The Application of NN-DS Theory in Natural Gas Pipeline Network Leakage Diagnosis

The Application of NN-DS Theory in Natural Gas Pipeline Network Leakage Diagnosis

Yang Yu, Yinfeng Wu, Jiangwen Wan, Bin Chen

School of Instrument Science and Opto-Electronic Engineering, Beihang University, Beijing , China 100083

Abstract

For the problems of low precision of conventional leakage diagnosis methods, the paper presents a novel natural gas pipeline network leakage diagnosis approach based on neural network and DS evidence theory. The principle is that different premonition information is dealt with by neural network and the result preprocessed is taken to the two levels information fusion by evidence theory for final diagnosis result. The approach makes full use of redundant and complementary leakage information. Computer simulation validates the approach is feasible. Detection results show that the approach improves the precision of leakage diagnosis and decreases the recognition uncertainty.

Keywords: NN-DS theory; pipeline network; leakage diagnosis; sensor node

1. Introduction

Natural gas supply system is one of the city's lifelines, and its main distribution mode is pipeline transmitting. With the increasing complexity of the pipeline network, the pipeline is getting longer and longer, the major security problems have become more and more prominent. Pipeline leakage and its resulting explosion occur from time to time. Timely and accurately identifying the pipeline leakage is of great significance to the urban security. Several methods have been developed for leakage detection, such as acoustic detection [1-3], negative pressure wave [4-6], flow balanced [7,8], real-time transient model method [9,10]and so on. These diagnosis methods have their own advantages and disadvantages. At present, acoustic detection and negative pressure wave methods have obvious advantages compared with other methods. But it is still difficult to the problem of leakage diagnosis precision with strong noise background and complex conditions.

Now urban natural gas pipeline leakage monitoring system based on WSN is adopted to solve the problems existed, which is called 'PipeWSN'. In the paper, with respect to the fusion of multi-source leakage information in the network, we adopt a leakage diagnosis model based on neural network and DS evidence theory. Neural network is a family of massively parallel architectures that solve difficult problems via the cooperation of highly interconnected but simple artificial neurons. Basically, the processing elements of a neural network are similar to the neurons in the brain, which consist of many simple computational elements arranged in layers [12,13]. Evidence theory is a powerful method for combining accumulative evidence or for changing prior opinions in light of new evidence [15,16]. It can solve the conflict problem caused by different diagnosis results. Combining neural network with evidence theory, the diagnosis approach has complementary advantages. The paper is organized as follows: the basic evidence theory is briefly introduced (Section 2). And the NN-DS theory approach is presented (Section 3). In Section 4, the leakage diagnosis model based on NN-DS is proposed. Experiment analysis is given in Section 6. Finally, the paper is concluded in Section 7.

2. DS evidence combining theory

DS evidence theory was presented by Dempster in 1967. He first put forward the definition of upper and lower bounds of the probability, and the theory was developed by Shafer in 1976. DS evidence theory can deal with the uncertainty problem caused by unknown. Briefly, the evidence theory can be summarized as follows.

Θ is supposed as the frame of discernment, containing N exclusive and exhaustive hypotheses, and 2^{Θ} is denoted as its power set, containing all subsets of Θ. A basic probability assignment(BPA) is a mapping m: $2^{\Theta} \rightarrow [0,1]$ which satisfy the following conditions: $m(\varnothing)=0$ and $\sum_{A\subseteq 2^{\Theta}} m(A)=1$, where $0\leqslant m(A)\leqslant 1 \quad \forall A\in 2^{\Theta}$. m(A) is called the mass of A and represents the degree of belief strictly assigned to A. A subset A with a non-null mass is called a focal element of m. A vacuous BPA has the frame of discernment itself as the only focal element, and will be denoted by $m_0(\Theta)=1$. Θ is sometimes abusively called the ignorance. From a BPA m, two main functions are generally defined from 2^{Θ} to [0, 1]: the belief function (Bel), given by

$$Bel(A)=\sum_{B\subseteq A} m(B) \quad \forall A\subseteq\Theta \tag{1}$$

and the plausibility function (Pl), given by

$$Pls(A)=1-Bel(\overline{A})=\sum_{A\cap B\neq\phi} m_j(B) \quad \forall A\subseteq\Theta \tag{2}$$

The belief function represents the lower limit of the probability and plausibility function provides the upper limit of the probability. The difference between the two functions represents the ignorance.

