

A New Evidential Distance Measure Based on Belief Intervals

V. Khatibi¹ and G.A. Montazer^{1,*}

Abstract. *So far, most of the evidential distance and similarity measures proposed in the Dempster-Shafer theory literature have been based on the basic belief assignment function, so as the belief and plausibility functions as two main results of the theory are not directly used in this regard. In this paper, a new evidential distance measure is proposed based on these functions according to nearest neighborhood concept. After assigning basic belief values to propositions and constructing the belief and plausibility functions or the belief interval, this evidential distance measure compares the similarity between the unknown pattern and class belief intervals. For this purpose, we first acquire the belief and plausibility functions or the belief intervals and then the distance between the belief intervals of uncertain pattern feature vectors and samples are calculated. We applied this novel distance measure to the bacillus colonies recognition and coronary heart disease patients classification problems to examine the proposed measure capability in contrast to other evidential measures. Our experiment illustrates that the belief interval distance measure yields the accuracy rates of 91.66 and 92.45 percent for unknown bacillus patterns recognition and coronary heart disease patients classification, respectively, which in contrast to other evidential measures shows superior performance.*

Keywords: *Evidence theory; Approximate reasoning; Pattern recognition; Belief interval distance; Bacillus colony recognition; Coronary heart disease patients classification.*

INTRODUCTION

Pattern classification is concerned with the assignment of unknown patterns, represented by feature vectors, to predefined categories or classes [1]. In fact, pattern recognition problems typically involve the classification of an unknown pattern L to a given set of K prototypes $P_k, k \in \{1, 2, \dots, K\}$ [2-4]. Each prototype P_k belongs to a given class $C_m, m \in \{1, 2, \dots, M\}$, which is specified by the indicator function A_k :

$$A_k = l_m, \text{ if } P_k \text{ belongs to the } m\text{th class } l_m. \quad (1)$$

Let $S(L, P_k)$ be a similarity measure which measures the degree of similarity, or compatibility, between the unknown pattern L and the k th prototype P_k . Then, formally we may write the process of classifying or

assigning the unknown pattern L to the class $l^* = A_{k^*}$, where:

$$k^* = \arg \max_k (S(L, P_k)). \quad (2)$$

The pattern recognition problems could be found in many scientific disciplines [5-7], and are considered as one of the most applicable topics in engineering [8-10]. On the other hand, several evidential measures have been proposed so far, which use basic belief assignment function as a basis in classification [11-18], whereas in evidence theory literature, there is no work which relies mainly and directly on the belief intervals in this regard. The belief interval comprised of belief and plausibility functions constitutes all the knowledge available in the problem for a proposition, regardless of baseless assumptions [19]. Hence, using belief intervals leads us to an approximate reasoning approach which is more accurate than other methods. In this paper, a novel evidential distance measure based on nearest neighborhood concept is proposed. Our goal is to develop an approach of approximate reasoning in which

1. Department of Information Technology, School of Engineering, Tarbiat Modares University, P.O. Box 14115-179, Tehran, Iran.

*. Corresponding author. E-mail: montazer@modares.ac.ir

Received 27 September 2009; received in revised form 25 May 2010; accepted 23 August 2010

uncertain feature vectors and samples are classified according to existing knowledge and partial truth while considering problem uncertainty. In this approach, we first acquire basic belief assignments for the propositions, and then construct the belief and plausibility functions or the belief intervals. This procedure is done for all the attributes, so as we have a feature vector of belief intervals for the attributes. On the other hand, the appropriate feature vector of belief intervals for the classes can be built on several mechanisms such as investigations, observations, expertise etc. Then, applying the proposed evidential distance, we could identify the most similar class to each unknown pattern according to nearest neighborhood concept. Also, this evidential distance is applied to two medical diagnosis problems, and the acquired results are compared with other evidential measures.

This paper is organized as follows: In the next section, the fundamental notions of evidence theory are presented briefly. Besides, the evidential distance measures in Dempster-Shafer theory literature are discussed. Also, a new evidential distance measure is proposed which relies mainly on belief intervals. Then, we apply the proposed measure to the bacillus colonies recognition and coronary heart disease patients classification problems, and the experimental results are illustrated and compared with other evidential measures in the last section.

DEMPSTER-SHAFER EVIDENCE THEORY

Evidence theory was first introduced by Dempster in 1968 [20]. Later, it was refined and further developed by Shafer in 1976 [21]. Also, Smets introduced Transferable Belief Model (TBM) which extends Dempster-Shafer model to represent nonprobabilistic beliefs [22,23].

Evidence theory starts by defining a frame of discernment that is a set of mutually exclusive “elementary” propositions. The finest subdivision of the set is considered as the elementary proposition. Therefore, the frame of discernment denoted by Ω or X consists of all elementary propositions [24].

In evidence theory, the basic propagation of information is through Basic Belief Assignment (BBA) function [19]. BBA expresses our degree of belief in a proposition. It is determined by various forms of information: sources, experimental methods, quantity and quality of information, experts’ opinions and so on. BBA is assigned by making use of a mapping function (m) in order to express our belief in a proposition with a number in the unit interval $[0,1]$, as shown below:

$$m : 2^\Omega \rightarrow [0, 1]. \quad (3)$$

This measure m (the basic belief assignment function)

must satisfy the following three properties:

$$i. \quad m(A) \geq 0 \text{ for any } A \in 2^\Omega, \quad (4)$$

$$ii. \quad m(\Phi) = 0, \quad (5)$$

$$iii. \quad \sum_{A \in 2^\Omega} m(A) = 1. \quad (6)$$

After assigning the basic belief degrees to the propositions, we can proceed to determine the belief and plausibility functions. Due to a lack of information, it is more reasonable to present bounds for the result of uncertainty quantification as opposed to a single value of belief. Our total degree of belief in a proposition “ A ” is expressed within an interval, $[Bel(A), Pls(A)]$, called the Belief Interval (BI) [19], which lies in the unit interval $[0,1]$, as shown in Figure 1, where Bel and Pls functions are explained forth.

The lower bound, Belief (Bel), also known as Support (Sup), for a set A is defined as the sum of all the basic belief assignments of the proper subsets C_i of the set of interest A , i.e., $C_i \subseteq A$. The general relation between BBA and belief functions can be written as:

$$Bel(A) = \sum_{C_i \subseteq A} m(C_i). \quad (7)$$

It can be shown that $Bel(\Phi) = 0$ and $Bel(\Omega) = 1$.

The upper bound, Plausibility (Pls), is the summation of basic belief assignments of the sets C_i whose intersection with the proposition A is not an empty set, i.e. $C_i \cap A \neq \Phi$, and therefore, it can be written as:

$$Pls(A) = \sum_{C_i \cap A \neq \Phi} m(C_i). \quad (8)$$

Also, it can be shown that $Pls(\Phi) = 0$ and $Pls(\Omega) = 1$. In fact, every proposition that allows for the proposition A to be included at least partially is considered to imply the plausibility of the proposition A .

