IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 11, 2017, accepted May 4, 2017, date of publication May 10, 2017, date of current version June 28, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2703122

Leak Detection and Location of Pipelines Based on LMD and Least Squares Twin Support Vector Machine

XIANMING LANG[1], PING LI[2], ZHIYONG HU[3], HONG REN[4], AND YAN LI[1]

[1]School of Automation, Northwestern Polytechnical University, Xi'an 710072, China
[2]School of Information and Control Engineering, Liaoning Shihua University, Fushun 113001, China
[3]School of Petrochemical Engineering, Liaoning Shihua University, Fushun 113001, China
[4]China Huanqiu Contracting & Engineering Corp., Shenyang 110167, China

Corresponding author: Xianming Lang (langxianming@mail.nwpu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61673199.

ABSTRACT In oil pipeline leak detection and location, noise in the pressure signal collected at the end of the pipeline affects the accuracy of leak detection and the error of leakage location. To reduce the noise interference, an improved local mean decomposition signal analysis method is proposed. The production functions (PFs) that are related to the leak signal can be exacted, and it is necessary to know the characteristics of leak signals or noise in advance. According to the cross-correlation function, there is a significant peak between the measured signals, which are decomposed into a number of PFs. These reconstructed principal PF components are obtained, and a wavelet analysis is used to remove the noise in the reconstructed signal. On this basis, the signal features are extracted according to the time-domain feature and the waveform feature, which are input into the least squares twin support vector machine (LSTSVM), to recognize pipeline leaks. According to the reconstructed signal after wavelet denoising, the time-delay estimate of the negative pressure signal at the end of the pipeline is obtained by the cross-correlation function, and the leak location is ultimately calculated by combining the time delay with the leak signal propagation velocity. A flow model for pipeline leakage is proposed based on the Flowmaster software, where the collected data of the different working conditions are processed. The experimental results show that the proposed method can effectively identify different working conditions and accurately locate the leakage point.

INDEX TERMS Local mean decomposition, wavelet analysis, least squares twin support vector machine (LSTSVM), leak aperture, leak location, Flowmaster software.

I. INTRODUCTION

Leakage from oil pipelines causes economic losses, contributes to environmental pollution, and is a threat to human health. With the rapid development of oil and gas pipeline networks, leaks may occur frequently due to aging pipelines, corrosion, weld defects and third-party damage [1]. Therefore, it is critical to research on pipeline leakage identification and location to ensure the safe operation of pipelines.

There are many methods for pipeline leak detection and location, which are mainly divided into hardware-based methods and software-based methods [2], [3]. Hardware-based methods primarily require equipment to detect fluid leaks and locate leakage points, such as optical fibers [4], cable sensors [5], magnetic flux leakage [6], etc. However, software-based methods primarily require fluid pressure, flow rate and other signals to detect fluid leaks and locate leakage points, such as negative pressure waves [7], mass balance methods, real-time model methods [8], [9], etc.

The pressure signals collected at the pipeline ends are nonlinear and non-stationary. The features of pressure signals that are the most effective parameters to reflect the characteristics of the various operation conditions are extracted and selected, as they are significant in improving the leakage alarm rate and reducing the false alarm rate [10], [11]. The pressure signals are collected in the pipeline surroundings, as they often contain significant noise and interference, which directly affects the application effect of the negative pressure wave. Therefore, the signals must be noise-reduced before leak detection can occur. If conducting leak signal processing with a traditional de-noising method, small leakage features

will usually be used to filter out noise, thus, the method can cause missed alarms and faulty alarms. Wavelet de-noising can suppress the external disturbances in the collected signals, but this method must choose the appropriate wavelet basis function [12]. Empirical mode decomposition (EMD) methods can self-adaptively decompose the signal into a list of intrinsic mode functions (IMF), each of which highlights the local characteristics of the signal, and the residual components reflect the slow change of the signal. The original signal feature can be extracted by analyzing these components. In reference [13], an improved EMD noise reduction method was proposed without the knowledge of the characteristics of leak signal or noise, and can obtain a more accurate delay time peak and higher leakage location precision. Although the EMD method has been widely applied to various fields, it has many problems such as the boundary effect, mode mixing, and envelope undershoot. For more complex signals, the accuracy of EMD decomposition is reduced.