Evidence can be combined by computing the orthogonal sum using Dempster's rule of combination. Here, two information sources A and B are considered. Let $m_1(A)$ and $m_2(B)$ be basic probability assignment (BPA) given by the sources A and B, respectively. The combination rule is written as

$$m(A)=(m_1\oplus m_2)(A)=\begin{cases}0 & A=\varnothing\\ \dfrac{\sum_{A_i\cap B_j} m_1(A_i)m_2(B_j)}{1-K} & A\neq\varnothing\end{cases} \tag{3}$$

where $K=\sum_{A_i\cap B_j=\varnothing} m_1(A_i)m_2(B_j)$ is called the weight of conflict between m_1 and m_2. For a more detailed analysis of evidence theory, the reader is referred to [14-17].

3. NN-DS Theory

BPA function is a bridge connecting neural network with evidence theory. Neural network plays a role of prior knowledge, it has a stronger generalization capability through learning a large number of samples. The principle of NN-DS theory can be concluded as follows: through calculating each measured value in multi-sensor system by neural network method, normalization output of network could be as the basic probability assignment of each proposition, and final result could be obtained by DS evidence combination rules. At the same time, considering that there are a large number of sensors in multi-sensor system, we can hardly ensure the stability and performance of the system by only one neural network. So sensors are divided into several different groups according to their feature information, and each group constructs a neural network individually. Then network output can be used as the evidence for combining in spatial domain after proper mathematical transition. Because each sub neural network only deals with one aspect, the network structure is a bit simpler and training speed is improved.

4. Leakage diagnosis model based on NN-DS

Pipeline network monitoring system based on wireless sensor network is made up of a great deal of sensor nodes, sink nodes and a control management center. The system adopts a clustering network structure (See fig 1) [11]. Installed in the natural gas pipeline, sensor nodes are responsible for signal collection and data pre-processing. The processing results are sent to sink nodes by a multi-hop way. The system adopts a way of in-tube wireless communication between sensor nodes. Because of stronger computing and storage capacity, sink nodes are installed on the ground. As the cluster leaders, they are responsible for managing and maintaining sensor nodes in their clusters as well as combining multi-source information. Whether there is a pipeline leakage or not is determined by decision results which are sent to the control management center to take precautions.

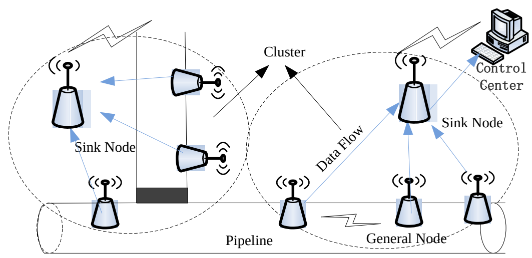

Figure 1 Network structure of monitoring system

Wavelet transform technique is used to preprocess leakage detection signals (acoustic emission and negative pressure wave signals) with leakage information. Leakage characteristic parameters extracted are sent to neural networks of stronger parallel processing capability for initial recognition. The set which is made up of leakage symptom in the course of pipeline network security monitoring is denoted as symptom space, and the set which is made up of state of pipeline security is denoted as recognition space. The course of fault diagnosis by neural network is a non-linear mapping from symptom space to recognition space.

Among all sorts of neural network models, BP neural network has a wide range of applications. When there are a large mass of input data, network structure tends to be more complex. Thus the network requires a great number of training samples which may lead to a slower training speed. While simple structure network only deals with one aspect of a whole question so that training samples could be obtained easily. So two networks of simple structure are adopted to separately deal with sub symptom spaces corresponding to acoustic emission and negative pressure wave signals. The belief degree of each proposition could be obtained from current leakage symptom information.