Belief functions cannot be directly used for decision making. Hence we can apply the pignistic transformation: a particular mapping of a belief function m to a Bayesian belief function (probability function) $BetP$. The resulting pignistic probability for the singletons $\theta_i \in \Theta$ is given by [5]:

$$BetP(\{\theta_i\}) = \sum_{\theta_i \in A \subseteq \Theta} \frac{1}{|A|} \frac{m(A)}{1 - m(\Phi)}. \quad (9)$$

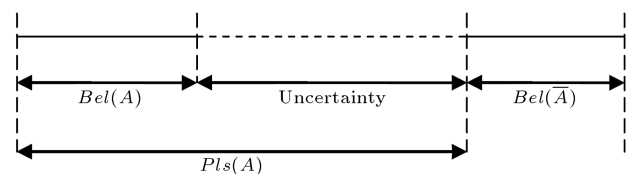


Figure 1. Belief (Bel), plausibility (Pls), and uncertainty.

In a sense, belief and plausibility measurements represent lower and upper bounds which surround the certain belief value. In other words, the Belief Interval (BI) is an interval between belief and plausibility functions representing range in which exact belief resides. A narrow belief interval represents more precise beliefs. It can be shown that the belief is uniquely determined if $Bel(A) = Pls(A)$. If $BI(A)$ has an interval $[0,1]$, it means that no information about hypotheses is available; on the other hand, if the interval is $[1,1]$, it means that A has been completely confirmed.

Therefore, the degree of uncertainty which is the difference between belief and plausibility functions becomes smaller as we obtain more information and knowledge. Even though evidence theory does not give us a single value, the given belief interval $[Bel, Pls]$ retains all the information without any excessive and baseless assumptions. That is, the result of evidence theory is consistent with given partial truth, and therefore leads us to approximate reasoning. Since the belief interval represents the current uncertainty situation based on available evidence, a decision maker can obtain insight into the problem and avoid mistakes made by misusing assumptions [25].

Evidential Distance Measures

Several works could be found in the literature which tried to propose distance measures through the evidence theory framework [11,12,14-16,18]. In the Evidential Distance-based Classifier (EDC) proposed by Denoeux [12], k-NN classification rule is addressed from the point of view of Dempster-Shafer theory. This evidence-theoretic k-NN rule considers each neighbor of a sample to be classified as an item of evidence that supports certain hypotheses regarding the class membership of that sample. The degree of support is defined as a function of the distance between the two feature vectors. Also, a learning algorithm was proposed by Zouhal and Denoeux [26] for determining the parameters of the proposed approach equation to reduce the classification errors. Denoeux argued that unknown samples can also be classified according to their distances to a limited number of reference patterns through multi-layer neural network to reduce computation time [13]. A characteristic of this approach resides in the possibility of taking into account partial knowledge of the class of training samples [27]. In this way, a Euclidean measure for the evidence theory has been proposed as follows:

$$d(m(A), m(B)) = \sum_{A \in F_i} \sum_{B \in F_j} m_i(A) m_j(B), \quad (10)$$

where F_i is the set of all focal elements of m_i . Another evidential measure has been represented by

Tessem [18]: The error due to an approximation of BBAs is quantified by the maximal deviation in the pignistic probabilities before and after approximation, so as we have:

$$d(i, j) = \max_{\theta_i} |BetP_i(\theta_i) - BetP_j(\theta_i)|, \quad (11)$$

where $BetP_i$ is the pignistic probability corresponding to BBA m_i .

Bauer [11], on the same problem, introduced two other measures of error to reflect the quality of a decision based on the pignistic probability distribution after approximation. The idea to define a measure on the power set Θ has been used by Petit-Renaud, where an error criterion between two belief structures based on the generalized Hausdorff distance has been defined [15].

Also, Denoeux introduced the Interval-valued Belief Structures (IBS) allowing to model the situation in which hypotheses may only be expressed with some imprecision [28]. Actually, IBS copes with imprecision in the belief value itself attached to each hypothesis which is only known to lie within a certain interval. Using the theory of fuzzy sets as a very efficient tool for representing the vague quantities, belief values could also be expressed as fuzzy numbers, so called fuzzy-valued belief structures [29].

Jousselme et al. [15] apply a classical similarity measure to achieve the comparison of the focal elements of two BBAs, in order to define a distance in a vector space generated by the focal elements. This distance is then used as a measure of performance of different algorithms and helps evaluate how fast they converge to the desired known solution. A BBA $_i$ can be seen as a vector \vec{m}_i in the mentioned vector space and the normalized distance is a function:

$$d(\vec{m}_1, \vec{m}_2) = \sqrt{\frac{1}{2}(\vec{m}_1, \vec{m}_2)^T \underline{S}(\vec{m}_1, \vec{m}_2)}, \quad (12)$$

where \underline{S} is a similarity matrix between the focal elements filled by the application of the Jaccard measure between the corresponding subsets. Also, another way to write Equation 12 is:

$$d(\vec{m}_1, \vec{m}_2) = \sqrt{\frac{1}{2} \left(\|\vec{m}_1\|^2 + \|\vec{m}_2\|^2 - 2 \langle \vec{m}_1, \vec{m}_2 \rangle \right)}, \quad (13)$$

where $\langle \vec{m}_1, \vec{m}_2 \rangle$ is the scalar product defined by:

$$\langle \vec{m}_1, \vec{m}_2 \rangle = \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} m_1(A_i) m_2(A_j) \frac{|A_i \cap A_j|}{|A_i \cup A_j|}, \quad (14)$$

where $A_i, A_j \in P(\Theta)$ for $i, j = 1, \dots, 2^N$. $\|\vec{m}\|^2$ is then the square norm of \vec{m} :

$$\|\vec{m}\|^2 = \langle \vec{m}, \vec{m} \rangle. \quad (15)$$

A popular approach has been to take the conflict of two Basic Belief Assignments (BBAs) as an indication of the belief that they do not belong to a same class, and then assume it as a measure of dissimilarity or distance between them. Schubert [17] presents a clustering of BBAs based on a measure of conflict, resulting from the aggregation of the BBAs members of the cluster using Dempster's rule of combination. This measure gives an indication of how likely it is that they belong to the same class or not. The conflicts from each cluster are then combined to find the overall meta-conflict of the partition. The clustering algorithm minimizes the metaconflict to obtain the most credible partition of the pieces of evidence. Once the clusters are defined, a BBA prototype is found among members of each cluster by analyzing the evolution of the cluster's conflict when removing each BBA. These prototypes can be then used to classify new samples to the cluster whose prototype has the smaller conflict with.

Also, another distance measure is Bhattacharya distance [14,16] which is defined in probability theory as:

$$d_B(p_1, p_2) = -\ln \left(\int_x \sqrt{p_1(x)p_2(x)} dx \right), \quad (16)$$

where p_1, p_2 represent the probability density functions of vector x of any dimension. This measure has the advantage that is designed to compare features for the two classes. It is a special case of the Chernoff bound of the error probability in binary classification [30]. It is well known that the Chernoff information gives the highest achievable exponent for the error probability. The Bhattacharya distance has the symmetric property, ($d(p_1, p_2) = d(p_2, p_1)$). The triangle property is only satisfied for specific configurations. Also, the discrete expression of Equation 16 is:

$$d_B(h_1, h_2) = -\ln \left(\sum_i \sqrt{h_1(i)h_2(i)} \right), \quad (17)$$

where i is an index of the bins of the normalized histograms h_1 and h_2 . Also, Bhattacharya distance definition in evidence theory [16] is as follows:

$$d(m_i, m_j) = \sqrt{1 - \left(\sum_{A \in F_i} \sum_{B \in F_j} \sqrt{m_i(A)m_j(B)} \right)}, \quad (18)$$

where F_i is the set of all focal elements of m_i . Diaz et al. [14] proposed a similarity measure between the focal elements, making way for the application of classical classification algorithms in this field. The properties of this measure are particular to its context, considering the characteristics of the focal elements,

their relationship with each other and their proximity to the vacuous belief function that represents the state of total ignorance.

