

A bi-level belief rule based decision support system for diagnosis of lymph node metastasis in gastric cancer



Zhi-Guo Zhou^{a,b}, Fang Liu^{a,b,*}, Li-Cheng Jiao^b, Zhi-Jie Zhou^c, Jian-Bo Yang^d, Mao-Guo Gong^b, Xiao-Peng Zhang^e

^aSchool of Computer Science and Technology, Xidian University, Xi'an 710071, PR China

^bKey Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071, PR China

^cHigh-Tech Institute of Xi'an, Xi'an, Shaanxi 710025, PR China

^dManchester Business School, The University of Manchester, Manchester M15 6PB, UK

^eKey laboratory of Carcinogenesis and Translational Research of Ministry of Education of China, Department of Radiology, Peking University School of Oncology, Beijing Cancer Hospital & Institute, Beijing 100142, PR China

ARTICLE INFO

Article history:

Received 27 February 2013

Received in revised form 30 August 2013

Accepted 2 September 2013

Available online 29 September 2013

Keywords:

Clinical decision support system

Gastric cancer

Lymph node metastasis

Belief rule base

Clonal selection algorithm

ABSTRACT

Lymph Node Metastasis (LNM) in gastric cancer is an important prognostic factor regarding long-term survival. As it is difficult for doctors to combine multiple factors for a comprehensive analysis, Clinical Decision Support System (CDSS) is desired to help the analysis. In this paper, a novel Bi-level Belief Rule Based (BBRB) prototype CDSS is proposed. The CDSS consists of a two-layer Belief Rule Base (BRB) system. It can be used to handle uncertainty in both clinical data and specific domain knowledge. Initial BRBs are constructed by domain specific knowledge, which may not be accurate. Traditional methods for optimizing BRB are sensitive to initialization and are limited by their weak local searching abilities. In this paper, a new Clonal Selection Algorithm (CSA) is proposed to train a BRB system. Based on CSA, efficient global search can be achieved by reproducing individuals and selecting their improved matured progenies after the affinity maturation process. The proposed prototype CDSS is validated using a set of real patient data and performs extremely well. In particular, BBRB is capable of providing more reliable and informative diagnosis than a single-layer BRB system in the case study. Compared with conventional optimization method, the new CSA could improve the diagnostic performance further by trying to avoid immature convergence to local optima.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Gastric cancer is one of the leading causes for cancer-related death in the world [25]. Lymph Node Metastasis (LNM) is one of the most important prognostic factors regarding long-term survival [7,24,30,34]. The TNM¹ staging system based on American Joint Committee on Cancer (AJCC) is taken as the evaluated standard and has been widely accepted [15]. Based on this standard, the 5-year survival rate of patients in the N0 stage after surgery is 86.1%, while the survival rate for the N1, N2 and N3 stage patients is just 58.1%, 23.3% and 5.9%, respectively [45].

Currently, doctors manually diagnose LNM according to the size of lymph nodes, which relies on various imaging methods, such as multi-slice spiral Computerized Tomography (CT), Magnetic

Resonance Imaging (MRI) and Positron Emission computed Tomography (PET). However, any of the above imaging tools cannot assess the lymph node status satisfactorily. For example, large lymph nodes may be caused by inflammation, while small ones may be metastatic. In fact, many studies have shown that LNM is not only associated with the size of lymph nodes but also with the number of lymph nodes. It is difficult for doctors to combine the above two factors to conduct a comprehensive analysis. Fortunately, Clinical Decision Support System (CDSS) from Information Technology (IT) can be used to help solve this problem, which has been applied to medical domains successfully [1,18].

In the current studies, several CDSSs have been proposed for diagnosing LNM, such as Support Vector Machine (SVM) based CDSS [46], and Artificial Neural Network (ANN) based CDSS [9]. However, there are some limitations when these methods are applied in reality. Firstly, these methods are block-box modeling approaches in nature and their internal structures are not directly linked to the reasoning logic or process, which makes it difficult for doctors to know how important each medical attribute is regarding prediction results. Secondly, ANN and SVM are implemented by

* Corresponding author at: School of Computer Science and Technology, Xidian University, Xi'an 710071, PR China. Tel.: +86 029 88204310; fax: +86 029 88201023.

E-mail address: f63liu@163.com (F. Liu).

¹ The international normative TNM classification describes the state of the tumor (T), the lymph nodes (N), and possible metastases (M).

pure nonlinear data fitting. However, as doctors play an essential role in determining the final stage, it is important to evaluate LNM using both clinical data and doctors' domain specific knowledge.

From the above discussions, it can be concluded that CDSSs can take into account human judgments are more appropriate to help diagnose LNM. For knowledge based CDSSs, representing and reasoning with uncertain medical knowledge are critical areas which need outstanding methodologies and techniques [18,23]. This is because uncertainties exist in clinical signs, clinical domain knowledge and clinical symptoms, and it is difficult to select domain knowledge and construct knowledge bases [11]. Several CDSSs have been developed by using domain knowledge. An early rule-based expert system named as MYCIN [29] was developed in the early 1970s. Another CDSS called Iliad [33] was developed to assist medical decision-making. In Iliad, a sequential Bayesian inference algorithm is used to generate a ranked list of diagnoses and assign a posterior probability to each diagnosis. Fuzzy logic was used in CDSS for renal transplantation assignment [44]. However, the above methods have their deficiencies in knowledge representation and reasoning. Complete knowledge about all parameters is required in a Bayesian inference model. This creates a bottleneck in a knowledge acquisition process that can be extremely time-consuming. As there is usually a fuzzification and de-fuzzification procedure in a fuzzy inference process, reasoning is controversial in most fuzzy logic based systems [11].

To support modeling and inference with clinical domain knowledge under uncertainties, a recently developed belief Rule-based Inference Methodology using the Evidential Reasoning approach (RIMER) [41] is employed for developing an intelligent CDSS [11] which is used for clinical risk assessment of cardiac chest pain. In this CDSS, Belief Rule Base (BRB) is used to model clinical domain knowledge and reasoning is implemented by the Evidential Reasoning (ER) approach [37,38,40–42]. Compared to other CDSSs as described above, it has the following advantages: (1) BRB can be initialized using domain knowledge or assigned randomly, because the belief rules can be adjusted by historical clinical data, which can help circumvent the bottleneck of estimating all initial parameters. (2) Different types of uncertainties can be included in initial BRB as ER can preserve the original features of uncertainties in inference process and the corresponding effects can also be represented in the final conclusion.