In order to make full use of information of different symptom spaces, diagnosis information from two BP sub neural networks in one node is regarded as independent evidence. Evidence is combined by evidence theory to deal with diagnosis fuzziness problems caused by single neural network. According to the network structure, there may be several nodes detecting the same leakage event simultaneously. So diagnosis results of nodes in different sites may conflict with each other. In order to avoid that sink nodes couldn’t recognize or make decision, fusion result of single-node is regarded as evidence to be combined again in sink node to improve focality of the correct proposition.

Figure 2 presents a leakage diagnosis model based on BP neural network and DS evidence theory [11]. There are three levels of fusion in the model. In the first level, different symptom information is dealt with by neural network to obtain initial recognition result. In the second level, outputs of neural network regarding as individual evidence are combined in the single-node and sink nodes for two levels fusion by evidence theory. At last, diagnosis result is obtained according to the decision rule.

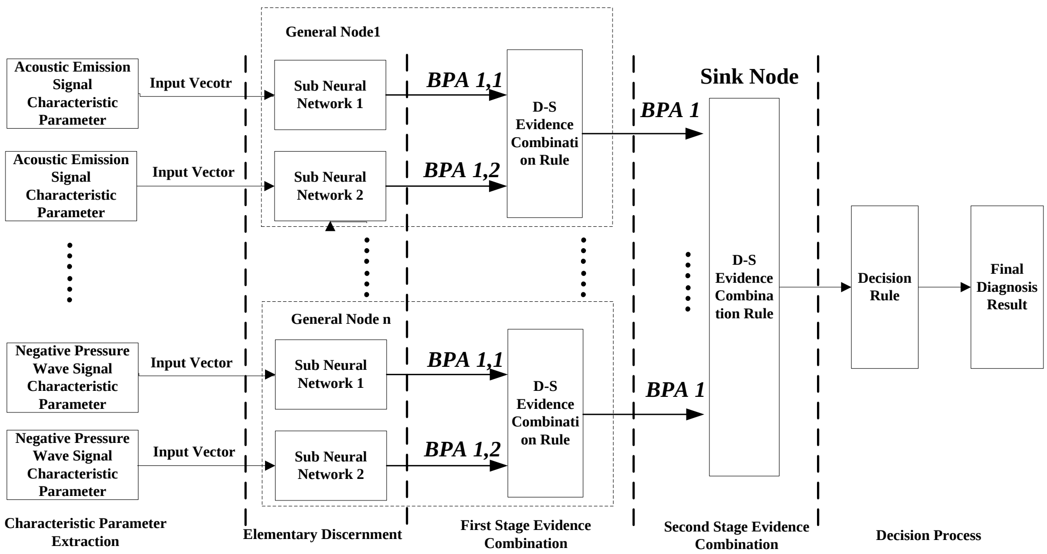

Figure 2 Fusion diagnosis model

5. Algorithm Design

Algorithm design consists of two sections. The first is the structure design of neural networks, and the second is the model design of making decision based on DS evidence theory [11].

5.1 Structure design of neural networks

According to requirement analysis of leakage diagnosis, the number of output nodes is 3. Its aim vector T=[100,010,001], separately represents three different states which are ‘pipeline normal’, ‘small leak’ and ‘gross leak’. Six parameters of mean value, virtual value, peak value, variance, skewness and spectrum maximum amplitude of negative pressure wave signals are extracted as input eigenvectors to constitute sub BP neural network 1 (BPN1), and eight parameters of average amplitude, root mean square, peak frequency, probability density and four frequency band energy distribution factors of acoustic emission signals are extracted as input eigenvectors to constitute sub BP neural network 2 (BPN2).