Another evidential measure has been proposed by Ristic and Smets [16] which derives the global cost of assignment (i.e. a dissimilarity measure) based on the plausibility of the global assignment. This measure is directly related to the conflict as described in the TBM. The plausibility of the global assignment V is given by:

$$Pl^{\Theta^{2n}}(\theta(o_{1i}) = \theta(o_{2\nu i}) : i = 1, 2, \dots, n) = \prod_{i=1}^n 1 - m_{1i \oplus 2\nu i}^{\Theta}(\phi), \quad (19)$$

where o_1 and o_2 represent bodies of evidence, ϕ the null set, and \oplus combination of the basic belief assignments or the global assignment. Also, they defined an additive global dissimilarity measure of an assignment as follows:

$$D = -\log \left[Pl^{\Theta^{2n}}(\theta(o_{1i}) = \theta(o_{2\nu i})) : i = 1, 2, \dots, n \right] = \sum_{i=1}^n d_{i\nu i}, \quad (20)$$

where:

$$d_{ij} = -\log(1 - m_{1i \oplus 2j}^{\Theta}(\phi)). \quad (21)$$

A NEW EVIDENTIAL DISTANCE MEASURE BASED ON THE BELIEF INTERVALS

In the real world problems, usually several features, components or criteria influence the problem, and play important roles in proposing a solution, whereas each of them may be uncertain [31]. For instance, a cardiologist has to consider several medical markers to assess the Coronary Heart Disease (CHD) risk in the patients, whereas each marker indicates CHD occurrence risk with uncertainty. Therefore, it is more intuitive and reasonable to exploit the evidence theory and construct a belief interval through determining the belief and plausibility functions.

In this paper, we propose a new evidential distance measure based on belief intervals. This measure presents the distance between belief intervals associated to the imprecise unknown patterns and classes. In other words, this measure is based on the difference between the analogous features' belief intervals which are assigned to unknown patterns and classes. Hereafter, we refer to this measure by the name of Belief Interval Distance (*BID*). In doing so, we develop an approach of approximate reasoning in which uncertain feature vectors are classified according to

existing knowledge and partial truth while considering problem uncertainty.

A convenient geometrical representation of the belief interval can be employed to illustrate this distance. As mentioned before, all the propositions are surrounded in the frame of discernment and after assigning BBAs, we quantify problem uncertainty through belief and plausibility functions which leads to the belief interval. This process could be expressed through a geometrical interpretation. We can assume a frame of discernment M and subset N in the Euclidean plane with the Cartesian coordinates. For a given belief interval, BI , a function f from M to N can be constructed, such that if $x \in M$, then:

$$p = f(x) \in N,$$

and the point $p \in N$ has the coordinates $x \in M \langle a, b \rangle$ for which:

$$0 \leq a, \quad b \leq 1,$$

where:

$$a = Bel(x) \quad \text{and} \quad b = Pls(x).$$

In fact, function f quantifies the problem uncertainty and assigns the appropriate values to the belief and plausibility functions, in other words, constructs the belief interval. As illustrated in Figure 2, this geometrical interpretation maps the frame of discernment as a polygon to the belief interval territory represented in a triangle area.

Now, we suggest an exact form of the belief interval which comprises of belief and plausibility functions, beside of uncertainty quantity which represents the status of problem uncertainty in each situation. Despite belief and plausibility functions could be used to reflect the uncertainty, none of them could express the uncertainty alone. Besides, we found that using three coordinates lead to better results in pattern recognition. Thus, the exact belief interval has three coordinates as follows: $\langle Bel, Pls, Uncer \rangle$. Hence, the most natural representation of a belief interval is

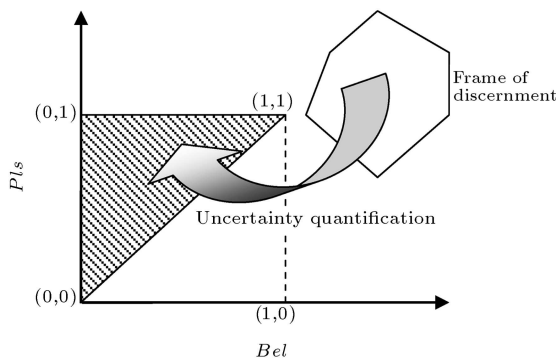


Figure 2. A geometrical interpretation of a belief interval.

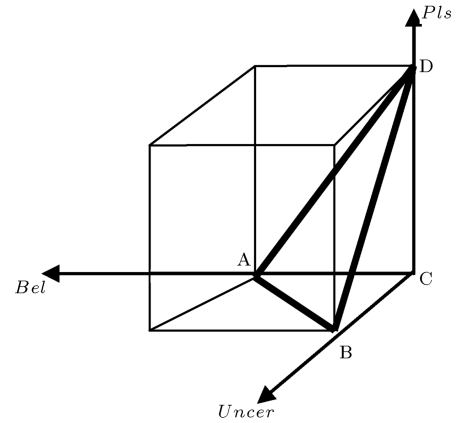


Figure 3. A three-dimensional representation of a belief interval.

to draw a cube with edge length equal to 1, as shown in Figure 3, so as triangle ABD represents a typical belief interval. This representation of a belief interval is used to investigate the belief interval distance.

To lighten up this geometrical interpretation, let us consider a conceptual instance which concerns comparing the opinions of several experts regarding a hypothesis. These experts described their beliefs concerning different perspectives of the hypothesis, and after uncertainty quantifications based on their beliefs, the belief intervals for each of them were formed. Each expert i is represented as a point having coordinates $\langle Bel, Pls, Uncer \rangle$. Expert $C: \langle 1, 1, 0 \rangle$ fully accepts the discussed hypothesis. Expert $A: \langle 0, 0, 0 \rangle$ fully rejects it. The experts on the segment AC fixed their points of view. In fact, their uncertainty measure equal zero, so that each expert is convinced to the belief Bel_i . Also, expert $B: \langle 0, 1, 1 \rangle$ is absolutely uncertain, whereas expert $C: \langle 1, 1, 0 \rangle$ is absolutely confirmed that the hypothesis is true and expert $A: \langle 0, 0, 1 \rangle$ believes it is absolutely false. An orthogonal real representation of a belief interval has been show in Figure 4.

A line parallel to AC describes a set of experts with the same level of uncertainty. For instance, two sets are presented with uncertainty level equal to U_1

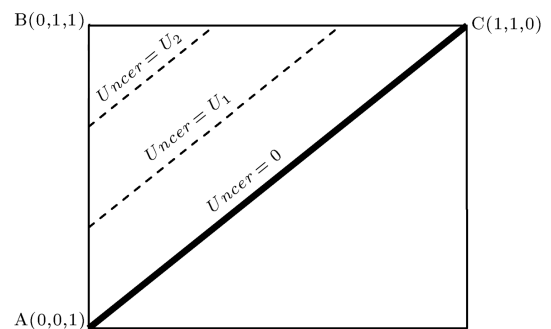


Figure 4. An orthogonal projection of the real representation of a belief interval.

and U_2 where $U_2 > U_1$. In other words, the triangle ABC in Figure 4 is an orthogonal projection of the real situation, i.e. the triangle ABD in Figure 3.

