GENERATING CONSISTENT FUZZY BELIEF RULE BASE FROM SAMPLE DATA *

JUN LIU[1], LUIS MARTINEZ[2], DA RUAN[3], HUI WANG[1]

[1] *School of Computing and Mathematics, University of Ulster, Northern Ireland, UK*

Email: j.liu@ulster.ac.uk

[2] *Department of Computer Science, University of Jaén, E-23071 Jaén, Spain*

[3] *Belgian Nuclear Research Centre (SCK•CEN), Boeretang 200, B-2400 Mol, Belgium*

A fuzzy rule-based evidential reasoning approach and it corresponding optimization algorithm have been proposed recently, where a fuzzy rule-base with a belief structure, called a fuzzy belief rule base (FBRB), forms a basis in the inference mechanism. In this paper, a new learning method for optimally generating a consistent FBRB based on the given data is proposed. The main focus is given on the consistency of FBRB knowing that the consistency conditions are often violated if the system is generated from real world data. The measurement of inconsistency of FBRB is provided and finally is incorporated in the objective function of the optimization algorithm. This process is formulated as a nonlinear constraint optimization problem and solved using the optimization tool provided in MATLAB. A numerical example is provided to demonstrate the effectiveness of the proposed algorithm.

1. Introduction

In order to extend the fuzzy logic framework [1] to cover credibility uncertainty as well, a Fuzzy Rule-Based Evidential Reasoning (FURBER) approach was proposed in [2] by combining fuzzy logic and Dempster-Shafer (D-S) theory of evidence [3, 4], which is mainly based on a generic Rule-base Inference Methodology using the Evidential Reasoning approach (RIMER) developed in [5]. In the FURBER framework, a fuzzy rule-base designed on the basis of a belief structure, called a Fuzzy Belief Rule Base (FBRB), is used to capture uncertainty and non-linear relationships between the parameters, and the inference is implemented using the evidential reasoning algorithm in [6].

In addition, optimal models for training the elements of general belief rule bases in RIMER have been proposed in [7], which has also been revised and applied into FURBER framework for engineering system safety analysis and leak detection [8, 9]. However, the consistency of generated rule-base are not addressed in the above work, which is important for a rule-based system to exhibit a reliable performance because the inconsistency generally exists in the knowledge itself provided by the domain experts and in the process of knowledge representation and acquisition as well. Hence, the main focus of this

* This work is supported by the Research Project TIN-2006-02121.

paper is to extend the optimal algorithm in [8] in order to generate the consistent FBRB from sample data. The measurement of inconsistency of FBRB is provided and incorporated into the objective function of the optimization algorithm. The paper is organized as follows. Section 2 briefly reviews the FURBER approach. The detailed optimization algorithm is provided in Section 3 including the way on how to measure the inconsistency of FBRB. A numerical example is provided in Section 4 to demonstrate the effectiveness of the proposed algorithm. Conclusions are drawn in Section 5.

2. Fuzzy Rule-Based Evidential Reasoning (FURBER) Approach

This section reviews the FURBER framework in [2]. A belief rule-base is given by $R=\{R_1, R_2, \ldots, R_L\}$, where the k^{th} rule can be represented as follows:

> R_k: *IF* U is A^k *THEN* D with belief degree β^k, with a rule weight θ_k and attribute weights $\delta_{k1}, \delta_{k2}, \ldots, \delta_{kT}$.

This is the vector form of *a belief rule*. Here U represents the antecedent attribute vector $(U_1, \ldots, U_T)$, A^k the packet antecedents $\{A_1^k, \ldots, A_T^k\}$ (A_i^k ($j=1,\ldots, T$) is the linguistic value of the j^{th} antecedent attribute in the k^{th} rule), T the number of antecedent attributes used in the rule, D the consequent vector $(D_1, \ldots, D_N)$, and β^k the vector of the belief degrees $(\beta_{1k}, \ldots, \beta_{Nk})$ for $k \in \{1, \ldots, L\}$. β_{ik} measures the degree to which D_i is the consequent if the input activates the antecedent A^k in the k^{th} rule for $i=1,\ldots, N$, $k=1,\ldots, L$. L is the number of rules and N is the number of possible consequents. The rule-base can be summarized using a belief rule expression matrix shown in Table 1.

Table 1. A belief rule expression matrix.