Determining the number of hidden nodes is a very important section in neural network design. In the paper, an empirical formula is used to calculate the number range of hidden nodes. The number of neurons in hidden layers is variable in BP neural network. Based on the same training precision, training samples and testing samples, the optimal number of hidden nodes is determined by comparing different training numbers and recognition precision. The empirical formula of optimal number of hidden nodes is given by

$$n_1 = \sqrt{p+q} + \beta \tag{4}$$

Where β=1~10, p presents the number of input nerve cells and q presents the number of output nerve cells.

During learning in neural network, normal BP algorithm may lead to slower convergence as a result of improper learning rate. In the paper, an improved BP algorithm is adopted to change learning rate in the course of training. Weight and threshold formulas are given by

$$\Delta W^d(k) = \gamma \Delta W^d(k-1) - (1-\gamma) a s^d (a^{d-1})^T \tag{5}$$

$$\Delta b^d(k) = \gamma \Delta b^d(k-1) - (1-\gamma) a s^d \tag{6}$$

Where γ is the momentum factor, the general value is 0.95; a is the learning rate; $W^d(K)$, $b^d(K)$, s^d respectively denote weight vector, threshold vector and sensitivity function in level d. Under the premise of algorithm stability, learning rate has been improved after joining the momentum term, instead of generating oscillation which may lead to non-convergence.

5.2 Model design of making decision

In the paper, frame of discernment of pipeline leakage $\Theta=\{A_1,A_2,A_3\}$. A_1 represents normal, A_2 represents small leakage, and A_3 represents gross leakage. Let $Y_i(j)$ be output j (j=1,2,3) of BPN i (i=1,2), and let α_i denote its network diagnosis reliability coefficient. A_i denotes reliability coefficient of neural network. The empirical values of α_1 and α_2 are 0.88 and 0.92 by a great deal of experiments.

In the first level evidence combination, outputs of BPN1 and BPN2 are regarded as evidence of frame of discernment after normalization processing. BPA function of evidence i is given by

$$m_i^1(A_j) = \frac{\alpha_i Y_i(j)}{\sum_{j=1}^{q} Y_i(j)} \tag{7}$$

$$m_i^1(\Theta) = 1 - \alpha_i \tag{8}$$

where $Y_i(j)$ presents output j of BPN i. Combination result of each single-node is regarded as the BPA function in the second level evidence combination.

Suppose there are N nodes in the cluster. Evidence set $\{m_\ell^2(A_j)\}$ $(\ell = 1,2,..N)$ is obtained after the first evidence combination. In the course of pipeline leakage diagnosis, the nearer the leakage points to the sensor node, the stronger the signal of leakage character and the higher the reliability of diagnosis result. On the contrary, diagnosis result tends to be unreliable because of disturbance in diagnosis signals. In order to reduce the effect of combination results resulting from unreliable evidence and increase the focality of correct propositions in the set, evidences are combined by the Dempster formula after preprocessing the evidence set. Evidence pretreatment method based on range is given as follows: suppose reliability of the node nearest to leakage point is 1, reliability of other nodes is given by

$$\begin{cases} d = \min(d_\ell) & \ell = 1,2,...,N \\ C_\ell = (d/d_\ell)^\tau & 0 < C_\ell \le 1 \end{cases} \tag{9}$$

where τ is the influence factor. Note that τ=5 by computer simulation. Mass function of each piece of preprocessed evidence is given by

$$\begin{cases} m_\ell^*(A_j) = C_\ell m_\ell^2(A_j), A_j \ne \varnothing \\ m_\ell^*(\Theta) = 1 - \sum_{A_j \subset \Theta} m_\ell^*(A_j) \end{cases} \tag{10}$$

Suppose $\exists A_1$、$A_2 \subset \Theta$, Ω denotes the unknown proposition and $m(A_1)=\max\{m(A_k), A_k \subset \Theta\}$, $m(A_2)=\max\{m(A_k), A_k \subset \Theta$, 且 $A_k \ne A_1\}$. If fulfilling the relationship given by

$$\begin{cases} m(A_1) - m(A_2) > \varepsilon_1 \\ m(\Omega) < \varepsilon_2 \\ m(A_1) > m(\Omega) \end{cases} \tag{11}$$

where ε_1 and ε_2 are threshold values. ε_1=0.3 and ε_2=0.3 after a large number of simulation experiments. As can be seen from the relationship, the supporting proposition has the largest reliability which is greater than that of unknown.