Now, the belief interval distance is proposed using the Hamming distance measure. Instead of using 1-D Hamming distance to measure the difference between two analogous crisp values, the differences between two intervals besides their uncertainty quantities have been measured through 3-D Hamming distance, so that the differences between two belief functions, two plausibility functions, and also two uncertainty quantities of both intervals are computed analogously. To formulate the distance, we depict it through the following definition:

Let A and B be two belief intervals associated to a class and unknown pattern, respectively. The Belief Interval Distance (BID) is defined as follows:

$$BID(A, B) = \frac{1}{2} \sum_{A_i, B_i} \left(\begin{array}{l} |Bel(A_i) - Bel(B_i)| + \\ |Pls(A_i) - Pls(B_i)| + \\ |Uncer(A_i) - Uncer(B_i)| \end{array} \right), \quad (22)$$

where A_i and B_i denote components (such as features) of A and B , respectively. This definition satisfies the appropriate properties of a distance metric which is proved in following.

Proof

The belief interval represents the upper and lower bounds which surround the certain belief value. Essentially, if we could not represent the belief associated to a proposition as a certain value, it is more intuitive to express the belief through an interval. Hence, the belief interval implies a certain belief value. With attention to this point and its geometric interpretation, measuring the distance between two belief intervals could be considered as the distance between two values which are represented through intervals.

The proposed evidential distance must satisfy four properties to be a distance metric and show its validity. The proposed distance is a function that satisfies the following properties for any belief intervals A and B :

I. Non-negativity:

$$BID(A, B) \geq 0.$$

Because:

$$|A| \geq 0 \rightarrow BID(A, B) \geq 0.$$

II. Nondegeneracy:

$$BID(A, B) = 0 \Leftrightarrow A = B,$$

$$A = B \Leftrightarrow Bel(A) = Bel(B),$$

$$Pls(A) = Pls(B),$$

and:

$$Uncer(A) = Uncer(B).$$

So:

$$\begin{aligned} BID(A, B) &= \frac{1}{2} \left(|Bel(A) - Bel(B)| + |Pls(A) - Pls(B)| + |Uncer(A) - Uncer(B)| \right) \\ &= 0, \end{aligned}$$

and vice versa.

III. Symmetry:

$$BID(A, B) = BID(B, A),$$

$$BID(A, B) = \frac{1}{2} \left(|Bel(A) - Bel(B)| + |Pls(A) - Pls(B)| + |Uncer(A) - Uncer(B)| \right)$$

$$= \frac{1}{2} \left(|Bel(B) - Bel(A)| + |Pls(B) - Pls(A)| + |Uncer(B) - Uncer(A)| \right)$$

$$= BID(B, A).$$

IV. Triangle inequality:

$$BID(A, B) \leq BID(A, C) + BID(C, B),$$

$$\forall C \in [0, 1],$$

because of:

$$|A \pm B| \leq |A| + |B|,$$

$$BID(A, C) + BID(C, B) = \frac{1}{2} \left(\begin{array}{l} |Bel(A) - Bel(C)| + |Pls(A) - Pls(C)| \\ + |Uncer(A) - Uncer(C)| \\ + |Bel(C) - Bel(B)| + |Pls(C) - Pls(B)| \\ + |Uncer(C) - Uncer(B)| \end{array} \right)$$

$$\geq \frac{1}{2} \left(\begin{array}{l} |Bel(A) - Bel(C) - (Bel(C) - Bel(B))| + \\ |Pls(A) - Pls(C) - (Pls(C) - Pls(B))| + \\ |Uncer(A) - Uncer(C) - (Uncer(C) - Uncer(B))| \end{array} \right)$$

$$= \frac{1}{2} \left(|Bel(A) + Bel(B)| + |Pls(A) + Pls(B)| + |Uncer(A) + Uncer(B)| \right)$$

$$\geq \frac{1}{2} \left(|Bel(A) - Bel(B)| + |Pls(A) - Pls(B)| + |Uncer(A) - Uncer(B)| \right)$$

$$= BID(A, B).$$

So, we have:

$$BID(A, B) \leq BID(A, C) + BID(C, B).$$

Example 1

Let us consider the belief intervals A, B, C, D and E as follows:

$$\begin{aligned} A &= (0, 0), & B &= (1/3, 1/3), \\ C &= (1/2, 2/3), & D &= (2/3, 4/5), & E &= (1, 1). \end{aligned}$$

The geometrical interpretation of these belief intervals is shown in Figure 5.

Let us calculate the BID distance between the belief intervals based on Equation 22:

$$\begin{aligned} BID(B, E) &= \frac{1}{2} \left(\left| \frac{1}{3} - 1 \right| + \left| \frac{1}{3} - 1 \right| + |0 - 0| \right) = \frac{2}{3}, \\ BID(C, E) &= \frac{1}{2} \left(\left| \frac{1}{2} - 1 \right| + \left| \frac{2}{3} - 1 \right| + \left| \frac{1}{6} - 0 \right| \right) = \frac{1}{2}, \\ BID(D, E) &= \frac{1}{2} \left(\left| \frac{2}{3} - 1 \right| + \left| \frac{4}{5} - 1 \right| + \left| \frac{2}{15} - 0 \right| \right) = \frac{1}{3}. \end{aligned}$$

As we expected intuitively, these distances are describing the real differences between the belief intervals. For instance, the distance between B and E is greater than the one between D and E , as Figure 5 implies.

NUMERICAL EXPERIMENTS

In this section, we apply the proposed evidential distance measure in two medical diagnosis problems to examine its capability in pattern recognition. Also, the acquired results are compared with the other evidential measures. For this purpose, the performance of the proposed method is evaluated using classification

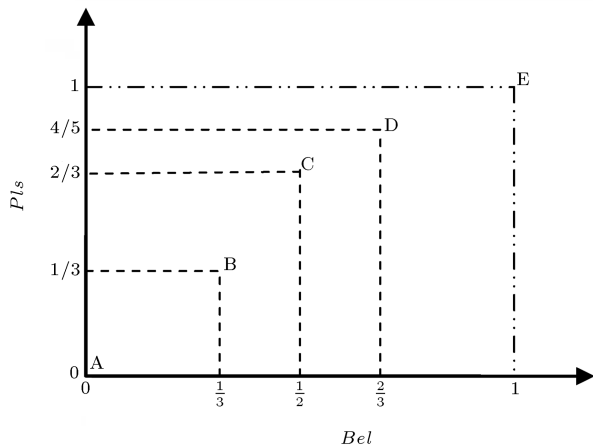


Figure 5. A geometrical interpretation of the belief intervals considered in Example 1.

accuracy, sensitivity and specificity [32], so as their representations in medical diagnosis are:

$$\text{Accuracy} = \frac{\# \text{ true positives} + \# \text{ true negative}}{\# \text{ all patients}}, \tag{23}$$

$$\text{Sensitivity} = \frac{\# \text{ true positives}}{\# \text{ all patients with the disease}}, \tag{24}$$

$$\text{Specificity} = \frac{\# \text{ true negatives}}{\# \text{ all patients without the disease}}. \tag{25}$$