Output \ Input	A^1 (w_1)	A^2 (w_2)	…	A^k (w_k)	…	A^L (w_L)
D_1	β_{11}	β_{12}	…	β_{1k}	…	β_{1L}
⋮	⋮	⋮	…	⋮	…	⋮
D_N	β_{N1}	β_{N2}	…	β_{Nk}	…	β_{NL}

In the matrix, w_k is the activation weight of A^k, which measures the degree to which the k^{th} rule is weighted and activated. w_k is calculated as follows:

$$w_k = (\theta_k * \prod_{j=1}^{T} \left(\alpha_j^k\right)^{\bar{\delta}_j}) / (\sum_{t=1}^{L} [\theta_t * \prod_{l=1}^{T} \left(\alpha_l^t\right)^{\bar{\delta}_l}]), \text{ here } \bar{\delta}_j = \delta_j / (\max_{j=1,\ldots,T} \{\delta_j\}), \quad (1)$$

where $A_j^k \in \{A_{js}; s=1,\ldots, S_j\}$ (S_j is the total number of linguistic terms for the attribute U_j) and $\alpha_j^k \in \{\alpha_{js}; s=1,\ldots, S_j\}$ ($j=1,\ldots, T$) is the degree of belief to which the input for U_j belongs to A_j^k of the j^{th} individual antecedent in the k^{th} rule. $\alpha_j^k = A_j^k(x_j)$ is the fuzzy membership degree of a given real input x_j for the attribute U_j to the linguistic term A_j^k. Continuous and differentiable Gaussian function is used in this paper, i.e.,

$$A_j^k(x_j) = \exp\left(-(1/2)\left((x_j - c_j^k)/(\sigma_j^k)\right)^2\right), \quad (2)$$

where c_j^k is the central value of the fuzzy membership function and σ_j^k is the variance at the central value.

Based on the above belief rule expression matrix, we can use the analytical evidential reasoning (ER) algorithm in [7] (which is equivalent to the recursive ER algorithm in [6]) to combine rules and generate final conclusions. The combined degree of belief β_j in D_j is generated as follows:

$$\beta_i = \frac{\mu * \left[\prod_{k=1}^{L} (w_k \beta_{i,k} + 1 - w_k \sum_{i=1}^{N} \beta_{i,k}) - \prod_{k=1}^{L} (1 - w_k \sum_{i=1}^{N} \beta_{i,k}) \right]}{1 - \mu * \left[\prod_{k=1}^{L} (1 - w_k) \right]}, \; i=1,\ldots,N \quad (3)$$

where $\mu = \left[\sum_{i=1}^{N} \prod_{k=1}^{L} (w_k \beta_{i,k} + 1 - w_k \sum_{i=1}^{N} \beta_{i,k}) - (N-1) \prod_{k=1}^{L} (1 - w_k \sum_{i=1}^{N} \beta_{i,k}) \right]^{-1}$.

3. Extended Optimization Algorithm for Generating Consistent FBRB

The performance of inference can be improved if the following parameters in Eq. (3) are adjusted by autonomous learning if they are not given *a priori* or only known partially or imprecisely: (1) rule weight θ_k (k=1,..., L) and attribute weights δ_j (j=1,..., T); (2) the degrees of belief β_{ik} (i=1,..., N; k=1,..., L); (3) the central value c_{js} of fuzzy membership function and the variance σ_{js} at the central value (j=1,.., T; s=1,..., S_j); (4) the utility $u(D_i)$ (i=1,..., N) of the linguistic term of the consequent. Some constraint conditions on each parameter above are detailed in [7, 8], so are skipped here because of space limitation.

3.1 Initial optimization algorithm

The objective of the optimization algorithm is to produces the parameter estimation which minimizes the mean square error criterion defined as

$$\text{MIN } \{\xi\} \quad (4)$$

where, $\xi = \frac{1}{M} \sum_{m=1}^{M} (y_m - \hat{y}_m)^2$, $y_m = \sum_{j=1}^{N} u(D_j) \beta_j(m)$, M is the total number of sample data in the training set, $\hat{y}_m$ is the expected confidence score or the actual output. $(y_m - \hat{y}_m)$ is the residual at the mth point. Eq. (4) is a multi-variable constrained nonlinear single-objective optimization problem and is solved using the existing nonlinear optimization tools in Matlab [10].

3.2 Extended optimization algorithm to minimize the inconsistency

Fuzzy rules are regarded as inconsistent, if they have very similar premise parts, but possess rather different consequents; or they conflict with the expert knowledge or heuristics [11]. Before we discuss the definition of the consistency, we first provide the definition of the similarity of rule premise (SRP) and the similarity of rule consequent (SRC) again with the help of fuzzy similarity measure and similarity measure of discrete probability distribution.