According to the clinical domain knowledge, LNM can be divided into four stages. The higher the stage is, the more severe the illness. Stage 0 represents that LNM does not occur, while stage 3 shows that metastasis is most serious. These two stages can be predicted easily by a single-layer BRB system. However, the prediction of stages 1 and 2 is difficult because they are easily confusing [13,17]. So a single-layer BRB system is difficult to diagnose LNM accurately. To solve this problem, a Bi-level Belief Rule Based (BBRB) prototype CDSS is proposed in this paper. The new prototype CDSS consists of two layers [10,16,28] which are modeled by a single-layer BRB. In the first layer, original stages 1 and 2 are integrated into a new stage, resulting in three stages which include original stages 0, stage 3 and the new stage. In the second layer, a specialized BRB is constructed to distinguish stage 1 from stage 2. Compared to single-layer BRB, BBRB can provide more reliable and informative diagnosis for LNM. As manually constructed belief rules may not be accurate, there is a need to train BBRB. However, training a BRB system is sensitive to system initialization and may lead to a locally optimized BRB system. In this paper, a novel Clonal Selection Algorithm (CSA) for training BRB system is proposed. Compared to conventional methods, new CSA is capable of improving performance significantly.

The rest of the paper is organized as follows. The problem formulation is shown in Section 2. In Section 3, the BBRB CDSS

prototype is described. A new CSA based method for optimizing BRB is developed in Section 4. In Section 5, the proposed CDSS prototype for diagnosing LNM is presented. The validation of BBRB is discussed in Section 6. This paper is concluded in Section 7.

2. Problem formulation

Suppose that $x = [x_1, x_2, \dots, x_M]$ is the diagnosis factor which are obtained from medical images such as CT images, where $x_i (i = 1, \dots, M)$ denotes an attribute and M is the number of attributes. It is assumed that D is the output and P is the corresponding parameter vector. The LNM diagnosis problem is in essence aimed to establish causal relationships between x and D . Such relationships are generally represented by

$$D = f(x, P) \tag{1}$$

where f is in general a function of D . As there are four stages, the output can also be represented as:

$$D = \{(D_0, \beta_0), (D_1, \beta_1), (D_2, \beta_2), (D_3, \beta_3)\} \tag{2}$$

where $D_i (i = 0, \dots, 3)$ is the stage of LNM, $\beta_i (i = 0, \dots, 3)$ is the corresponding belief degree and the following constraints are satisfied.

$$\sum_{i=0}^3 \beta_i = 1 \quad \text{and} \quad 0 \leq \beta_i \leq 1, \quad i = 0, \dots, 3 \tag{3}$$

As mentioned in Section 1, D_1 and D_2 is difficult to distinguish. Therefore, a new BBRB model will be presented to solve this problem. On the other hand, the parameter P is initialized by expert knowledge and may be inaccurate. Although the optimal learning methods have been proposed, they are sensitive to initialization and may lead to local optimization. Therefore, a new CSA based method is proposed, which will be described in Section 4.

3. BBRB prototype CDSS

In this section, the BBRB model is presented as a two layer system.

3.1. Proposed BBRB model

The new CDSS prototype is shown in Fig. 1, where x_1, x_2, \dots, x_M denote the inputs and stage 0–stage 3 is the final output. BRB_1 represents the first layer in BBRB, while BRB_2 represents the second layer. In the first layer, the original stages D_1 and D_2 are integrated into a new stage, and the output O_1 is represented as:

$$O_1 = \{(D_1^1, \beta_1^1), (D_2^1, \beta_2^1), (D_3^1, \beta_3^1)\} \tag{4}$$

where $D_i^1 (i = 1, \dots, 3)$ is the output in BRB_1 and $\beta_i^1 (i = 1, \dots, 3)$ is the belief degree which satisfies:

$$\sum_{i=1}^3 \beta_i^1 = 1 \quad \text{and} \quad 0 \leq \beta_i^1 \leq 1, \quad i = 1, \dots, 3 \tag{5}$$

If the result is D_1^1 or D_3^1 , it shows that LNM will be predicted as D_0 or D_3 and the decision process ends. Otherwise, the second layer will be activated. Let O_2 represent the output of BRB_2 as follows:

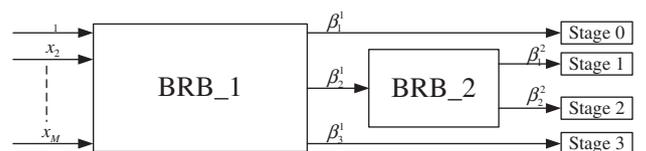


Fig. 1. The new BBRB CDSS prototype.

$$O_2 = \left\{ (D_1^2, \beta_1^2), (D_2^2, \beta_2^2) \right\} \quad (6)$$

where $D_i^2 (i = 1, 2)$ is the output of BRB_2, and $\beta_i^2 (i = 1, 2)$ is the assigned belief degree that should satisfy:

$$\sum_{i=1}^2 \beta_i^2 = 1 \quad \text{and} \quad 0 \leq \beta_i^2 \leq 1, \quad i = 1, 2 \quad (7)$$

In BRB_2, the result D_1^2 shows that the final result is D_1 , while D_2^2 shows that it is D_2 . Since this layer is specially constructed, D_1 and D_2 may be predicted more accurately. After integrating O_1 and O_2 , the final result D can be obtained. As each layer in BBRB is a BRB system, basics about BRB will be described in the following subsection.

3.2. A brief description of BRB

The proposed BBRB prototype CDSS is a two-layer system. Each layer is modeled by a single-layer BRB system, which consists of belief rules for domain knowledge representation. A typical belief rule is defined as follows [41]:

$$R_k^i: \text{if } x_1 \text{ is } A_1^k \wedge x_2 \text{ is } A_2^k \cdots \wedge x_M \text{ is } A_M^k, \text{ Then } \{ (D_1, \beta_{1,k}), \dots, (D_N, \beta_{N,k}) \}$$

With a rule weight θ_k and attribute weight $\delta_{1,k}, \delta_{2,k}, \dots, \delta_{M,k}$

$$(8)$$

where x_1, x_2, \dots, x_M denotes the antecedent attributes in the k th rule. The referential value of the j th antecedent attribute in the k th rule is denoted by $A_j^k (j = 1, \dots, M, k = 1, \dots, L)$. M represents the number of antecedent attribute, and L is the number of rules in BRB. Further, “ \wedge ” represents the “AND” relationship. In addition, $D_j (j = 1, \dots, M)$ is the consequence of LNM stage, and $\beta_{j,k} (j = 1, \dots, N, k = 1, \dots, L)$ is the belief degree assigned to the stage D_j . It should be noted that the k th rule is complete if $\sum_{j=1}^N \beta_{j,k} = 1$; otherwise, it is incomplete. $\theta_k (k = 1, \dots, L)$ is the relative weight of the k th rule, while $\delta_{j,k} (j = 1, \dots, M)$ represents the relative weight of the j th antecedent attributes in the k th rules, which shows the relative importance of the attribute.

The inference of BRB is implemented by the ER approach. It is originally proposed to deal with Multiple Attribute Decision Analysis (MADA) problems [14] having both qualitative and quantitative attributes under uncertainty [37,38]. The kernel of the ER approach is the ER algorithm which is based on the decision theory and the Dempster-Shafer (D-S) theory of evidence [8,27]. It is utilized to aggregate nonlinear information under uncertainty. It consists of the following two steps.

