–194.
[22] Y. Deng, Y. Liu, D. Zhou, An improved genetic algorithm with initial population strategy for symmetric TSP, Math. Probl. Eng. 501 (2015) 212794.
[23] Y. Deng, W. Shi, Z. Zhu, Q. Liu, Combining belief functions based on distance of evidence, Dec. Supp. Syst. 38 (3) (2004) 489–493.
[24] T. Denoeux, A k-nearest neighbor classification rule based on Dempster-Shafer theory, IEEE Trans. Syst. Man Cybern. 25 (5) (1995) 804–813.
[25] L. Diaz-Mas, R. Munoz-Salinas, F. Madrid-Cuevas, R. Medina-Carnicer, Shape from silhouette using Dempster-Shafer theory, Patt. Recognit. 43 (6) (2010) 2119–2131.
[26] D. Dubois, H. Prade, A set-theoretic view of belief functions logical operations and approximations by fuzzy sets? Int. J. Gen. Syst. 12 (3) (1986) 193–226.
[27] X.F. Fan, M.J. Zuo, Fault diagnosis of machines based on D-S evidence theory. Part 2: application of the improved D-S evidence theory in gearbox fault diagnosis, Pattern Recog. Lett. 27 (5) (2006) 377–385.
[28] M.C. Florea, A.-L. Jousselme, É. Bossé, D. Grenier, Robust combination rules for evidence theory, Inf. Fusion 10 (2) (2009) 183–197.
[29] R. Ilin, E. Blasch, Information fusion with belief functions: a comparison of proportional conflict redistribution PCR5 and PCR6 rules for networked sensors, in: Proceedings of the 18th International Conference on Information Fusion (FUSION), Washington, DC, 2015, pp. 2084–2091.
[30] S. Jayasinghe, K. Hristova, S.H. White, Mptopo: a database of membrane protein topology, Protein Sci. 10 (2) (2001) 455–458.
[31] C. Kahraman, S.C. Onar, B. Oztaysi, Fuzzy multicriteria decision-making: a literature review, Int. J. Comput. Intell. Syst. 8 (4) (2015) 637–666.
[32] B. Kang, Y. Deng, R. Sadiq, S. Mahadevan, Evidential cognitive maps, Knowl.-Based Syst. 35 (2012) 77–86.
[33] E. Lefèvre, Z. Elouedi, How to preserve the conflict as an alarm in the combination of belief functions? Decis. Support Syst. 56 (2013) 326–333.
[34] W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics 22 (13) (2006) 1658–1659.
[35] Y. Li, X. Lan, X. Deng, R. Sadiq, Y. Deng, Comprehensive consideration of strategy updating promotes cooperation in the prisoner's dilemma game, Physica A: Stat. Mech. Appl. 403 (2014) 284–292.
[36] Q. Liu, Y.S. Zhu, B.H. Wang, Y.X. Li, A HMM-based method to predict the transmembrane regions of β-barrel membrane proteins, Comput. Biol. Chem. 27 (1) (2003) 69–76.
[37] Z. Liu, Q. Pan, J. Dezert, A new belief-based K-nearest neighbor classification method, Pattern Recog. 46 (3) (2013) 834–844.
[38] Z.-g. Liu, Q. Pan, J. Dezert, A belief classification rule for imprecise data, Appl. Intell. 40 (2) (2014) 214–228.
[39] C.K. Murphy, Combining belief functions when evidence conflicts, Decis. Supp. Syst. 29 (1) (2000) 1–9.
[40] C.R. Parikh, M.J. Pont, N.B. Jones, Application of Dempster-Shafer theory in condition monitoring applications: a case study, Pattern Recog. Lett. 22 (6-7) (2001) 777–785.
[41] C.R. Parikh, M.J. Pont, N.B. Jones, F.S. Schlindwein, Improving the performance of CMFD applications using multiple classifiers and a fusion framework, Trans. Inst. Measurement Control 25 (2) (2003) 123–144.
[42] B.S. Reddy, O.A. Basir, Concept-based evidential reasoning for multimodal fusion in human computer interaction, Appl. Soft Comput. 10 (2) (2010) 567–577.
[43] G. Rogova, Combining the results of several neural network classifiers, Neural Netwo. 7 (5) (1994) 777–781.
[44] C. Sammut, G.I. Webb, Encyclopedia of Machine Learning, Springer, 2011.
[45] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.
[46] F. Smarandache, J. Dezert (Eds.), Advances and Applications of DSmT for Information Fusion vol. 1–4, American Research Press, Rehoboth, NM, USA, 2004–2015. http://fs.gallup.unm.edu//DSmT.htm.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033

[47] F. Smarandache, J. Dezert, On the consistency of PCR6 with the averaging rule and its application to probability estimation, in: Proceedings of the 16th International Conference on Information Fusion (FUSION), Istanbul, 2013, pp. 1119–1126.