3.2.1 *Similarity measures*

For any two fuzzy sets A and B, the set-theoretic similarity measure usually used in interpretability analysis is in the following form [12]:

$$S(A,B) = (|A \cap B|)/(|A| + |B| - |A - B|), \tag{5}$$

where $|\cdot|$ denotes the cardinality of the set. For the Gaussian fuzzy sets A and B defined in (1), the cardinality calculation becomes an integration as follows [12]:

$$|A| = \int_{-\infty}^{\infty} A(x)dx = \int_{-\infty}^{\infty} \exp\left[-((x - c_A)/\sigma_A)^2\right]dx = \sqrt{\pi}\sigma_A$$

$$|B| = \int_{-\infty}^{\infty} B(x)dx = \int_{-\infty}^{\infty} \exp\left[-\left(\frac{x - c_B}{\sigma_B}\right)^2\right]dx = \sqrt{\pi}\sigma_B, \; |A \cap B| = \frac{\sqrt{\pi}}{2}[2\sigma_{\min} + \Omega],$$

$$\Omega = (\sigma_{\max} - \sigma_{\min})\mathrm{erf}\left(\frac{c_{\max} - c_{\min}}{\sigma_{\max} - \sigma_{\min}}\right) - (\sigma_{\max} + \sigma_{\min})\mathrm{erf}\left(\frac{c_{\max} - c_{\min}}{\sigma_{\max} + \sigma_{\min}}\right)$$

$\sigma_{\max} = \max(\sigma_A, \sigma_B)$ (others are defined similarly), erf(x)= $(2/\sqrt{\pi})\int_0^x e^{-t^2}dt$.

For the comparison functions $m(P, Q)$ for two probability distributions P and Q, we use the Minkowski's (or Euclidean) distance given by

$$m(P,Q) = \sum_{y \in Y} |p(y) - q(y)|^2, \text{ So } S(P, Q)=1-m(P, Q). \tag{6}$$

The more options about m can be found in [13].

3.2.2 *Consistency measure of FBRB*

Consider two rules in the rule base:

R_i: IF $\boldsymbol{U}_1$ is A_1^i and … and $\boldsymbol{U}_T$ is A_T^i THEN $\boldsymbol{D}$ is $\{(D_1, \beta_{1i}), \ldots, (D_N, \beta_{Ni})\}$,

R_k: IF $\boldsymbol{U}_1$ is A_1^k and … and $\boldsymbol{U}_T$ is A_T^k THEN $\boldsymbol{D}$ is $\{(D_1, \beta_{1k}), \ldots, (D_N, \beta_{Nk})\}$.

Then SRP of these two rules is defined as follows:

$$\mathrm{SRP}(i,k) = \min_{j=1}^{T} S(A_j^i, A_j^k), \text{ where } S(A, B) \text{ is defined in (5).}$$

The SRC of these two rules is defined as follows:

$$\mathrm{SRC}(i,k) = S[\{(D_1, \beta_{1i}), \ldots, (D_N, \beta_{Ni})\}, \{(D_1, \beta_{1k}), \ldots, (D_N, \beta_{Nk})\}],$$

where S is defined in (6).

Then the consistency of rule R_i and R_k is defined by [11]:

$$\mathrm{Cons}(R_i, R_k) = \exp\left\{-(\mathrm{SRP}(i,k)/\mathrm{SRC}(i,k) - 1.0)^2/(1/\mathrm{SRP}(i,k))^2\right\}. \tag{7}$$

A degree of inconsistency of a rule base is suggested based on the consistency index provided in (7). At first, a degree of inconsistency for the ih rule is calculated as follows:

$$\mathrm{Incons}(i) = \sum_{\substack{1 \le k \le L \\ k \ne i}} [1.0 - \mathrm{Cons}(R_i^1, R_k^1)] + \sum_{1 \le l \le M} [1.0 - \mathrm{Cons}(R_i^1, R_l^2)], \; i=1,\ldots,L, \tag{8}$$

where R^1 and R^2 denote the rule base generated from data and the rule base extracted from prior knowledge, L and M are the rule numbers of R^1 and R^2 respectively. The degree of inconsistency of each rule is then summed up to indicate the degree of inconsistency of a rule base:

$$\xi_{\text{Incons}} = \sum_{i=1}^{L} \text{Incons}(i)\,, \tag{9}$$

which can be incorporated in the objective function of the algorithm.

3.2.3 Extended optimization formulation

Combining the inconsistency indices, the quality of a generated FBRB is evaluated with the following objective function in order to minimize the mean square error criterion and also minimize the inconsistency level:

$$\text{MIN } \{\xi + \lambda \xi_{\text{Incons}}\},$$

where ξ is the same as in (4), ξ_{Incons} is provided in (9). λ is a weighting constant to control the consistency level.