(1) Calculating the activation weight.

The activation weight of the k th rule ω_k is calculated by Eq. (9):

$$\omega_k = \theta_k \prod_{i=1}^M (\alpha_i^k)^{\delta_{i,k}} / \sum_{l=1}^L \theta_l \prod_{i=1}^M (\alpha_i^l)^{\delta_{i,k}} \quad \text{and} \quad \bar{\delta}_{i,k}$$

$$= \delta_{i,k} / \max_{i=1, \dots, M} \{ \delta_{i,k} \} \quad (9)$$

where $\bar{\delta}_{i,k} (i = 1, \dots, M)$ are the normalized attribute weights. $\alpha_i^k (i = 1, \dots, M)$, which is called the individual matching degree, is the degree of belief to which the input for the i th antecedent attribute belongs to its j th referential value A_i^k in the k th rule. It can be generated using various ways, depending on the nature of an antecedent attribute and data that are available such as a qualitative attribute using linguistic values [43]. α_k is the normalized combined matching degree as follows:

$$\alpha_k = \prod_{i=1}^M (\alpha_i^k)^{\bar{\delta}_{i,k}} \quad (10)$$

(2) Combining rules by the ER approach.

The final conclusion D which is generated through aggregating all rules using the ER analytical algorithms [32] can be represented as follows:

$$D = f(x) = \left\{ (D_j, \hat{\beta}_j), \quad j = 1, \dots, N \right\} \quad (11)$$

where $\hat{\beta}_j$ denotes the belief degree in D_j , and are calculated as follows:

$$\hat{\beta}_j = \mu \times \left[\prod_{k=1}^L \left(\omega_k \beta_{j,k} + 1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) - \prod_{k=1}^L \left(1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) \right] / \left[1 - \mu \times \prod_{k=1}^L (1 - \omega_k) \right] \quad (12)$$

$$\mu = \left[\sum_{j=1}^N \prod_{k=1}^L \left(\omega_k \beta_{j,k} + 1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) - (N-1) \prod_{k=1}^L \left(1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) \right]^{-1} \quad (13)$$

Assume P represents a vector composed of the above parameters in BRB. P includes rule weight, attribute weights and consequent belief degrees as follows:

$$P = [\theta_1, \dots, \theta_L, \delta_1, \dots, \delta_M, \beta_{1,1}, \dots, \beta_{N,L}]^T \quad (14)$$

The parameters should satisfy the following constraints:

$$0 \leq \theta_k \leq 1, 0 \leq \delta_m \leq 1, 0 \leq \beta_{j,k} \leq 1, \sum_{j=1}^N \beta_{j,k} = 1, \quad k = 1, \dots, L, \quad m = 1, \dots, M, \quad j = 1, \dots, N \quad (15)$$

The initial parameters of two BRBs in BBRB are given using domain knowledge and may not be accurate. Therefore, it is necessary to adjust them. In the following section, a new optimization method is proposed to fine-tune these parameters for optimal performance.

4. New optimization model and CSA for training BBRB

In this section, a new optimization method that is composed of an optimization model and a new CSA based optimization algorithm is proposed.

4.1. Optimization model of BBRB

The aim of training BBRB is to find a set of parameters which can accurately represent domain knowledge. Fig. 2 shows the optimization model of BBRB. In Fig. 2, U is the given input, O is the simulated output, and \hat{O} is the observed output. $\zeta(P)$ is the diagnosing error rate. Thus the objective function is described as follows:

$$\min_P \{ \zeta(P) \} \quad (16)$$

4.2. The new CSA based optimization algorithm

As two BRBs in BBRB are independent, they can be optimized, respectively. The optimal learning method for training BRB was

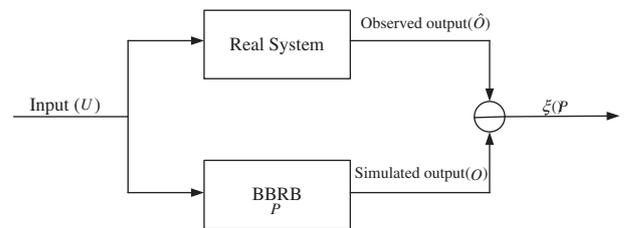


Fig. 2. Optimization model for training BBRB.

proposed by Yang *et al.* [43]. Zhou *et al.* proposed the recursive optimal training methods to update BRB [48–51]. These training schemes have been applied to several practical problems [5,36,47] successfully. However, current strategies are sensitive to initialization and are limited by their weak local searching abilities.

It is well known that Evolutionary Algorithms (EAs) utilize computational models of evolutionary processes as key elements and are adaptive search schemes which are suitable for optimization in a nonlinear and discrete space. Because this type of algorithms can obtain optimal or near-optimal solutions, they have been applied to many real-world optimization problems. Several different computing paradigms have also been developed, such as genetic algorithm, and particle swarm optimization simulating bird flight. Artificial Immune System (AIS) is another adaptive system inspired by human immune system. AIS makes use of the mechanism of vertebrate immune system in terms of the model of information processing, and constructs new intelligent methods for problem solving. These methods provide evolutionary learning mechanisms including robustness, unsupervised learning, self-organization, and memory. Therefore, they have received a significant amount of interest from researchers and industrial sponsors in recent years. Clonal Selection Algorithm (CSA) is a major part for optimization in AIS, which has received wide attention [19–21]. The potential of CSA for global optimization inspires us to use it to train a BRB system. To solve this problem, a new CSA, which is a global and efficient optimization algorithm, is proposed in this paper.

It is well known that CSA should initialize population in the first step. Most of the existing CSAs initialize population randomly. However, BRB allows direct intervention of human judgments and thus domain knowledge should be embedded. The initial population is a suitable place to embed knowledge from problem domain.

In our case, initial rules in BRB are extracted by examining historical data and domain knowledge. Different individuals are generated in initial population to keep the diversity of individuals. To obtain a sufficient number of individuals, we utilize individuals provided by experts to generate other individuals which are added to the initial population. This process will not stop until sufficient individuals are generated. So the initial population consists of two parts: individuals provided by experts and generated antibodies [26]. The initial population is represented as

$$Pop(j) = \{I_j^1, I_j^2, \dots, I_j^{Num}\}, \quad j = 0 \tag{17}$$

where the population number is *Num*. Suppose that *G* is the maximum number of evolutionary generation. The initial objective-function values of population are

$$\zeta^j = \{\zeta_1^j, \zeta_2^j, \dots, \zeta_M^j\}, \quad j = 0 \tag{18}$$

The whole process is shown as follows.

Step 1: Clone operator is executed. Assume that *CS* represents clonal scale. Each individual in *Pop(j)* is cloned *CS* times.