The true positives are all patients with the disease and positive test result, whereas the true negatives are all patients without the disease and negative test result. We use the accuracy, sensitivity and specificity to perform ROC analysis. ROC (Receiver Operating Characteristic) graphs have long been used in signal detection theory to depict tradeoffs between hit rate (sensitivity) and false alarm rate (1-specificity) [32]. ROC analysis has lately been extended for use in visualizing and analyzing the behavior of diagnostic systems, and is used for visualization in medicine, where specificity-sensitivity relations are often analyzed. In the usual setting, the machine learning algorithms are tuned to maximize classification accuracy. In our case, the sensitivity and specificity are more important. The clinicians especially wanted to see if it is possible to increase the specificity of the diagnostic process without affecting the sensitivity too much (this may lead to the reduction of number of patients that are being unnecessarily submitted to invasive preoperative examinations). It must be noted that the ROC analysis is only used in two-class classifications, so as multiclass problems cannot be analyzed through it. In analyzing a classifier through ROC graph, we should note that by giving samples from the negative class higher scores, a good classifier should become more specific and less sensitive. On the other hand, by giving examples from the positive class higher scores, a good classifier should become more sensitive and less specific. In fact, a good classifier ROC curve takes place in upper left corner of chart, so as the more upper-left the ROC graph, the more accurate the recognition rate.

BID Application in Bacillus Colony Recognition

Microbiologists broadly classify bacteria according to their shapes [33-35]. Most bacteria come in one of three shapes: rod, sphere or spiral. Rod-shaped bacteria are called bacilli, Spherical bacteria, cocci, and spiral or corkscrew-shaped bacteria are called spirilla, too. Bacteria may be further classified according to whether they require oxygen (aerobic or anaerobic) and how they react to a test with Gram’s stain. Bacteria in

which alcohol washes away Gram's stain are called Gram-negative, while bacteria in which alcohol cause the bacteria's wall to absorb the stain, are called Gram-positive [33,34]. As noted before, bacillus is a rod-shaped bacterium which is active only in the presence of the oxygen (aerobic bacterium). Bacilli occur mainly in chains, produce spores, and include many saprophytes, some parasites and the bacterium that causes anthrax [36]. In this research, four intestinal bacilli named "*Shigella*, *Salmonella*, *Bacillus coli* and *Klebsiella*" are considered which have some similarity in culture medium and are Gram-negative.

Shigella [36] is a rod-shaped Gram-negative bacterium (bacillus) that lives in the intestinal tracts of human beings and animals and causes bacillary dysentery. There are four species, all causing dysentery but with varying degrees of severity. *Salmonella* [36] is also a rod-shaped bacterium found in the intestine that can cause food poisoning, gastroenteritis and typhoid fever. Other two bacilli considered herein *Klebsiella* and *Bacillus coli* are also Gram-negative and found in the intestine. Primal features used for these bacilli classification comprise of macro shape which is domical in these four bacillus colonies, single microscopic shape, double microscopic shape and existence of flagellum. Our target is to use the proposed evidential distance to recognize the bacillus type of the samples.

To confront this problem using *BID* measure, the next steps were followed. First, a data set consisted of 283 samples of different bacillus colonies were obtained from microbiology section of Resalat

Laboratory in Tehran, Iran. Then, the basic belief assignment formulation has been performed by the medical experts, so as we rely mainly on the knowledge and expertise of the microbiologists. Consequently, the belief and plausibility functions were acquired. Hence, the bacillus patterns' features were expressed in the belief intervals. Second, we pursued to determine each class's feature vector which depict the relationship between the bacillus type and four features, as shown in Table 1. In fact, the classes are constructed based on the beliefs acquired from the microbiologists' medical knowledge and expertise.

To classify each unknown sample, we used *BID* distance. In fact, a sample's distances opposing four classes were calculated using *BID* distance measure. Having compared four *BID* distance values acquired, according to nearest neighborhood concept, the class which has the smallest *BID* distance to the unknown sample is chosen as the appropriate class for the sample. To review the recognition process, an example is denoted. According to the four bacillus colonies patterns exhibited in Table 1, the following classes are depicted:

$$\begin{aligned} \textit{Bacillus coli} &= \\ &\{(0.9, 0.95, 0.05), (0.9, 1, 0.1), (0, 0, 0), (0.9, 0.94, 0.04)\}, \\ \textit{Shigella} &= \\ &\{(0.9, 0.92, 0.02), (0.9, 0.95, 0.05), (0.05, 0.08, 0.3), \\ &(0.08, 0.1, 0.02)\}, \end{aligned}$$

Table 1. Imprecise feature vectors of four classes (patterns).

Bacillus class	Feature	$Bel(A_i)$	$Uncer(A_i)$	$Pls(A_i)$
<i>Bacillus coli</i>	Domical shape	0.9	0.05	0.95
	Single microscopic shape	0.9	0.1	1
	Double microscopic shape	0	0	0
	Flagellum	0.9	0.04	0.94
<i>Shigella</i>	Domical shape	0.9	0.02	0.92
	Single microscopic shape	0.9	0.05	0.95
	Double microscopic shape	0.05	0.03	0.08
	Flagellum	0.08	0.02	0.1
<i>Salmonella</i>	Domical shape	0.8	0.1	0.9
	Single microscopic shape	0.8	0.1	0.9
	Double microscopic shape	0.1	0.05	0.15
	Flagellum	0.9	0.09	0.99
<i>Klebsiella</i>	Domical shape	0.8	0.05	0.85
	Single microscopic shape	0.7	0.15	0.85
	Double microscopic shape	0.2	0.05	0.25
	Flagellum	0.1	0.05	0.15

Table 2. Classification result using BID distance versus microbiologists opinions.

S. No.	<i>Bacillus coli</i>	<i>Shigella</i>	<i>Salmonella</i>	<i>Klebsiella</i>	Microbiologists
S_1	0.306	1.247	0.357	1.357	<i>Bacillus coli</i>
S_2	0.520	1.293	0.405	1.273	<i>Salmonella</i>
S_3	1.444	0.575	1.444	0.568	<i>Shigella</i>
S_4	1.421	0.558	1.250	0.274	<i>Klebsiella</i>
S_5	0.282	0.986	0.350	1.177	<i>Bacillus coli</i>
S_6	1.311	0.563	1.227	0.655	<i>Shigella</i>
S_7	0.413	1.086	0.408	1.242	<i>Bacillus coli</i>
S_8	1.500	0.597	1.242	0.310	<i>Klebsiella</i>
S_9	1.166	0.336	1.15	0.470	<i>Shigella</i>
S_{10}	0.354	1.144	0.401	1.414	<i>Bacillus coli</i>

Salmonella =

$$\{(0.8, 0.9, 0.1), (0.8, 0.9, 0.1), (0.1, 0.15, 0.05), (0.9, 0.99, 0.09)\}, \text{ and}$$

Klebsiella =

$$\{(0.8, 0.85, 0.05), (0.7, 0.85, 0.15), (0.2, 0.25, 0.05), (0.1, 0.15, 0.05)\}.$$

If we consider a sample $A = \{(0.8, 0.9, 0.1), (0.95, 1, 0.05), (0, 0, 0), (1, 1, 0)\}$, based on Equation 22, we have:

$$BID(A, Shigella)$$

$$= \frac{1}{2} \left\{ \begin{array}{l} |0.8 - 0.9| + |0.1 - 0.02| + \\ |0.9 - 0.92| + |0.95 - 0.9| + \\ |0.05 - 0.05| + |1 - 0.95| + \\ |0 - 0.05| + |0 - 0.03| + \\ |0 - 0.01| + |1 - 0.08| + \\ |0 - 0.02| + |1 - 0.1| \end{array} \right\} = 1.15.$$

Also, For other three classes, similar calculations are performed which lead to:

$$BID(Bacillus coli, A) = 0.25,$$

$$BID(Salmonella, A) = 0.4,$$

$$BID(Klebsiella, A) = 1.45.$$

From the above results, it is evident that sample A has the less distance or the most similarity to *Bacillus coli* pattern, and therefore is classified as a *Bacillus coli* colony.