The LMD method is a new type of non-linear and non-stationary signals processing method, which is similar to the EMD method but is improved compared to EMD in some aspects. The LMD method can be used to adaptively decompose any complicated multi-component signal into a range of product functions (PFs). The PFs component not only preserve more frequency, but also envelope more information than EMD. The LMD method is suitable for analyzing non-linear and non-stationary signals, so it is used as a time-frequency analysis tool for EEG signal processing, machinery fault diagnosis and other fields [14], [15]. According to the report in [16], a small gas leak signal was decomposed by LMD, and then the PFs kurtosis values were calculated. According to these values, the envelope spectrum entropy, which was used as the leakage feature, was input the SVM to select the principle PF components, and the different leak apertures were identified. However, the LMD method is greatly influenced by noise, and is unable to correctly decompose the desired PF component [17].

Features of the vibration signals along the pipeline were extracted, and then the features were input into the SVM to achieve the operating conditions. The results showed that the method has higher recognition accuracy [18]. Recently, the Twin Support Vector Machine [19]–[21], representing one of the new emerging machine learning approaches suitable for both classification and regression problems, has been widely researched. It is four times faster than traditional SVM, and similar to traditional SVM, it has the advantage of small sample learning, and good generalization ability. The TWSVM and improved TWSVM have been used in pattern recognition and classification [22]–[24]. LSTSVM is faster and shows enhanced generalization performance. In place of solving a pair of complex QPPs, LSTSVM generates two non-parallel hyper-planes by solving two linear equations, so the computational complexity is simplified and the operation speed is improved [25], [26].

Leak location belongs to time delay estimation [27]. When pipeline leaks occur, due in part to differential pressure between the inside and outside of the pipeline, the rapid loss of fluid leaks may occur while the pressure drop and the liquid on both sides of the leak point can quickly add to the leak point. This process transmits upstream and downstream in turns, which produces a negative pressure wave (NPW) of a certain speed... To calculate the leak location, we must typically obtain the pressure signal time delay between the upstream and downstream sensors caused by the leak and NPW propagation velocity. The generalized correlation analyses method is adopted for leak location [28].

This paper proposed that pressure signals that are collected at the pipeline ends in the different working conditions can be decomposed into a list of PFs based on LMD. According to the correlation peak analysis of each component with the reference signal, the components with more information are selected as the main PF component, and the pressure signals are reconstructed. Furthermore, the reconstructed signals are de-noised by wavelet analysis, and the features of the signals are extracted according to the characteristics of the time domain and the waveform. The features are input into the LSTSVM to identify pipeline leaks. On the basis of this analysis, the time delay estimation is calculated by the generalized correlation analysis method, and it is adopted for leak location.

II. PIPELINE LEAKAGE FEATURES EXTRACTION BASED ON LMD AND WAVELET ANALYSIS

A. LOCAL MEAN DECOMPOSITION METHOD

Local mean decomposition is used to decompose non-stationary and non-linear signals into a set of PF components, each of which is the product of an envelope signal and a frequency modulated (FM) signal. A detaileds introduction of the LMD algorithm can be found in [29] and [30]. Based on LMD, any complicated signal can be reconstructed according to

$$x(t) = \sum_{p=1}^{k} \mathrm{PF}_p(t) + r_k(t) \tag{1}$$

where k is the number of PF components and $r_k(t)$ is the residue.

It can be seen from expression (1) that the original signal $x(t)$ can be reconstructed by all the PFs and a residual $r_k(t)$, so LMD ensures the information integrity of the complicated signal $x(t)$; thus, LMD contributes by making it efficient to extract useful and comprehensive signal change information.

B. IMPROVED LMD NOISE CANCELLATION

In the existing LMD noise cancellation methods, assuming a desired signal is $\bar{x}(t)$ and a noise $n(t)$, the measured signal is $x(t)$ in expression (2). $x(t)$ is decomposed with LMD according to expression (1), taking into account this representation (1). In fact, in noise cancellation, the goal is to find a $\tilde{x}(t)$ approximately equal to the desired signal $\bar{x}(t)$. When $\bar{x}(t)$ is a low-frequency signal and $n(t)$ is a high-frequency

signal, after removing parts of the high-frequency PFs, $\tilde{x}(t)$ is achieved as expression (3).

$$x(t) = \bar{x}(t) + n(t) \tag{2}$$

$$\tilde{x}(t) = \sum_{p=m_{th}}^{k} \mathrm{PF}_i(t) + r_k(t) \quad (p = m_{th}, \cdots, k) \tag{3}$$

If the desired signal is the low-frequency signal and the noise and the disturbance are the high-frequency signals, we can accurately reconstruct the original signal by removing part of the high-frequency signals.

When the frequency band of the desired signal change and each frequency band contain the desired signal, the method can remove the noise and desired signal. If the desired signal and noise signals are random signals of characteristics of the unknown, the measured signal is used for decomposition and will not be able to distinguish if the PF component contains either the desired signal or the noise signal. Therefore the method cannot effectively select the PF containing the desired signal.