Step 2: Mutation is operated on the cloned population. As different parameters in BRB have different constraints, mutation cannot be executed simultaneously. So a novel Group based Mutation Operator (GMO) is devised.

GMO is a two-stage operator. In the first stage, certain individuals are selected for mutation with probability $1/CS$. Then the parameters in BRB are divided into several groups according to the types of constraints. Attribute weights $\delta_j, j = 1, \dots, M_k$ and rule weights $\theta^i, i = 1, \dots, L$ are two groups, and they satisfy

$$0 \leq \theta^i \leq 1 \quad \text{and} \quad 0 \leq \delta_j \leq 1 \tag{19}$$

Meanwhile, there are *L* groups of belief degrees as *L* rules are generated. Every group should satisfy

$$\sum_{j=1}^N \beta_j^i = 1, i = 1, 2, \dots, L \quad \text{and} \quad 0 \leq \beta_j^i \leq 1 \tag{20}$$

Therefore, there are *L* + 2 groups in total. According to prior knowledge, each group is chosen by the probability of 0.5 in the selected individuals. If a group is chosen, it will be mutated by a set of randomly generated constrained data.

Step 3: Updating population is performed. After the mutation operation is finished, several new individuals are generated and added to the population. To keep the number of original population unchanged, a new population is generated by selecting *M* individuals. The objective-function values of cloned population are calculated firstly. Then the minimal objective-function value of every cloned individual will be selected and represented by $\zeta^c = \{\zeta_1^c, \zeta_2^c, \dots, \zeta_M^c\}$. The cloned individual will replace the original individual only if

$$\zeta_i^j > \zeta_i^c \tag{21}$$

Let $j = j + 1$. The objective-function of the new population is $\zeta^j = \{\zeta_1^j, \zeta_2^j, \dots, \zeta_M^j\}$ and the corresponding population *Pop(j)*.

Step 4: Step 1–3 is executed repeatedly until $j = G$. When the algorithm stops, the individual with minimal objective-function value is considered as output.

After constructing the BBRB model and describing the proposed optimization algorithm, we will utilize them for diagnosing LNM which is presented in the following section (see Table 1).

5. BBRB prototype CDSS for diagnosing LNM

In this section, BBRB prototype CDSS is used for diagnosing LNM. The proposed prototype CDSS consists of two aspects: (i) constructing each component in a BBRB system and (ii) the training process.

5.1. Determination of the initial belief rules in BBRB

5.1.1. Referential points of the antecedents and consequence

Since the size of LN (represented as *LNsize*) and the number of LN (represented as *LNnumber*) are the two important factors in LNM diagnosis, they are chosen as the medical antecedent attributes. The two medical antecedent attributes are described as follows:

- (1) The number of LN: The number of all visible gastric regional lymph nodes in CT images by groups is counted.
- (2) The size of LN: The short axis of the largest lymph node detected in CT images is measured.

To represent the domain knowledge in BBRB, the referential points are given. Note that the number and the value of referential points are set on the basis of clinical domain knowledge [22,35]. As there are two BRBs in the BBRB, two groups of referential points are given.

In BRB_1, for *LNsize*, eleven referential points are used, and they are zero (Z), super small (SS), very small (VS), small (S), a little small (LS), middle (M), a little large (LL), large (L), very large (VL), greatly large (GL), and super larger (SL). That is

$$A_1^1 \in \{Z, SS, VS, S, LS, M, LL, L, VL, GL, SL\} \tag{22}$$

For *LNnumber*, nine referential points are used, and they are zero (Z), very small (VS), small (S), a little small (LS), middle (M), a little larger (LL), large (L), very large (VL), and greatly large (GL), i.e.,

Table 1
CSA for training BRB.

Input: Set $j = 0$, initialize population $Pop(j)$ and calculate the objective-function values of φ^j .

while $j < G$

 Clone each individual in $Pop(j)$ CS times.

 Perform GMO on the cloned population.

 Compute the objective-function values of the cloned population.

 Select individuals from the combined population, and form next $Pop(j)$,
 $j = j + 1$.

End while

Output: the individual with minimal objective-function value.

$$A_2^1 \in \{Z, VS, S, LS, M, LL, L, VL, GL\} \quad (23)$$

As the number of output in BRB_1 is three, three referential points are used and shown as follows:

$$D^1 = (D_1^1, D_2^1, D_3^1) \quad (24)$$

However, the referential points defined above are in linguistic terms. So they should be quantified. Table 2 shows the quantified value of $LNsize$, while Table 3 shows the quantified value of $LNnumber$.

BRB_2 is mainly used for evaluating original stage 1 and stage 2. For $LNsize$ in BRB_2, nine referential points are used, and they are zero (Z), very small (VS), small (S), a little small (LS), middle (M), a little larger (LL), large (L), very large (VL), and greatly large (GL). That is

$$A_1^2 \in \{Z, VS, S, LS, M, LL, L, VL, GL\} \quad (25)$$

For $LNnumber$ in BRB_2, eight referential points are used. They are zero (Z), very small (VS), small (S), middle (M), a little larger (LL), large (L), very large (VL), and greatly large (GL), i.e.,

$$A_2^2 \in \{Z, VS, S, M, LL, L, VL, GL\} \quad (26)$$

The output in BRB_2 is two, and there are two referential points as follows. In addition, the corresponding quantified values are shown in Tables 4 and 5.

$$D^2 = (D_1^2, D_2^2) \quad (27)$$

The final output of BBRB is four stages, so there are four referential points. They are

$$D = (D_0, D_1, D_2, D_3) \quad (28)$$

5.1.2. Construction of the rule base

According to the above description, the prototype CDSS can be generated. A belief rule in BRB_1 is represented as follows.

$$R_k^1 : \text{if } LNsize \text{ is } A_1^k \text{ and } LNnumber \text{ is } A_2^k$$

$$\text{THEN stage is } \left\{ (D_1^1, \beta_{1k}^1), (D_2^1, \beta_{2k}^1), (D_3^1, \beta_{3k}^1) \right\}, \left(\sum_{i=1}^3 \beta_{ik}^1 \leq 1 \right)$$

with a rule weight θ_k^1 ($k = 1, 2, \dots, 99$) and attribute weight δ_1^1, δ_2^1

$$(29)$$

The belief rule in BRB_2 can be represented as shown in Eq. (30), where A_1^k and A_2^k are the referential values in BRB_1 and BRB_2 as defined in Tables 2–5. As $LNsize$ is divided into 11 terms and $LNnumber$ 9 terms, the combinations of the two antecedents generate 99 rules in total in BRB_1. Similarly, 72 rules are generated in BRB_2. In the two BRBs, we assume that all belief rules have equal rule weight, all medical antecedent attributes have equal weight, and the initial belief degrees assigned to each stage is based on the statistics of available historical patient data and clinical domain knowledge.