After testing phase completion, all the unknown patterns were classified. The final results of classification have been checked with microbiologists' opinions.

Table 2 summarizes the classification result of ten samples which are compared with the microbiologists' opinions, where the shaded cells indicate the appropriate class of the sample based on the BID distance measure.

Having compared all the classified samples with medical experts' ideas, we observed that 91.66 percent of the unknown samples have been classified correctly. For the current research data set, the following numbers of cases have been identified correctly through BID evidential distance: *Bacillus coli* $\frac{34}{40}$, *Shigella* $\frac{36}{40}$, *Salmonella* $\frac{20}{20}$ and *Klebsiella* $\frac{20}{20}$. Also, the error intervals of misclassified unknown patterns of *Bacillus coli* and *Shigella* classes reside in the following intervals, respectively: $[0.01, 0.18]$ and $[0.01, 0.09]$. The bacillus recognition using the BID distance for a number of samples is represented in Figure 6 in which the most similar class has obtained a lower rank, meanwhile compared with the microbiologists' opinions.

As shown in Figure 6, the belief distance measure has been applied for each of the samples, so as its similarity to each of the classes has been assessed via

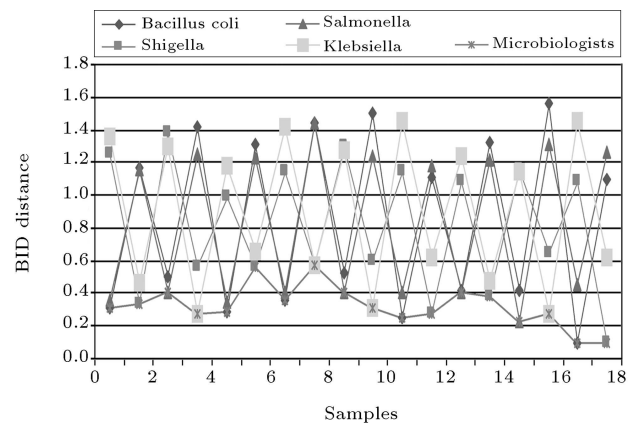


Figure 6. Bacillus colonies recognition using BID measure compared with microbiologists' opinions.

Table 3. Comparison between BID measure and other evidential measures in bacillus colonies recognition.

Measure	BID	Bahatacharya	Euclidean	Tessem	Jousellme et al.	Ristic and Smets
Accuracy rate	91.66%	48.33%	16.66%	76.66%	85.00%	16.60%

determining their distances to the classes. In this way, acquiring the least distance to a class means that the sample belongs to that class. Also, the microbiologists' opinions are represented in this figure which shows the high conformance of the proposed measure with real classifications. The other evidential distance measures were applied to the bacillus colonies recognition. Jousellme et al.'s measure [15] could distinguish 85 percent of the samples properly, while Bahatacharya [14,16], Euclidean [27] and Tessem [18] evidential distance measures provided us with 48.33%, 16.66% and 76.66% accuracy rates in recognition. Also, Ristic and Smets measure [16] yielded weak recognition rate of 16.6% in this problem. Table 3 compares the BID distance measure result with the other evidential measures results in bacillus colonies recognition.

BID Application in Coronary Heart Disease Patients Classification

Coronary heart disease is an important disease in heart medicine, so as we observe proliferation in CHD occurrence in humanity, nowadays, because of machine presence in various aspects of human being life which is along with obesity and psychological diseases such as depression. Annually, twelve million people around the world die because of heart disease and it is predicted that cause of 75 percent of mortalities in 2020 will be these diseases [37]. Recently it is announced that 317 people die in Iran daily because of coronary and heart diseases among which 47% are young [38].

Coronary heart disease is a chronic disease in which the coronary arteries, responsible for transporting oxygenized blood to heart muscles, get narrowed and are not able to convey enough fresh blood to this bloodpumping organ [39]. Often, narrowing of blood vessels is because of arteriosclerosis that is a common arterial disease in which raised areas of degeneration and cholesterol deposit plaques form on the inner surfaces of the arteries obstructing blood flow. When the blood supply of the heart is reduced, it does not receive enough oxygen and nutrition to operate properly. This oxygen shortage leads to two important and critical results: angina pectoris and heart attack [40].

CHD patients classification assists the physicians to identify the patients with low and high CHD risk, and in this way control CHD progress in the appropriate patients [41]. One of the best and common ways of the CHD patients classification is to investigate CHD medical markers or risk factors [42]. In fact, physicians through studying the CHD medical markers

in their medical research can diagnose this disease before being acute. Most important CHD medical markers are total cholesterol, systolic and diastolic blood pressures, highdensity cholesterol, age, smoking and diabetes mellitus [42]. In this section, we want to apply the BID distance measure to classify the patients into two classes: healthy and unhealthy, in association to CHD.

After a survey study on the heart and CHD literature, it was found that various researches are done on CHD risk factors and medical markers, so as they could determine each marker's effect on the CHD occurrence. With reviewing these sources and interviewing with heart disease experts, the key medical markers affecting CHD were obtained [42]. After studying heart and CHD literature, importance and affection degree of the medical markers on CHD occurrence were acquired. It is worthy to note that none of the markers imply the CHD occurrence certainly and we observe uncertainty in the relation between the medical markers and CHD occurrence. To represent the relations between the markers and CHD risk, Framingham heart study [42] has been used, so as 28 association rules were extracted which models the uncertainty in the CHD occurrence, according to the respective markers, as shown in Table 4.

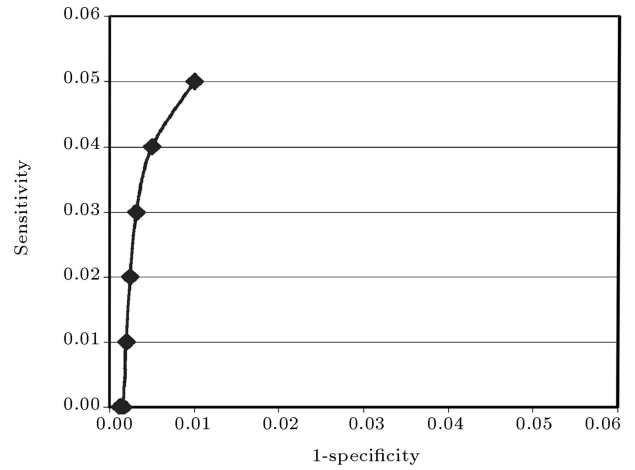
Taking into account the base patterns or classes for healthy and unhealthy people in association to coronary heart disease, the unknown patterns are classified using evidential distance measures. For this purpose, we acquired the feature vectors of healthy and unhealthy people from Framingham heart study [42]. Then, the BBAs for the unknown feature vectors of patients were assigned according to Table 4. Having applied the different evidential distance measures to the unknown patterns brought us their respective distances to healthy and unhealthy classes which were used as criteria to classify people in the classes, so as less distance to a class implies assignment to that class. Having applied the BID distance measure to the Hungarian institute of cardiology's heart disease data set in the university of California, Irvine's machine learning repository [43] comprised of 294 samples, the acquired CHD patients classification results were compared with the real diagnoses, so as it was found that the proposed measure could classify the samples with 92.54% accuracy rate.