[48] P. Smets, R. Kennes, The transferable belief model, Artif. Intell. 66 (2) (1994) 191–234.
[49] X. Su, S. Mahadevan, P. Xu, Y. Deng, Dependence assessment in human reliability analysis using evidence theory and AHP, Risk Anal. 35 (2015) 1296–1316.
[50] R. Sun, H.Z. Huang, Q. Miao, Improved information fusion approach based on D-S evidence theory, J. Mech. Sci. Technol. 22 (12) (2008) 2417–2425.
[51] M. Tabassian, R. Ghaderi, R. Ebrahimpour, Knitted fabric defect classification for uncertain labels based on Dempster–hafer theory of evidence, Expert Syst. Appl. 38 (5) (2011) 5259–5267.
[52] M. Tabassian, R. Ghaderi, R. Ebrahimpour, Combining complementary information sources in the Dempster-Shafer framework for solving classification problems with imperfect labels, Knowl. Based Syst. 27 (2012) 92–102.
[53] Z. Tao, X. Liu, H. Chen, Z. Chen, Group decision making with fuzzy linguistic preference relations via cooperative games method, Comput. Indust. Eng. 83 (2015) 184–192.
[54] C. Thiel, F. Schwenker, G. Palm, Using Dempster-Shafer theory in MCF systems to reject samples, Lect. Notes Comput. Sci. 3541 (2005) 959–961.
[55] H. Viklund, A. Elofsson, Best α-helical transmembrane protein topology predictions are achieved using hidden Markov models and evolutionary information, Protein Sci. 13 (7) (2004) 1908–1917.
[56] H. Viklund, A. Elofsson, OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar, Bioinformatics 24 (15) (2008) 1662–1668.
[57] J.-Q. Wang, R.-R. Nie, H.-Y. Zhang, X.-H. Chen, Intuitionistic fuzzy multi-criteria decision-making method based on evidential reasoning, Appl. Soft Comput. 13 (4) (2013) 1823–1831.
[58] Z. Wang, S. Kokubo, M. Jusup, J. Tanimoto, Universal scaling for the dilemma strength in evolutionary games, Phys. Life Rev. 14 (2015) 1–30.
[59] Z. Wang, S. Kokubo, J. Tanimoto, E. Fukuda, K. Shigaki, Insight into the so-called spatial reciprocity, Phys. Rev. E 88 (4) (2013) 042145.
[60] Z. Wang, A. Szolnoki, M. Perc, Evolution of public cooperation on interdependent networks: the impact of biased utility functions, EPL (Europhys. Lett.) 97 (4) (2012) 48001.
[61] Z. Wang, A. Szolnoki, M. Perc, Interdependent network reciprocity in evolutionary games, Sci. Rep. 3 (2013) 1183.
[62] Z. Wang, A. Szolnoki, M. Perc, Rewarding evolutionary fitness with links between populations promotes cooperation, J. Theor. Biol. 349 (2014) 50–56.
[63] Z. Wang, L. Wang, A. Szolnoki, M. Perc, Evolutionary games on multilayer networks: a colloquium, Eur. Phys. J. B 88 (5) (2015) 1–15.
[64] Z. Wang, C.-Y. Xia, S. Meloni, C.-S. Zhou, Y. Moreno, Impact of social punishment on cooperative behavior in complex networks, Sci. Rep. 3 (2013) 3055.
[65] L. Xu, A. Krzyzak, C.Y. Suen, Methods of combining multiple classifiers and their applications to handwriting recognition, IEEE Trans. Syst. Man Cybern. 22 (3) (1992) 418–435.
[66] R.R. Yager, On the Dempster-Shafer framework and new combination rules, Inf. Sci. 41 (2) (1987) 93–137.
[67] R.R. Yager, Combining various types of belief structures, Inf. Sci. 303 (2015) 83–100.
[68] J.-B. Yang, Y.-M. Wang, D.-L. Xu, K.-S. Chin, L. Chatton, Belief rule-based methodology for mapping consumer preferences and setting product targets, Expert Syst. Appl. 39 (5) (2012) 4749–4759.
[69] X. Zhang, Y. Deng, F.T. Chan, P. Xu, S. Mahadevan, Y. Hu, IFSJSP: a novel methodology for the job-shop scheduling problem based on intuitionistic fuzzy sets, Int. J. Prod. Res. 51 (17) (2013) 5100–5119.

Please cite this article as: X. Deng et al., An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.01.033