$$R_k^2 : \text{if } LNsize \text{ is } A_1^k \text{ and } LNnumber \text{ is } A_2^k$$

$$\text{THEN stage is } \left\{ (D_1^2, \beta_{1k}^2), (D_2^2, \beta_{2k}^2) \right\}, \left(\sum_{i=1}^2 \beta_{ik}^2 \leq 1 \right)$$

with a rule weight θ_k^2 ($k = 1, 2, \dots, 72$) and attribute weight δ_1^2, δ_2^2

$$(30)$$

Compared to traditional rules, reasoning with belief rules is more informative because BBRB shows a distributed clinical conclusion which provides an overall view about the situation of LNM. This is helpful for each patient to determine a therapeutic regimen. For example, an evaluating result for one patient in BRB_1 is $\{(D_1^1, 0.00), (D_2^1, 0.80), (D_3^1, 0.20)\}$ and BRB_2 is $\{(D_1^2, 0.65), (D_2^2, 0.35)\}$. The final result is $\{(D_0, 0.00), (D_1, 0.52), (D_2, 0.28), (D_3, 0.20)\}$. Although the result is mainly stage 1, we still should consider the changeable tendency of state of an illness because the condition may become worse in the future. Therefore, doctors not only should provide treatment for illness at stage 1, but also think about how to treat it if patient's condition becomes stage 2.

5.2. CSA based optimization for training BBRB

As the two BRBs in BBRB are independent, they can be optimized, respectively. The parameters that need to be trained are shown in Eq. (4). In BRB_1, $N = 3$, $L = 99$ and $M = 2$, while $N = 2$, $L = 72$ and $M = 2$ in BRB_2. The inputs for the two layers are both $LNsize$ and $LNnumber$, that is

$$U^i = (LNsize, LNnumber), \quad i = 1, 2 \quad (31)$$

The training process is outlined in the following steps.

Step 1: Set initial parameters

When CSA is used, an initial population should be generated. Meanwhile, the clinical domain knowledge should be embedded. In this paper, the initial belief degrees are obtained based on the statistics of available historical patient data and clinical domain knowledge. Then two populations which both consist of 100 individuals in two BRBs are generated, respectively. To simplify the initial process, we use the *fminimax* function provided by the optimization toolbox in MATLAB environment to generate individuals. Different intermediate results in optimizing process are considered as individuals. The process will not stop until generating the required number of individuals. The maximum number of evolutionary generation is set as 50.

Step 2: Input data transformation

A rule-based transformation technique is applied to the transformation of quantitative data [39]. The input value U^1 in BRB_1 needs to be transformed and represented in terms of the referential values as defined in Tables 2 and 3 using belief degrees representing the degrees to which the input values match the referential values [36]. The input U^2 in BRB_2 is also transformed and represented based on Tables 4 and 5.

Step 3: Combine the activated rules

By utilizing the ER analytical algorithm [32], the activated rules can be combined to generate the final result. For BRB_1, the result is as follows:

$$O^1(U^1) = \left\{ (D_j^1, \beta_j^1) \mid j = 1, \dots, 3 \right\} \quad (32)$$

The result in BRB_2 is

$$O^2(U^2) = \{(D_j^2, \beta_j^2) | j = 1, 2\} \tag{33}$$

Step 4: Calculate the prediction error rate

The objective functions that need to be optimized are calculated as follows:

$$\xi^i(P^i) = \frac{C}{M}, \quad i = 1, 2 \tag{34}$$

where M is the number of training samples, and C is the number of correct prediction samples in training set.

Step 5: Update the parameters

We use the proposed CSA for updating. The objective is to update P^i in two BRBs so as to minimize the prediction error rate, that is

$$\min_{P^i} \{\xi^i(P^i)\}, \quad i = 1, 2 \tag{35}$$

Then a new population will be generated. Steps 2–5 will be executed iteratively until it achieves the maximum number of evolutionary generation. As a result, there will be a new set of parameters for two BRBs. To validate the proposed model and CSA based training algorithm, a case study is given in the next section.

6. An experimental case study

The proposed prototype CDSS is validated by a set of real patient data. To test the system, the 2-cross-validation approach is used [31]. In detail, we divide the data into two subsets with similar size. Then we run each experiment twice, each time using one subset as training set, and using the other subset as test set. Different from traditional cross-validation which randomly split available data into training and testing set, we split the data into two similar sets to ensure most belief rules in BBRB can be trained and also tested in the testing process [11].

To better demonstrate the overall performance of system, we use Confusion Matrix (CM) [3] to show the results. Each column of the matrix represents the instances in a predicted stage, while each row represents the instances in an actual stage. There are two evaluated indexes: User Accuracy (UA) and Procedure Accuracy (PA). UA is the ratio of the number of correct predicted stage to the number of test sets in this stage. PA is the ratio of the number of correct predicted stage to the number of a part of test data whose predicting result is this stage. All the experiments are performed in MATLAB 2009a and run on PC with 2.99 GHz and 2 G memory.

In this section, the real patient data are described firstly. Then BBRB is validated by comparing it with a traditional BRB system. The new optimization method is validated in the following subsection. Finally, a comparative study among SVM, ANN and BBRB is presented.

6.1. Description of the real patient data

255 Real patient data provided by Beijing Cancer Hospital (Beijing, PR China) are taken as samples. It should be noted that the informed consent from all selected patients were obtained prior to the routine clinical course. The details of patient characteristics

Table 2
The referential points of $LNsize$ in BRB_1.

Linguistic terms	Z	SS	VS	S	LS	M	LL	L	VL	GL	SL
Numerical values	0	5	6	7	8	9	10	12	17	21	50

Table 3
The referential points of $LNnumber$ in BRB_1.

Linguistic terms	Z	VS	S	LS	M	LL	L	VL	GL
Numerical values	0	5	7	9	12	16	20	30	50

Table 4
The referential points of $LNsize$ in BRB_2.

Linguistic terms	Z	VS	S	LS	M	LL	L	VL	GL
Numerical values	0	5	6	7	8	10	14	20	50

Table 5
The referential points of $LNnumber$ in BRB_2.

Linguistic terms	Z	VS	S	M	LL	L	VL	GL
Numerical values	0	6	8	10	12	13	20	50

Table 6
Patient Characteristics.

Clinic pathological features	Value
Number of patients	255
Average age (y)	57 (29–85)
Ratio of men to women	185:69
<i>Lymph node metastasis</i>	
Stage 0	46 (18.1%)
Stage 1	49 (19.2%)
Stage 2	88 (34.5%)
Stage 3	72 (28.2%)

Table 7
Feature description of data.