In oil pipeline leak detection and location, the characteristics of leak signals and noise are unknown and are changed by various pipeline states (pipeline sizes, pipeline thickness, materials, etc.) and surroundings. The existing LMD de-noising methods cannot be used for decomposition and reconstruction of measured signals directly due to a lack of knowledge of leak signals and noise, so the methods will not be able to extract the feature of leak signals accurately. An adaptive de-noising method is proposed based on LMD to extract the features of leak signals and noise.

The collected pressure signals at the pipeline ends contain the noise and disturbances which are considered unrelated. When the upstream pressure signal $x_A(t)$ is used as a measured signal, the downstream pressure signal $x_B(t)$ is used as a reference signal (and vice versa). According to the cross-correlation analysis between the reference signal and the PFs of the measured signal, the PFs containing leak signal can be extracted.

The collected pressure signals are $x_A(t)$ and $x_B(t)$, which are denoted by expression (4).

$$\begin{cases} x_A(t) = \alpha\bar{x}(t - \tau_1) + n_1(t) \\ x_B(t) = \beta\bar{x}(t - \tau_2) + n_2(t) \end{cases} \tag{4}$$

where $\bar{x}(t)$ is the leak original signal, α and β are the attenuation factors of the leak original signal propagating along the pipeline, τ_1 and τ_2 are propagation time delay of leak original signals in two collected pressure signals, and n_1 and n_2 are external noises, which are mutually unrelated.

The desired PF components of a measured signal are correlated with the reference signal in the correlation analysis. According to expression (5), the cross-correlation functions between the PFs of the measured signal $x_A(t)$ and the reference signal $x_B(t)$ are calculated, and each PF of the measured signal $x_A(t)$ with the reference signal $x_B(t)$ are cross-correlation analyzed. In other words, each PF of $x_A(t)$ containing the leak signal while $x_B(t)$ containing noise are cross-correlation analyzed. R_{1i} is the cross-correlation function corresponding to the i-th PF. In equation (6), the computational formula of cross-correlations can be used to identify the PFs containing either the components of the leak signal or noises, where b_i^{noise} is the PF containing noise and b_i^{signal} is the PF containing leak signal. Because the cross-correlation function is zero when the PF contains only noise, the corresponding PF will not be selected. If the cross-correlation function is non-zero, the PF will be reserved.

$$R_{1i}(\tau) = \int_0^T b_i(t) x_B(t + \tau) dt \tag{5}$$

$$\begin{cases} R_{1i}(\tau) = \int_0^T b_i^{noise}(t) x_B(t + \tau) dt = 0 \\ R_{1j}(\tau) = \int_0^T b_j^{signal}(t) x_B(t + \tau) dt \neq 0 \end{cases} \tag{6}$$

In adaptive de-noising, the reconstructed signal $\tilde{x}(t)$ is described by

$$\tilde{x}(t) = \sum_i b_i^{signal}(t) \tag{7}$$

In actual working conditions, the cross-correlation function of two unrelated measures of the pressure signals is not absolutely zero. However, at a certain time delay value, the cross-correlation function of two related signals will have a remarkable peak. Therefore the presence of a remarkable peak in a cross-correlation function can be selected by the PFs.

When the cross-correlation function has a remarkable peak, the majority of the data are far less than the peak (maximum) in the cross-correlation function. When the cross-correlation function has no obvious peak, there is no significant difference (they are approximately equal) between the maximum value and minimum value. Based on this distinction, the definition δ is used to select PFs, which can be written as:

$$\delta = \sqrt{\frac{1}{T} \sum_{\tau=1}^{T} \{\max[|R(\tau)|] - |R(\tau)|\}^2} \tag{8}$$

where $\delta < 1$, and when the cross-correlation function has an obvious peak, δ is larger. When the cross-correlation function has not a significant peak, δ is smaller. When $\delta \geq \delta_0$, it means the cross-correlation function has an obvious peak. By selecting threshold δ_0, we can reserve the corresponding PFs that contain leak information signals. In this paper, we chose the threshold $\delta_0 = 0.6$.

C. WAVELET ANALYSIS

Currently, wavelet analysis is considered to be one of the most powerful tools for the analysis of non-stationary signals, as it shows the different characteristics of signals which contain

noise in the time domain and frequency domain. First, the collected signals use wavelet transform to decompose the signal into different time domains and frequency domains. Second, the wavelet coefficients corresponding to noise are removed. Finally, through wavelet analysis, and the signal is reconstructed, and the goal of wavelet de-noising is accomplished. A detailed introduction of wavelet analysis can be seen in [31] and [32].