Also, Jousellme et al. measure [15] yielded a classification with the same accuracy rate, while Bahatacharya [14,16] provided us the weak recognition rate of 29.24%. Besides, Euclidean [27], Tessem [18]

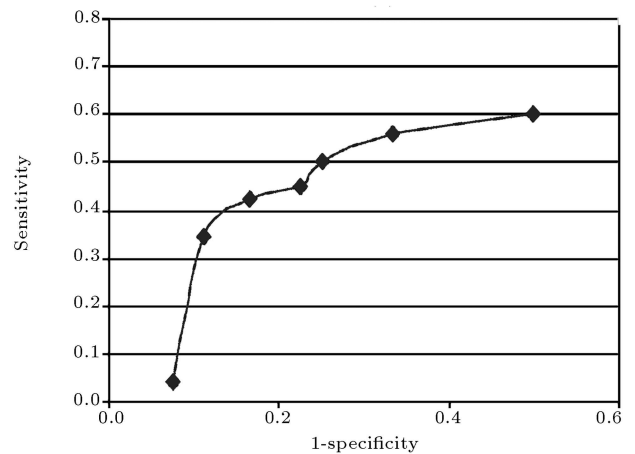
Table 4. Coronary heart disease medical markers and their basic belief assignments in CHD occurrence.

Medical Markers	Reference Range	CHD Occurrence BBA	
Age	30-34	0	
	35-39	0	
	40-44	0.02	
	45-49	0.05	
	50-54	0.09	
	55-59	0.13	
	60-64	0.18	
	65-69	0.23	
	70-74	0.27	
Total cholesterol	<160	0	
	160-199	0	
	200-239	0.03	
	240-279	0.07	
	>280	0.12	
Smoking	No	0	
	Yes	0.09	
High density cholesterol (HDL)	<35	0.08	
	35-44	0.04	
	45-49	0.01	
	50-59	0	
Diabetes mellitus	No	0	
	Yes	0.07	
Blood pressure	Systolic	Diastolic	
	<120	<80	0
	120-129	80-84	0
	130-139	85-89	0.03
	140-159	90-99	0.05
	>160	>100	0.11

and Ristic and Smets [16] evidential distance measures could distinguish 91.5%, 88.67% and 91.5% of the samples correctly, respectively. Since this problem is twoclass classification, we could apply ROC analysis. For this purpose, sensitivity and specificity of the BID measure were acquired, as shown in Figure 7. As this figure shows, the BID's ROC graph takes place in the upper left corner that implies its great capability in classification, while Tessem measure ROC graph is nearly in middle of the chart. Also, ROC graph of Jousellme et al. measure is nearly the same as BID measure which implies its good capability in classification too. Table 5 compares the BID distance



(a)



(b)

Figure 7. ROC analysis of evidential distance measures. a) BID distance measure; b) Tessem distance measure.

measure result with the other evidential measures in coronary heart disease patients classification.

Also, the BID measure experimental results in CHD patients classification have been compared with the similar researches, as shown in Table 6, which implies the proposed measure encounters the problem more efficiently. In one of these researches, various machine learning methods such as Bayesian classification and neural networks have been applied to the heart diseases diagnoses [32], so as the accuracy rate of 80% has been reported for coronary heart diseases diagnoses based on the appropriate medical markers. In another research, neural networks have been used for recognizing the coronary heart diseases, so as the designed neural network could identify 84% of the CHD

Table 5. Comparison between BID measure and other evidential measures in CHD patients classification.

Measure	BID	Bahatacharya	Euclidean	Tessem	Jousellme et al.	Ristic and Smets
Accuracy rate	92.45%	29.24%	91.5%	88.67%	92.45%	91.50%

Table 6. Comparison between the proposed measure outcome and the similar researches.

No.	Research	Methods	Accuracy Rate (%)
1	Kukar et al. [32]	Bayesian classification and neural network	80
2	Akay [44]	Neural network	84
3	Haddad et al. [45]	Neural network	48
4	Detrano et al. [46]	Probability theory (logistic regression)	77
5	Proposed measure	Evidence theory	92.45

occurrence in the patients [44]. Also, a research has been exploited the heart rate study using the neural networks which only could identify 48% of the disease in the patients [45]. Besides, the acquired result of the proposed evidential distance measure has been compared with the similar researches applied to the same data set among which we can mention Detrano et al. [46] and Gennari et al. [47] that yielded 77 and 78.9 percent accuracy rates, respectively.

CONCLUSION

The evidential measures proposed so far rely on the basic belief assignment function and do not exploit the belief intervals directly, whereas the belief interval could depict all the knowledge available regarding a pattern's feature. This paper proposed a new evidential distance measure based on the belief interval. Having obtained the belief and plausibility functions or the belief interval, the distance between the belief intervals of uncertain classes' pattern feature vectors and unknown patterns are calculated, so as the classification could be take place according to nearest neighborhood concept.

This evidential distance measure can be used in classification for imprecise and uncertain feature vectors with a realistic look at problem uncertainty. In the proposed approach, we first construct the belief intervals, and then examine the similarity between an unknown pattern and the classes based on the belief interval distance measure. As practical applications of proposed distance measure in classification, it was applied to the bacillus colony recognition and coronary heart disease patients classification. Experimental results showed that the proposed distance provided us with 91.66 and 92.54 percent accuracy rates in these applications, respectively. Comparison the acquired results with other evidential measures results depicts the proposed measure superior performance.

ACKNOWLEDGMENTS

The authors would like to thank Ms. Hoda Davarzani, the Ph.D. student at the Tarbiat Modares University, Tehran, Iran for her precious assistance in data gathering of their research. This work was supported by the Iran Telecommunications Research Center (ITRC), Tehran, Iran, under grant No. TMU 87-07-48.

REFERENCES

1. Han, J. and Kamber, M., *Data Mining: Concepts and Techniques*, Elsevier, Morgan Kaufmann, San Mateo, CA (2006).
2. Bishop, C.M., *Neural Nnetworks for Pattern Recognition*, Oxford University Press, Oxford, England (1995).
3. Devijer, P.A. and Kittler, J., *Pattern Recognition, a Statistical Approach*, Prentice-Hall, London, England (1982).
4. Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis*, John Wiley, New York (2001).
5. Haji, M.M. and Katebi, S.D. "Machine learning approaches to text segmentation", *Scientia Iranica*, **13**(4), pp. 31-42 (2006).
6. Yazdchi, M.R. et al. "A new bidirectional neural network for lexical modeling and speech recognition improvement", *Scientia Iranica*, **14**(6), pp. 65-73 (2007).
7. Khatibi, V. and Montazer, G.A. "Intuitionistic fuzzy set application in bacteria detection", *14th Int'l Computer Conference of Iranian Computer Society (CS-ICC2009)*, Tehran, Iran (2009).
8. Khatibi, V. and Montazer, G.A. "Bacillus colonies recognition using intuitionistic fuzzy sets", *Notes on IFS*, **14**(2), pp. 91-99 (2008).
9. Gitizadeh, M. and Kalantar, M. "Genetic algorithm based fuzzy multi-objective approach to FACTS devices allocation in FARS regional electric network", *Scientia Iranica*, **15**(6), pp. 35-42 (2008).
10. Montazer, G.A. et al. "A neurofuzzy inference engine for farsi numeral characters recognition", *Expert Systems with Applications*, **37**(9), pp. 6327-6337 (2010).
11. Bauer, M. "Approximation algorithms and decision making in the Dempster-Shafer's theory of evidence - an empirical study", *International Journal of Approximate Reasoning*, **17**(2-3), pp. 217 (1997).
12. Denoeux, T. "A nearest neighbor classification rule based on Dempster-Shafer theory", *IEEE Transactions on Systems, Man, and Cybernetics*, **25**(5), pp. 804-813 (1995).
13. Denoeux, T. "A neural network classifier based on Dempster-Shafer", *IEEE Transactions on Systems, Man, and Cybernetics A*, **30**(2), pp. 131-150 (2000).
14. Diaz, J. et al. "A similarity measure between basic belief assignments", *9th International Conference on Information Fusion*, Florence, Italy (2006).