Patient data	Stage 0	Stage 1	Stage 2	Stage 3
The size of LN (mm)	6.3±18.4	9.6±43	9.7±23	10.7±32
The number of LN	6.4±16.0	8.5±32	11.9±37	13.9±41

Table 8
Confusion matrix of BBRB.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	38	7	1	0	46	82.61
Stage 1	3	33	11	2	49	67.35
Stage 2	0	7	76	5	88	86.36
Stage 3	0	0	2	70	72	97.22
Total	41	47	90	77	255	
PA (%)	92.68	70.21	84.44	90.91	85.1	

Table 9
Confusion matrix of BRB.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	37	9	0	0	46	80.43
Stage 1	7	25	13	4	49	51.02
Stage 2	0	12	67	9	88	76.14
Stage 3	0	0	3	69	72	95.83
Total	44	46	83	82	255	
PA (%)	84.09	54.35	80.72	84.15	77.65	

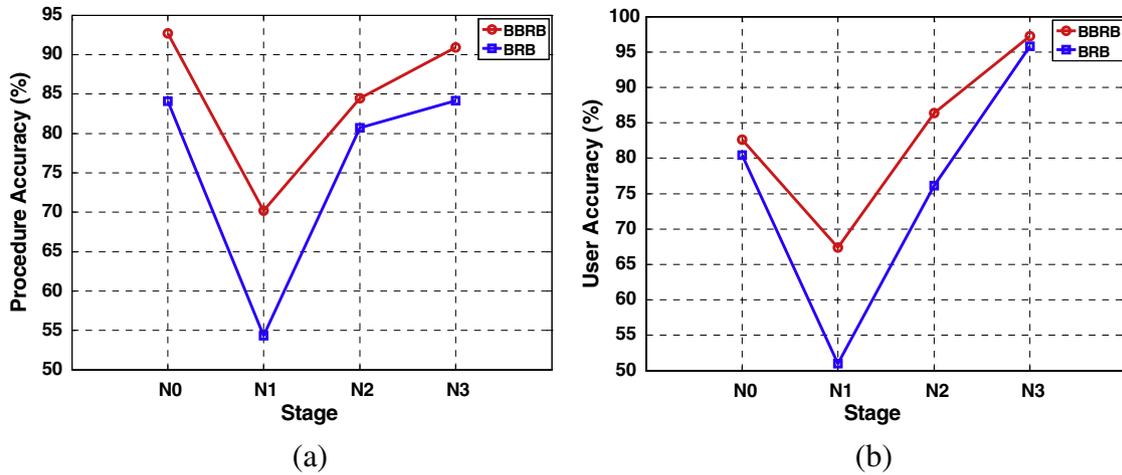


Fig. 3. PA and UA of BBRB and BRB. (a) is the PA and (b) is the UA.

Table 10
Confusion matrix of BBRB-minmax.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	37	8	1	0	46	80.43
Stage 1	3	31	13	2	49	63.27
Stage 2	0	7	73	8	88	82.95
Stage 3	0	0	4	68	72	94.44
Total	40	46	91	78	255	
PA (%)	92.5	67.39	80.22	87.18	81.96	

Table 12
Confusion matrix of BANN.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	35	11	0	0	46	76.09
Stage 1	8	21	17	3	49	42.86
Stage 2	0	7	74	7	88	84.09
Stage 3	0	0	9	63	72	87.50
Total	43	39	100	73	255	
PA (%)	81.39	53.84	74.00	86.30	75.69	

Table 11
Confusion matrix of BBRB-pre.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	34	12	0	0	46	73.91
Stage 1	3	43	0	3	49	87.76
Stage 2	0	76	0	12	88	86.36
Stage 3	0	2	0	70	72	97.22
Total	37	133	0	85	255	
PA (%)	91.89	32.33	0	87.5	57.65	

Table 13
Confusion matrix of BSVM.

	Stage 0	Stage 1	Stage 2	Stage 3	Total	UA (%)
Stage 0	34	11	1	0	46	73.91
Stage 1	3	25	18	3	49	51.02
Stage 2	0	5	76	7	88	86.36
Stage 3	0	2	8	62	72	86.11
Total	37	43	103	72	255	
PA (%)	91.89	58.14	73.79	86.11	77.25	

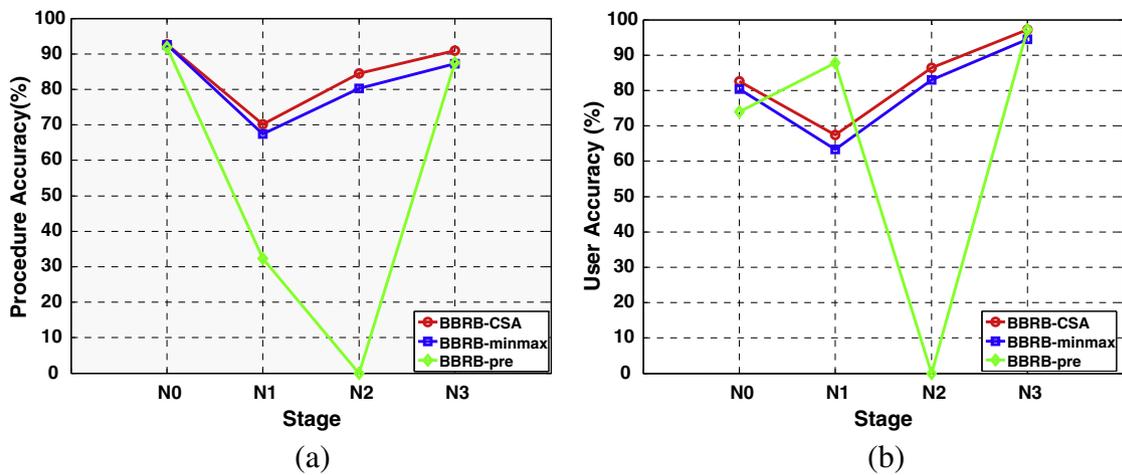


Fig. 4. PA and UA of three conditions. (a) is the PA and (b) is the UA.

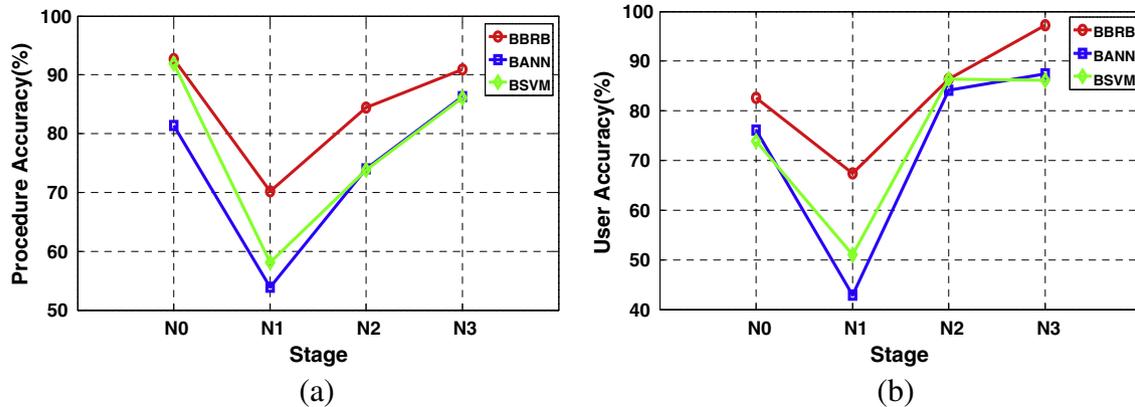


Fig. 5. PA and UA of three CDSSs. (a) is the PA and (b) is the UA.

are shown in Table 6 and two medical antecedent attributes are described in Table 7.