4. A Numerical Example

An example for oil pipeline leak detection in [9] is used here. We use the data to train a FRBB system for detecting and estimating the leak sizes. The difference between inlet flow and outlet flow, denoted by FlowDiff, and the average pipeline pressure change over time, denoted by PressureDiff, are the two important factors in detecting a leak in the pipeline. A sample rule could be:

IF FlowDiff is Negative Medium AND PressureDiff is Negative Large THEN LeakSize is {(High; 0.2), (Very High, 0:8)}

During the leak trial, 2008 samples were collected for training purpose and 1135 data for test. The test results based on the trained FRBB with and without considering the consistency of the FRBB, are given in Figs. 1 and 2 respectively.

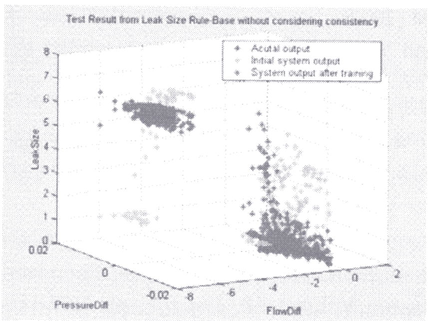

Fig. 1 Test results without consistency

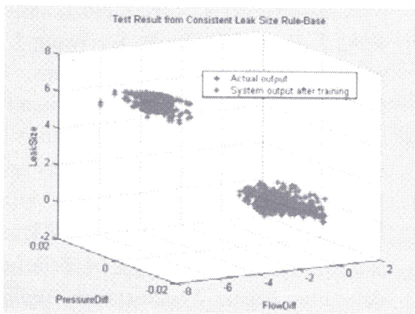

Fig. 2 Test results with consistency

It demonstrates that the estimated outcomes match the observed ones very closely. And the accuracy of the test output has been effectively improved due to minimize the inconsistency in an optimal way. The consistency of the system is violated because of the noise data, which may be caused by turbulence and dynamic changes in the pipeline, and possible instrument and data communication errors. Such noise is intrinsic to almost all pipeline operation data and poses significant challenges to developing pipeline leak detection systems. To minimize the inconsistency can be an effective way to minimize those noises.

5. Conclusions

An optimization method for generating consistent fuzzy rule base with the belief structure was proposed, which provided a practical and reliable support for the proposed FURBER approach, the detailed case study for application in engineering, e.g., offshore safety and risk estimation or oil pipeline leak detection, will be provided in the extended version of this paper.

References

[1] Zadeh, L.A., Fuzzy sets, *Information and Control*, Vol. 8, pp. 338-353 (1965).

[2] Liu J., Yang J.B., Wang J., Sii H.S., Wang Y.M., Fuzzy rule-based evidential reasoning approach for safety analysis, *International Journal of General Systems*, 33 (2-3), 183-204 (2004).

[3] Dempster A.P., A generalization of Bayesian inference, *Journal of the Royal Statistical Society, Series B*, Vol.30, pp. 205–247 (1968).

[4] Shafer G., *A Mathematical Theory of Evidence*. Princeton, N.J.: Princeton University Press, (1976).

[5] Yang J. B., Liu J., Wang J., and Sii H.S., A generic rule-base inference methodology using the evidential reasoning approach – *RIMER, IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 36, No. 2, pp. 266-285 (2006).

[6] Yang J.B. and Xu D.L., On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics-Part A*, Vol.32, No.3, 289-304 (2002).

[7] Yang J. B., Liu J., Xu D. L., Wang J. and Wang H. W., Optimization models for training belief rule based systems. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans,* Vol. 37, No. 4, pp. 569-585 (2006).

[8] Liu J., Yang J.B., Ruan D., Martinez L., Wang J., Self-tuning of fuzzy belief rule bases for engineering system safety analysis, *Annals of Operations Research*, Vol. 163, No. 1, Pages 143-168 (2008).

[9] Xu D.L., Yang J. B., Liu J., Liu G.P., Wang J., Jenkinson I., and Ren J., Inference and learning methodology of belief-rule-based expert system for pipeline leak detection, *Expert Systems with Applications*, Volume 32, Issue 1, pp. 103-113 (2007).

[10] Coleman T., Branch M.A, and Grace A., *Optimization Toolbox—for Use with Matlab*, The Mathworks Inc., (1999).

[11] Jin Y.C., von Seelen W., and Sendhoff., On Generating FC^3 Fuzzy Rule Systems from Data Using Evolution Strategies, *IEEE Transactions on Systems, Man, and Cybernetics* Part B: Cybernetics, Vol. 29, No. 6, pp. 829-845 (1999).

[12] Hefny H.A., Comments on "Distinguishability quantification of fuzzy sets", *Information Sciences,* 177, 4832–4839(2007).

[13] Annalisa A. *et al.*, Classification of symbolic objects: A lazy learning approach, Intelligent Data Analysis, 10, 301–324(2006).