6. Experiment analysis

120 groups of sample data are selected to train BPN1 and BPN2, where training precision is 1 and learning rate is 0.95. Transfer function logsig and purelin are respectively defined for hidden-layer and output-layer. After network training, 120 groups of test samples are inputted to BPN1 and BPN2. Output results are normalized for testing network leakage detection according to the decision rule [11].

On the basis of the empirical formula, hidden neuron numbers of two sub network are all in the range from 4 to 13. Training times and false accept rate (FAR) of BPN1 and BPN2 under different hidden node numbers could be calculated by Matlab simulation at the same training precision. Table 1 shows the result.

Table 1 Training times and FAR of different hidden layer neuron number

	BPN1		BPN2	
Num	Training time	FAR	Training time	FAR
4	5.63×10^4	6%	3.71×10^4	5%
5	7.12×10^4	6%	2.85×10^4	3%
6	0.84×10^4	4%	2.27×10^4	7%
7	4.05×10^4	10%	1.92×10^4	6%
8	0.96×10^4	9%	2.13×10^4	6%
9	2.15×10^4	13%	3.85×10^4	9%
10	3.52×10^4	5%	2.82×10^4	7%
11	1.21×10^4	6%	1.39×10^4	3%
12	1.98×10^4	7%	2.37×10^4	10%
13	1.15×10^4	4%	1.64×10^4	8%

According to testing results, structure parameter of BPN1 is determined to be '6-6-3' and the parameter of BPN2 is '8-11-3'. In this state, the network has faster convergence rate as well as higher recognition rate. As seen from table 1, recognition precision exist difference in different sub neural networks.

Suppose the number of sensor nodes detected is 5 when there is a small leakage somewhere in the pipeline. The ranges between each node to leakage point are 50m, 120m, 155m, 200m and 95m. Table 2 gives the BPA functions after initial recognition by neural network. As seen from table 2, diagnosis results dissatisfy the decision rules. Two main causes may contribute to the condition: differences of testing mechanism and noise sensitivity lead to the deviation of input eigenvector in the network; training samples couldn't include all possible cases when the network training.

Table 2 Diagnosis results of neural network

BPA	A_1	A_2	A_3	Θ
$m_{1,1}$	0.1794	0.4475	0.2532	0.1200
$m_{1,2}$	0.2126	0.6197	0.0876	0.0800
$m_{2,1}$	0.0393	0.4027	0.4381	0.1200
$m_{2,2}$	0.1088	0.3261	0.4851	0.0800
$m_{3,1}$	0.1658	0.6677	0.0465	0.1200
$m_{3,2}$	0.1088	0.3261	0.4851	0.0800
$m_{4,1}$	0.0766	0.6424	0.1611	0.1200
$m_{4,2}$	0.0668	0.7531	0.1001	0.0800
$m_{5,1}$	0.0071	0.4039	0.4690	0.1200
$m_{5,2}$	0.0656	0.7914	0.0630	0.0800

Table 3 and 4 present respectively the first-level and second-level combination results based on initial recognition results.

Table 3 Combining result of the first level evidence

BPA	A_1	A_2	A_3	Θ
m_1	0.1477	0.7338	0.1003	0.0182
m_2	0.0380	0.3764	0.5679	0.0178
m_3	0.0419	0.8265	0.1145	0.0171
m_4	0.0277	0.8996	0.0590	0.0138
m_5	0.0165	0.8276	0.1382	0.0178

Table 4 Combining result of the second level evidence

BPA	A_1	A_2	A_3	Θ
m	0.0013	0.9919	0.0066	0.0001

As seen from table 3, support of proposition A_2 increases to 0.8276 and support of proposition A_3 decreases to 0.1382 in node 5 after the first evidences combining. So the result tends to be reasonable. But in node 2, because $m_2(A_2) < m_2(A_3)$, the evidence is in contradiction with other four pieces of evidence. So sink node is difficult to make decision. According to table 4, support of proposition A_2 adds to 0.9919 and uncertainty declines to 0.0001. The diagnosis result supports proposition A_2 which is 'small leakage'.