6.2. Validation of BBRB²

In this study, the proposed CSA is utilized for training BBRB. In the training process, all the training data can be used in BRB_1. In BRB_2, training data whose original stage is 1 or 2 are utilized. Table 8 shows the CM of BBRB. It can be seen the overall correct rate achieves 85.1%. To demonstrate the performance better, several comparative studies are given in the following subsection.

6.3. Comparative study with single-layer BRB

A single-layer BRB system is used for comparison. The same CSA is utilized for training both single-layer BRB and BBRB. The reference points of BRB are the same as BRB_1 in BBRB. The initial belief degrees in BRB are also given by the statistics of available historical patient data and domain knowledge, and other parameters are the same as BBRB. The way to generate initial population and the number of initial population in BRB are also the same as in BBRB.

Table 9 shows the CM of BRB. From PA and UA, it can be seen that BBRB performs better than BRB. The overall correct rate of BRB is just 77.65%, which is lower than BBRB. By comparing the PA and UA of the two systems in stage 1 and stage 2, BBRB outperforms BRB as well. This shows that BBRB is more effective for distinguishing stage 1 from stage 2 in this case study. To better show the evaluated results, Fig. 3 shows the PA and UA of the two systems. It shows that BBRB performs better in all four stages.

6.4. Comparative study between two optimization algorithms

To show the performance of CSA based training algorithm, we compare CSA (represented as BBRB-CSA) with the minmax method (represented as BBRB-minmax) [42] which are realized by MATLAB optimization toolbox [6]. The initialization of belief degrees and other parameters are the same in both methods. Meanwhile, the diagnostic performance in test set before training (represented as BBRB-pre) is also presented.

Tables 10 and 11 shows the CM of BBRB-mimax and BBRB-pre. Before training, the overall correct rate is only 57.65%. More seriously, the UA and PA of stage 2 are both 0. Therefore, training procedure is very important. Fig. 4 shows the UA and PA of three conditions. BBRB-CSA performs better than BBRB-minmax in all

cases. This may be due to the fact that the proposed CSA overcomes the drawback of conventional methods.

6.5. Comparison among two-layer SVM, ANN and BBRB

As SVM and ANN have been applied to evaluating LNM, the following comparative study is carried out. On the other hand, we develop the same two-layer system for SVM (represented as BSVM) and ANN (represented as BANN) for comparisons because BBRB is the two-layer system. Firstly, the type of SVM in BSVM is LibSVM2.91 [4]. The kernel function is Radial Basis Function (RBF), in which the regularization and kernel parameters are set to $\{2^{-3}, 2^{-2}, \dots, 2^{10}\}$. The highest recognition rate is regarded as the output of BSVM. Secondly, for BANN, the feed forward neural network [2,12] in MATLAB tool box is adopted for two layers. Note that each ANN has a single hidden layer and the number of nodes is 5. Moreover, As ANN is always sensitive to initial parameters, the best results in 10 different runs are considered as the output. The 2-fold-cross validation method is still adopted. Training and test data in each fold is the same among the three CDSSs.

Tables 12 and 13 shows the CM of BANN and BSVM. The overall correct rate of BANN and BSVM is 75.69% and 77.25%, respectively. It is clear that the proposed prototype CDSS is significantly better than BANN and BSVM. Fig. 5 shows the PA and UA of the three CDSSs. In most cases, BBRB can obtain best performance, especially stage 1 and stage 2. This may be partly because the domain knowledge is embedded into the BBRB.

7. Conclusions

In this paper, a Bi-level Belief Rule Based (BBRB) CDSS prototype was proposed for diagnosing LNM in gastric cancer. The proposed prototype CDSS consists of two layers which are both modeled by BRB. Moreover, as the current methods for optimizing BRB are sensitive to initial values and are limited by their weak local searching abilities, a novel CSA strategy was developed, which is also taken as a general method for optimizing BRB. The validation of the proposed prototype CDSS was conducted using a set of real patient data. Based on the developed CDSS prototype and the system validation results, it is concluded that (1) BBRB is more suitable in evaluating LNM than traditional BRB system, especially for distinguishing stage 2 from stage 3. (2) Compared to traditional optimization method, the proposed CSA can improve performance further. (3) A comparative study among available CDSSs demonstrated that BBRB is capable of generating better diagnostic performance.

² The source code can be downloaded from the following link: http://web.xidian.edu.cn/fliu/files/20130830_193849.zip.

The effectiveness of the BBRB prototype CDSS for diagnosing LNM has been validated by a real patient data set. However, this is a preliminary study and its capability still needs to be tested by more case studies. Meanwhile, the proposed prototype CDSS can be generalized to other medical diagnosis problems. The proposed CSA method can also be designed by other EAs. Those need to be studied in future research.

Acknowledgments

The authors thank the editor and two anonymous referees for their constructive comments and suggestions, which have been very help in enhancing the paper quality. The authors would like to thank their collaborators of Beijing Cancer Hospital for their suggestions and impressive work. This research was supported in part by the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (Grant No. B07048), in part by the National Natural Science Foundation of China (Nos. 60970067, 61072139, 61272282, 61173092), in part by the Program for Cheung Kong Scholars and Innovative Research Team in University (No. IRT1170), and in part by the Fundamental Research Funds for the Central Universities (No. K5051203002).

