Viser Technology Pte. Ltd.
Mechanical Engineering Science
DOI: 10.33142/mes.v6i2.15056

Research Article Open Access

Experimental Study on Classification Method of Leakage Signal in Industrial Boiler Pipeline

Tianlong YUAN, Sitong HOU, Xidan ZHANG, Peng LI, Guangchen XU*

Liaoning Provincial Key Laboratory of Energy Storage and Utilization, Yingkou Institute of Technology, Yingkou, Liaoning, 115014, China

***Corresponding Author:** *Guangchen XU, E-mail: 59701706@qq.com*

Abstract

In this paper, a new acoustic emission detection technology for boiler pipeline leakage using random forest and KNN classifier is proposed. The signal parameter index is processed as feature vector, which overcomes the shortcoming of the traditional method which requires a large number of sample data for training and classification. First, the characteristic parameters of boiler pipeline leakage acoustic emission signal are extracted, and then the extracted characteristic parameters are input into random forest and KNN classifier as feature vectors for classification processing. Eight indexes including amplitude, ringing count, duration, energy, rise count, rise time, RMS voltage and average signal level are selected and input into the classifier as feature vectors. As a diagnostic for pipeline leak classification. The experimental results show that this method is effective and feasible in pipeline leak diagnosis, and the feasibility of applying random forest and KNN algorithm to the classification of acoustic emission signals in pipeline leak detection is verified.

Keywords: *pipeline; leakage signal; algorithm; detection*

1 Introduction

Acoustic Emission (AE) refers to the physical phenomenon that when a material is subjected to local deformation or external force, it quickly releases elastic energy and generates transient stress waves, also known as elastic wave emission. At present, according to the characteristics of pipeline leakage detection, the commonly used methods mainly include negative pressure wave method, pressure point analysis method, mass/volume balance method, etc.[1], among which negative pressure wave method has better sensitivity and accuracy, low cost, and can reduce the false alarm rate by using a lower threshold value. However, its shortcomings lie in the fact that it requires sudden and massive leakage. If there is a small slow leak in the pipeline, it is difficult to detect, resulting in a failure of diagnosis[2-3]. Acoustic emission technology has its own advantages compared with traditional non-destructive testing methods: first, the material itself emits defect information, rather than external equipment to supply its energy; Thirdly, the existing leakage can be continuously detected, and the requirement for real-time diagnosis is not high. The leak signal is not required to be detected when the leak just occurs, but can be detected after a period of time when the leak occurs, which greatly improves the convenience and correctness of diagnosis[4]. Acoustic emission technology has been widely used in aerospace, petrochemical, railway, automobile, construction, electric power and other fields. Industrial waste heat boiler works in extremely harsh environment, the medium in the tube is high temperature and high pressure fluid, and the outside of the tube needs to withstand the radiation and convection heat transfer of high temperature flue gas and continuous erosion of ash particles, so it is more prone to leakage accidents. Leakage generally does not suddenly occur in a large area, the initial leak development is slow, non-destructive leakage, after a few days or weeks, to a certain extent will become destructive leakage.

At present, online acoustic monitoring technology is mainly used in China[5]. This method is to open an air acoustic propagation hole on the wall of the furnace tube, weld an air acoustic waveguide on the propagation hole and install an air acoustic sensor. In order to master this advanced detection technology, an acoustic emission sensor was installed on the boiler in the laboratory, and the pipeline leakage and noise were simulated by lead breaking, percussion and sandpaper friction[6]. The acoustic emission signals were classified by random forest and KNN algorithm, and the leakage signals were effectively identified. It lays a foundation for the application of acoustic emission technology in pipeline leakage of industrial boilers[7]. The operation parameters of industrial boiler system are numerous and affect each

Copyright © 2024 by author(s). This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.
Received on September 8th, 2024; Accepted on October 16, 2024

other, and some parameters that can directly reflect the operation of boiler are often difficult to measure. Therefore, we need to find a new fault diagnosis method to monitor the boiler system in real time and get the expected results.

2 Basic Principle of Classifier

2.1 Random forest classifier

2.1.1 Definition of random forest classifier

A random forest is a classifier that uses multiple trees to train and predict samples, uses a recursive method to generate many trees, and then votes on the results to determine the sample category. Each tree is constructed independently by randomly extracting samples from the training data set. Because random forest uses the method of randomly selecting features at each node, the correlation between trees in the forest is reduced, and the error rate is reduced.

2.1.2 The construction process of random forest classifier

(1) Select some samples and features randomly from the sample set;

(2) Train a decision tree using selected samples and features;

(3) Repeat step 1 and Step 2 several times to build multiple decision trees;

(4) For each sample, the classification result of each decision tree is voted, and the classification result with the most votes is selected as the final result.

2.1.3 Advantages and disadvantages of random forest classifier

Advantages:

(1) Random selection of features and samples reduces the risk of overfitting;

(2) It can process high-dimensional data without the need for feature selection;

(3) Can deal with missing values and outliers;

(4) The importance of each feature can be assessed for feature selection and interpretation models.

Disadvantages:

(1) The training time of random forest classifier is longer than that of a single decision tree, so multiple decision trees need to be built.