15. Jousselme, A.L. et al. "A new distance between two bodies of evidence", *Information Fusion*, **2**, pp. 91-101 (2001).
16. Ristic, B. and Smets, P. "The TBM global distance measure for the association of uncertain combat ID declarations", *Information Fusion*, **7**, pp. 276-284 (2006).
17. Schubert, J. "On nonspecific evidence", *International Journal of Artificial Systems*, **8**, pp. 711-725 (1993).
18. Tessem, B. "Approximations for efficient computation in the theory of evidence", *Artificial Intelligence*, **61**(2), pp. 315-329 (1993).
19. Guan, J.W. and Bell, D.A., *Evidence Theory and Its Applications*, Elsevier Science Publisher B.V., Amsterdam (1991).
20. Dempster, A.P. "Upper and lower probabilities induced by a multivalued mapping", *Annals of Mathematical Statistics AMS*, **38**, pp. 325-339 (1967).
21. Shafer, G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ (1976).
22. Smets, P. "Belief functions; the disjunctive rule of combination and the generalized Bayesian theorem", *International Journal of Approximate Reasoning*, **9**, pp. 1-35 (1993).
23. Smets, P. and Kennes, R. "The transferable belief model", *Artificial Intelligence*, **66**, pp. 191-243 (1994).
24. Khatibi, V. and Montazer, G.A. "Coronary heart disease risk assessment using Dempster-Shafer evidence theory", *14th Int'l Computer Conference of Iranian Computer Society (CSICC2009)*, Tehran, Iran (2009).
25. Khatibi, V. and Montazer, G.A. "A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment", *Expert Systems with Applications*, **37**(12), pp. 8536-8542 (2010).
26. Zouhal, L.M. and Denoeux, T. "An evidence-theoretic k-NN rule with parameter optimization", *IEEE Transactions on Systems, Man, and Cybernetics- Part C*, **28**(2), pp. 263-271 (1998).
27. Denoeux, T. and Zouhal, L.M. "Handling possibilistic labels in pattern classification using evidential reasoning", *Fuzzy Sets and Systems*, **122**(3), pp. 47-62 (2001).
28. Denoeux, T. "Reasoning with imprecise belief structures", *International Journal of Approximate Reasoning*, **20**, pp. 79-111 (1999).
29. Denoeux, T. "Modeling vague beliefs using fuzzy-valued belief structures", *Fuzzy Sets and Systems*, **116**, pp. 167-199 (2006).
30. Young, T. and Fu, K.S., *Handbook of Pattern Recognition and Image Processing*, Academic Press, London (1986).
31. Khatibi, V. and Montazer, G.A. "Intuitionistic fuzzy set vs. fuzzy set application in medical pattern recognition", *Artificial Intelligence in Medicine*, **47**(1), pp. 43-52 (2009).
32. Kukar, M. et al. "Analysing and improving the diagnosis of ischaemic heart disease with machine learning", *Artificial Intelligence in Medicine*, **16**, pp. 25-50 (1999).
33. Facklam, H. and Facklam, M., *Bacteria*, Twenty-First Century, Kent (1995).
34. Ricciuti, E.R., *Microorganisms: The Unseen World*, Blackbirch, San Diego (1993).
35. Singleton, P., *Bacteria in Biology, Biotechnology, and Medicine*, Wiley, Chichester (1999).
36. Sankaran, N., *Microbes and People: An A-Z of Microorganisms in Our Lives*, Oryx, Phoneix (2001).
37. WHO "International coronary and heart diseases report", Report World Health Organization (2008).
38. IMHMT "Iranian population's coronary and heart diseases report", Report Iran's Ministry of Health and Medical Treatment, Tehran, Iran (2008).
39. Simon, S., *The Heart: Our Circulatory System*, Morrow Junior (1996).
40. McGoon, M.D., *Mayo Clinic Heart Book*, William Morrow (2000).
41. McMahan, C. et al. "Risk scores predict atherosclerotic lesions in young people", *Arch Intern Med*, **165**, pp. 883-890 (2005).
42. Wilson, P. et al. "Prediction of coronary heart disease using risk factor categories", *Circulation*, **97**, pp. 1837-1847 (1998).
43. Janosi, A. *University of California, Irvine; Machine Learning Repository; Heart Disease Data Set; Hungarian Institute of Cardiology's Heart Disease Data Set*. 1988 [cited; Available from: <http://archive.ics.uci.edu/ml/machinelearning-databases/heart-disease/>]
44. Akay, M. "Noninvasive diagnosis of coronary artery disease using a neural network algorithm", *Biological Cybernetics*, **67**, pp. 361-367 (1992).
45. Haddad, M. et al. "Feasibility analysis of a casebased reasoning system for automated detection of coronary heart disease from myocardial scintigrams", *Artificial Intelligence in Medicine*, **9**, pp. 61-78 (1999).
46. Detrano, R. et al. "International application of a new probability algorithm for the diagnosis of coronary artery disease", *American Journal of Cardiology*, **64**, pp. 304-310 (1989).
47. Gennari, J.H. et al. "Models of incremental concept formation", *Artificial Intelligence*, **40**, pp. 11-61 (1989).

BIOGRAPHIES

Vahid Khatibi received his B.S. degree in Computer Software Engineering from University of Applied Science and Technology, Tehran, Iran in 2006, and his M.S. degree in Information Technology Engineering from Tarbiat Modares University, Tehran, Iran, in

2009. Also, he is a researcher of the Iran Telecommunication Research Center (ITRC), Tehran, Iran. His areas of research include Information Systems Engineering, Knowledge Discovery, Pattern Recognition, e-Business, IT Development Models and Using Soft Computing and Data Mining Techniques in the Relative Systems Analysis.

Gholam Ali Montazer received his B.S. degree in Electrical Engineering from K.N. Toosi University of

Technology, Tehran, Iran, in 1991, his M.S. degree in Electrical Engineering from Tarbiat Modares University, Tehran, Iran, in 1994, and his Ph.D. degree in Electrical Engineering from the same university in 1998. He is an associate professor of the Information Technology department in Tarbiat Modares University, Tehran, Iran. His areas of research include Information Engineering, Knowledge Discovery, Intelligent Methods, System Modeling, e-Learning and Image Mining.