7. Conclusion

The paper puts forward a two levels fusion pipeline leakage diagnosis model based on NN-DS theory. Taking initial recognition results as individual evidences, the model realizes the objectivity of BPA. Two-level evidence fusion algorithm could increase the focality of correct propositions, solves the problems of low recognition and difficult making decision in neural network and improves the leakage diagnosis precision. Effectiveness of algorithm is validated by computer simulation and it meets the requirement of pipeline leakage diagnosis in engineering.

8. Acknowledgments

This paper is partially supported by Beijing Municipal Education Commission of China.

References

[1] Fuchs H V, Riehle R. Ten years of experience with leak detection by acoustic singal analysis [J]. Applied Acoustics, 1991, 33(1):1-19.

[2] Ellul I R. Advances in pipeline leak detection [J]. Petroleum Review, 1989, 514(43):582-285.

[3] Pollock A A, Hsu S Y S. Leak detection using acoustic emission [J]. Journal of Acoustic Emission, 1982, 7(4):237-243.

[4] Freitas R, Maitelli A, Salazar A. A wavelet approach to pipeline leak detection by pressure analysis [C]. In: Proceedings of 23rd International Conference onOffshore Mechanics and Arctic Engineering. Vancouver, Canada, 2004, 1-4.

[5] Henrique V, Morooka C, Guilherme I, et al. Leak detection in petroleum pipelines using a fuzzy system [J]. Journal of Petroleum Science and Engineering, 2005, 49:223-238.

[6] Muggleton J M, et al. Wave number prediction of waves in buried pipes for water leak detection [J]. Journal of Sound and Vibration, 2003, 249(5):939-954.

[7] Liou J C P. Leak detection by mass balance effective for Norman wells line [J]. Journal of Oil&Gas, 1996, 3:69-74.

[8] Liou J C P. Mass imbalance errors of water-hammer equations and leak detection [J]. Journal of Fluid. Engineering Transactions of ASME, 1994, 116:103-019.

[9] Kiuchi T. A leak localization method of pipeline by means of fluid transient model [J]. Journal of Energy Resources, 1993, 115(33):162-167.

[10] Verde C. Multi-leak detection and isolation in fluid pipeline [J]. Control Engineering Practice, 2001, 9(6):673-682.

[11] Bin Chen, et al. Pipeline leakage diagnosis method fusing neural network and evidence theory [J]. Journal of Beijing University of Posts and Telecommunications, 2009, 32(1): 5-9.(In Chinese)

[12] Fatih Ozcan, et al. Comparison of artificial neural network and fuzzy logic models for prediction of long-term compressive strength of silica fume concrete [J]. Advance in Engineering Software, 2009, 40:856-863.

[13] Pala M, Oztas A, Yuce MI. Appraisal of long-term effects of fly ash and silica fume on compressive strength of concrete by neural networks. Construct Build Mater 2005; 21(2):384-94.

[14] Buede D.M., Girardi P.A.. Identification comparison of Bayesian and Dempster-Shafer multisensor fusion, IEEE Trans. Syst. Man, Cybernet.-Part A: Syst. Humans 1997,27:569-577.

[15] Dempster A.. Upper and lower probabilities induced by multi-valued mapping, Ann. Math. Statist. AMS-38 1967: 325-339.

[16] G. Shafer. A mathematical theory of evidence[M], Princeton University Press, Princeton, NJ, 1976.

[17] Huiyong Guo, Ling Zhang. A weighted balance evidence theory for structural multiple damage localization [J]. Computer Methods in Applied Mechanics and Engineering, 2006, 195:6225-6238.

Email: sensory@buaa.edu.cn

Phone: (010) 82338089