(2) The model of random forest classifier is complex and difficult to explain.

2.2 KNN classifier

2.2.1 Definition of KNN classifier

K-nearest Neighbor (KNN) algorithm is a commonly used supervised learning method, originally proposed by Trevor Hastie. The idea of the algorithm is that for a given sample d, K Nearest Neighbor samples are found by calculating the distance, and the category of sample d is the category that most of the samples belong to.

2.2.2 Workflow of KNN classifier

(1) Collect data: collect sample data of known categories;

(2) Calculation distance: Calculate the distance between the sample to be classified and the sample of the known class, usually using Euclidean distance or Manhattan distance;

(3) Select K value: Select the nearest K known class samples;

(4) Determine the category: according to the category of K nearest neighbors, determine the category of the sample to be classified. The majority voting method is usually adopted, that is, the category with the most occurrences among the K nearest neighbors is selected as the category of the samples to be classified.

2.2.3 Advantages and disadvantages of KNN classifier

Advantages:

(1) The theory is mature and the thought is simple, which can be used to do classification or regression;

(2) Can be used for nonlinear classification;

(3) The training time complexity is lower than that of other algorithms such as support vector machines, only O(n);

(4) Compared with naive Bayes and other algorithms, there are no assumptions about the data, high accuracy and insensitive to anomalies;

(5) Because KNN method mainly depends on the limited neighboring samples, rather than the method of discriminating the class domain to determine the category, KNN method is more suitable for the sample set with more crossover or overlap of class domains;

(6) The algorithm is more suitable for the automatic classification of class domains with large sample size, and those class domains with small sample size are more prone to misclassification by using this algorithm.

Disadvantages:

(7) Large amount of calculation;

(8) The problem of sample imbalance (i.e., some categories have a large number of samples, while others have a small number of samples);

(9) Requires a lot of memory.

3 Evaluation Standard

For evaluating and detecting the performance of boiler leakage signal system, it is usually divided into True Positive example, False Positive example and true negative example according to the combination of its real category and the prediction category of intrusion detection system.In the four cases of (True Negative) and (False Negative), the Precision is calculated according to the following formula (1) to evaluate the performance of the boiler abnormal detection system.

$$\text{Precision}=\frac{TP}{TP+FP} \tag{1}$$

4 Experiment

First, the boiler data is standardized, and then the features are extracted by empirical wavelet transform, and the parameters of the AR model are estimated by selecting the features with large information content. By adjusting the order p, AR models under different orders are obtained. Finally, various classifiers are used to predict the intrusion data.

Three types of simulated leakage acoustic emission signals were obtained by tapping, sandpaper friction and lead breaking on laboratory boiler pipes. By analyzing the time-domain waveform and frequency spectrum of the three types of acoustic emission signals, random forest and KNN algorithm were used to analyze the characteristic parameters of the leakage signals. Eight indexes, including amplitude, ringing count, duration, energy, rise count, rise time, RMS voltage and average signal level, were selected and input into the classifier as the diagnosis of pipeline leakage classification to verify the feasibility of applying random forest and KNN algorithm to the classification of acoustic emission signals. Specific parameters are shown in Table 1.

In order to verify the stability of the algorithm, the K-fold cross-validation method is used for experiments. The basic idea of K-fold cross-validation is to divide the original data into k groups, take turns taking group K-1 as the training set, and the remaining group as the test set. k tests are repeated to obtain k classification models, and the final results of these k models are averaging, so as to analyze the performance of the algorithm. In this experiment, the value of k is 3. In this chapter, two parameters are used to test the effectiveness of the algorithm, namely the order p of the AR model and the internal parameters of the classifier. For the AR model, the order p is taken from 2 to 8. For a random forest, the parameter RF takes the value 1,2,3,4,5. For KNN, the values of parameter k are 1,2,3,4,5. Figure 1 shows the precision values of various attack types when random forest and KNN classifier are used.