References

- [1] O. Aase et al., Decision support by computer analysis of selected case history variables in the emergency room among patients with acute chest pain, *Europe Heart Journal* 14 (1993) 433–440.
- [2] C. Bishop, *Pattern Recognition and Machine Learning*, Springer Press, 2006.
- [3] C.M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, 1995.
- [4] C.-C. Chang, C.-J. Lin, LIBSVM : a library for support vector machines, *ACM Transactions on Intelligent Systems and Technology* 27 (2011) 1–27.
- [5] Y.W. Chen et al., Belief rule-based system for portfolio with nonlinear cash-flows and constraints, *Europe Journal of Operational Research* 6 (2012) 775–784.
- [6] T. Coleman et al., *Optimization Toolbox For-Use with MATLAB*, The Mathworks Inc, 1999.
- [7] D. Cunningham, Y.J. Chua, East meets west in the treatment of gastric cancer, *New England Journal of Medicine* 357 (2007) 1863–1865.
- [8] A.P. Dempster, A generalization of Bayesian inference, *Journal of the Royal Statistical Society, Series B* 30 (1968) 205–247.
- [9] H. Elfriede et al., Artificial neural network for prediction of lymph node metastases in gastric cancer: a phase II diagnostic study, *Annals of Surgical Oncology* 11 (2004) 506–511.
- [10] A. Gonzalez, F. Herrera, Multi-stage genetic fuzzy systems based on the iterative rule learning approach, *Mathware & Soft Computing* 4 (1997) 233–249.
- [11] Guilan Kong et al., A belief rule-based decision support system for clinical risk assessment of cardiac chest pain, *European Journal of Operational Research* 219 (2012) 564–573.
- [12] M.T. Hagan, M. Menhaj, Training feed-forward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks* 5 (1999) 989–993.
- [13] H.J. Kim et al., Gastric cancer staging at multi-detector row CT gastrography: comparison of transverse and volumetric CT scanning, *Radiology* 236 (2005) 879–885.
- [14] C.L. Huang, K. Yong, *Multiple Attribute Decision Making Methods and Applications. A State-of Art Survey*, Springer-Verlag, Berlin, 1981.
- [15] F.L. Greene et al., *AJCC Manual of Staging of Cancer*, sixth ed., Springer-Verlag, New York, 2002.
- [16] H.S. Kim, Y.T. Im, Expert system for multi-stage cold-forging process design with a re-designing algorithm, *Journal of Materials Processing Technology* 54 (1995) 271–285.
- [17] I.J. Lee et al., Diagnostic performance of 64-channel multidetector CT in the evaluation of gastric cancer: differentiation of mucosal cancer (T1a) from submucosal involvement (T1b and T2), *Radiology* 255 (2010) 805–814.
- [18] L. Lin et al., A decision support system for lower back pain diagnosis: uncertainty management and clinical evaluations, *Decision Support Systems* 42 (2006) 1152–1169.
- [19] L.N. de Castro, F.J. Von Zuben, Learning and optimization using the clonal selection principle, *IEEE Transactions on Evolutionary Computation* 6 (2002) 239–251.
- [20] Maoguo Gong et al., Multi-objective immune algorithm with non-dominated neighbor-based selection, *Evolutionary Computation* 16 (2008) 225–255 (MIT Press).
- [21] Maoguo Gong et al., A population-based artificial immune system for numerical optimization, *Neurocomputing* 72 (2008) 149–161.
- [22] S.P. Monig et al., Staging of gastric cancer: correlation of lymph node size and metastatic infiltration, *American Journal of Roentgenology* 173 (1999) 365–367.
- [23] M. Musen et al., *Clinical decision-support systems*, Biomedical Informatics (2006).
- [24] K. Ohkuma, K. Hiramatsu, Computed tomography for stage of gastric cancer, *Japan Journal of Cancer Clinics* 32 (1986) 1176–1179.
- [25] D.M. Parkin et al., Global cancer statistics, *A Cancer Journal for Clinicians* 55 (2005) 74–108.
- [26] P. Piero et al., Evolutionary algorithms + domain knowledge = real-world evolutionary computation, *IEEE Transactions on Evolutionary Computation* 10 (2006) 256–280.
- [27] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.
- [28] h. Shayeghi et al., Robust modified GA based multi-stage fuzzy LFC, *Energy Conversion and Management* 48 (2007) 1656–1670.
- [29] E.H. Shortliffe, *Computer-Based Medical Consultations: MYCIN*, Elsevier, New York, 1976.
- [30] D.D. Smith et al., Impact of total lymph node count on stage and survival after gastrectomy for gastric cancer: data from a large US-population database, *Journal of Clinical Oncology* 23 (2005) 7114–7124.
- [31] M. Stone, Cross-validatory choice and assessment of statistical predictions, *Journal of the Royal Statistical Society, Series B (Methodological)* 36 (1974) 111–147.
- [32] Y.M. Wang et al., Environmental impact assessment using the evidential reasoning approach, *European Journal of Operational Research* 174 (2006) 1885–1913.
- [33] H.R.J. Warner et al., Iliad: moving medical decision-making into new frontiers, *Methods of Information in Medicine* 28 (1989) 370–372.
- [34] C.W. Wu et al., Nodal dissection for patients with gastric cancer: a randomized controlled trial, *Lancet Oncol* 7 (2006) 309–315.
- [35] C.W. Wu et al., Relation of number of positive lymph nodes to the prognosis of patients with primary gastric adenocarcinoma, *International Journal of Gastroenterology and Hepatology* 38 (1996) 526.
- [36] D.L. Xu et al., Inference and learning methodology of belief-rule-based expert system for pipeline leak detection, *Expert Systems with Applications* 32 (2007) 103–113.
- [37] J.B. Yang, P. Sen, A general multi-level evaluation process for hybrid MADM with uncertainty, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 24 (1994) 1458–1473.
- [38] J.B. Yang, M.G. Singh, An evidential reasoning approach for multiple attribute decision making with uncertainty, *IEEE Transactions on Systems, Man, and Cybernetics- Part A: Systems and Humans* 24 (1994) 1–18.
- [39] J.B. Yang, Rule and utility based evidential reasoning approach for multi-attribute decision analysis under uncertainties, *European Journal of Operational Research* 131 (2001) 31–61.
- [40] J.B. Yang, D.L. Xu, On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 32 (2002) 289–304.
- [41] J.B. Yang et al., Belief rule-base inference methodology using the evidential reasoning approach-RIMER, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 36 (2006) 266–285.
- [42] J.B. Yang et al., The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties, *European Journal of Operational Research* 171 (2006) 309–343.
- [43] J.B. Yang et al., Optimal learning method for training belief rule based systems, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 37 (2007) 569–585.
- [44] Y. Yuan et al., The development and evaluation of a fuzzy logic expert system for renal transplantation assignment: is this a useful tool?, *European Journal of Operational Research* 142 (2002) 152–173.
- [45] X.F. Zhang et al., Surgical treatment and prognosis of gastric cancer in 2613 patients, *World Journal of Gastroenterology* 10 (2004) 3405–3408.
- [46] X.P. Zhang et al., Support vector machine model for diagnosis of lymph node metastasis in gastric cancer with multi-detector computed tomography: a preliminary study, *BMC Cancer* 11 (2011) 1471–2407.
- [47] Z.J. Zhou et al., Online updating belief-rule-based system for pipeline leak detection under expert intervention, *Expert Systems with Applications* 36 (2009) 7700–7709.
- [48] Z.J. Zhou et al., A sequential learning algorithm for online constructing belief-rule-based systems, *Expert Systems with Applications* 37 (2010) 1790–1799.
- [49] Z.J. Zhou et al., A model for real-time failure prognosis based on hidden Markov model and belief rule base, *European Journal of Operational Research* 207 (2010) 269–283.
- [50] Z.J. Zhou et al., New model for system behavior prediction based on belief rule based systems, *Information Sciences* 180 (2010) 4843–4846.
- [51] Z.J. Zhou et al., Online updating belief-rule-bass using the RIMER approach, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 41 (2011) 1225–1243.