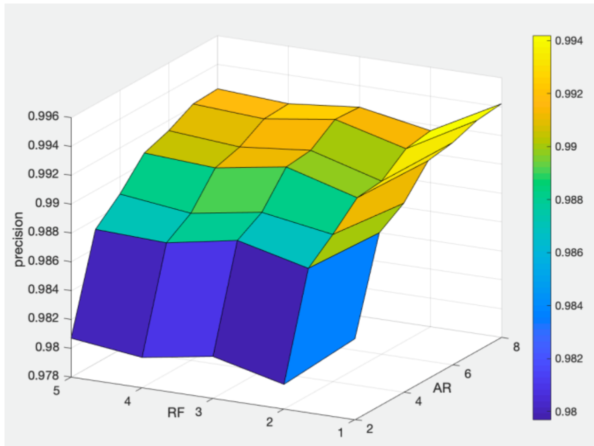

Table 1 The training and testing data

Category	Range	Ringing count	Duration	Energy	Rise count	Rise time	RMS	ASL
1	85.6	1606	74837	22464.3982	84	2006.5	0.761	49.5
1	83.8	1258	50542	10079.3015	116	2116	0.632	46
1	83.8	1525	69005	16824.2737	64	1958.5	0.624	47.7
1	83.9	1073	48143.5	11045.0485	92	2020	0.733	47.2
1	82.9	1381	55881	9138.7009	56	1945.5	0.503	44.3
2	80.1	8086	138328.5	100970.4834	3510	37337	1.206	57.3
2	82	11282	195532	138810.8475	7832	87377.5	1.321	57
2	77	8819	188334	96524.1531	5171	92933.5	0.888	54.2
2	72	8323	148967.5	49133.5358	4917	65165.5	0.531	50.4
2	77.1	12334	196223	93520.4971	8038	96712	0.823	53.6
3	89.8	4455	103773.5	118017.1829	59	474.5	2.503	61.1
3	85.8	4157	91367	68001.5442	61	473.5	1.578	57.4
3	85.3	4464	90805	70743.3914	66	473.5	1.654	57.8
3	86.5	4986	94894.5	89484.1476	812	7328	1.976	59.5
3	86.4	4456	85698	67500.8041	67	473.5	1.663	57.9
1	87.9	1791	84824.5	28623.2269	61	1972	0.93	50.6
1	87.8	1696	80705	27050.116	62	1983	0.912	50.5
2	81.3	10798	176710	106503.9536	6282	74547	1.068	55.6
2	81.7	10451	159592.5	111884.9625	5241	51598	1.205	56.9
3	83.8	4783	85828.5	54494.8898	68	474	1.33	56.1
3	89.5	4963	117936.5	118972.7081	482	17246.5	2.391	60.1

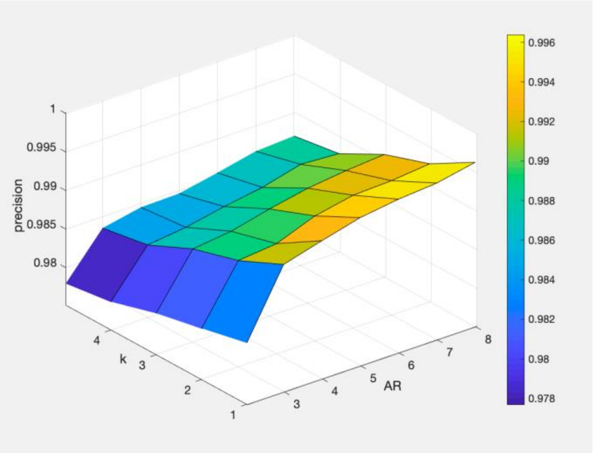

Figure 1 Random forest and KNN classifier accuracy variation with AR

As can be seen from Figure 1, for the random forest classifier, with the increase of AR value, the decrease of k value, the higher the accuracy; For KNN classifier, with the increase of AR value and the decrease of k value, the accuracy is higher. When both the random forest classifier and KNN classifier take k to 8, the results tend to be stable.

5 Conclusion

Acoustic emission testing technology is a dynamic non-destructive testing method, which has many advantages and characteristics that traditional testing methods can not compare. Based on this, this paper proposes to extract the characteristic parameters of acoustic emission signals of boiler pipeline leakage, and then input the extracted characteristic parameters into random forest and KNN classifier as feature vectors for classification processing. 8 indicators are selected to form feature vectors and input into the classifier, so as to successfully identify different kinds of acoustic emission signals. The feasibility of applying this method to the classification of boiler pipeline leakage signals is verified.

Fund Projects: The Foundation of Liaoning Provincial Key Laboratory of Energy Storage and Utilization (CNNK202406), and Yingkou Institute of Technology campus level research project — Development of food additive supercritical extraction equipment and fluid transmission system research.

References

[1] WANG Zhao-hui, ZHANG Lai-bin, XIN Ruo-jia, Application of acoustic emission technique in pipeline leakage detection[J]. Journal of the University of Petroleum China(Edition of Natural Science), 2007,31(5):87-90.

[2] WANG Zhao-hui, ZHANG Lai-bin. Research of small leakage diagnostic technique for liquid delivery pipeline [J]. China Petroleum Machinery, 2003,31(8):37-40.

[3] LIANG Wei, ZHANG Lai-bin, WANG Zhao-hui. State of research on negative pressure wave techniques applied to leak detection in liquid pipelines[J]. Pipeline Technique and Equipment, 2004(6):16-19.

[4] JIAO Jing-pin, HE Cun-fu, Wu Bin. Advance in acoustic emission techniques for pipeline leak detection[J]. Nondestructive Testing, 2003,25(10):519-522.

[5] Ma Liangyu, WANG Bingshu, TONG Zhensheng. Fuzzy recognition and Neural Network Method for Fault Diagnosis of condenser[J]. Proceedings of the CSEE, 2001,21(8):68-73.

[6] Wang Ming-hui.The application of digital temperature sensor DS18B20 in the measuring of the chemical industry[J]. Computers and Applied Chemistry, 2007,24(9):1249-1252.

[7] Yang Hongying, Hua Ke, Ye Hao, et al. Study of the leak diagnosis based on acoustic signal for gas pipelines[J]. Computers and Applied Chemistry, 2010,27(8):1009-1012.