Although the measured signal is decomposed and reconstructed by the improved LMD, noise is inevitable at some frequency domains. Because wavelet analysis has good performance in both the time domain and frequency domain, in the process of signaling, wavelet analysis is used for de-noising the reconstructed signal. In this way, the noise can be further suppressed, the original signal is restored, and the exact position of the falling edge of the pressure signals at the ends of pipeline is determined.

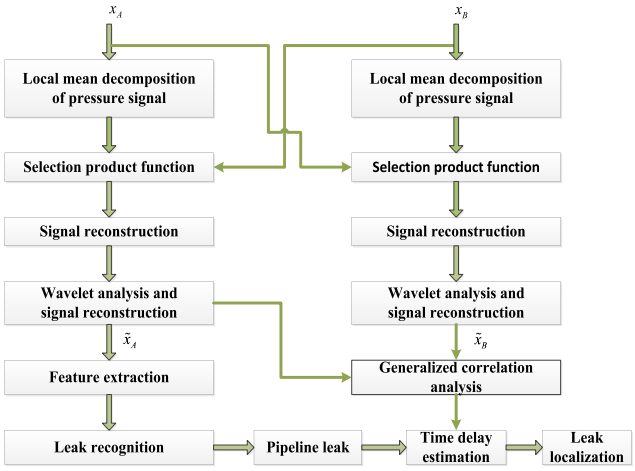

FIGURE 1. Flowchart of the pressure signal analysis and localization based on LMD.

The flowchart of the pressure signal analysis and location based on improved LMD is shown in Fig. 1. x_A is the measured signal, x_B is the reference signal, and $\tilde{x}_A$ and $\tilde{x}_B$ are the reconstructed signals based on improved LMD and wavelet analysis.

D. FEATURES EXTRACTION USING IMPROVED LMD AND WAVELET ANALYSIS

In an oil pipeline, the pressure fluctuation contains abundant information at the pipeline ends under leakage and different working conditions. When the working condition is changed, the amplitude of pressure waveform will change accordingly.

Under normal operation pressure, the pressure signals at the ends of a pipeline show random fluctuation. However, when leakage occurs, the pressure signals drop slowly, and when regulating valves are actuated, the pressure signals show a significantly larger jump [33]. According to this phenomenon, the transient process of pressure drops with different working conditions are selected as the object of feature extraction [34]. In this paper, the feature parameters of the leakage signal are extracted with the time domain feature and waveform feature.

In the time domain feature, when leakage occurs, the time domain waveform of the negative pressure wave signal contains a significant amount of leakage information. The average amplitude can reflect the size of the negative pressure wave caused by leakage, and the effective value can reflect the magnitude of the vibration energy. The average amplitude and the effective value are respectively denoted by expressions (9) and (10).

$$X_{am} = \frac{1}{n}\sum_{i=1}^{n}|x_i| \tag{9}$$

$$E = \sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2} \tag{10}$$

In expressions (9) and (10), X_{am} is the average amplitude, E is the effective value, x_i is the waveform amplitude, and n is data size.

In the waveform feature, kurtosis is used to characterize the distribution of the leakage pressure wave amplitude, and the pulse factor can reflect the feature of the pressure waveform change under sudden leakage. The peak coefficient can also be used to reflect the change of signal amplitude. In the kurtosis, the pulse factor and the peak coefficient can be obtained as follows:

$$X_k = \left(\frac{1}{n}\sum_{i=1}^{n}|x_i|^4\right)^{1/4} \tag{11}$$

$$X_{imf} = \max(|x_i|)/\sqrt{\frac{1}{n}\sum_{i=1}^{n}|x_i|} \tag{12}$$

$$F = \sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}/(\max(x_i) - \min(x_i)) \tag{13}$$

In expressions (11), (12) and (13), X_k is the kurtosis, X_{imf} is the pulse factor, F is the peak coefficient, x_i is waveform amplitude, and n is data size.

Based on the analysis of the negative pressure wave in the time domain feature and the waveform feature, five characteristics are extracted as the feature values of different working conditions.

III. IDENTIFICATION OF PIPELINE LEAKS BASED ON LSTSVM

SVM is a learning algorithm derived from the statistical learning theory, which shows superior performance for dealing with high dimensional and nonlinear problems. SVM is widely used in the field of fault diagnosis. The fault feature is input into SVM for the purpose of fault identification.

Twin support vector machines (TWSVM) generate two nonparallel hyper-planes by solving two smaller-sized Quadratic Programming Problems (QPPs), so that each hyper-plane is closer to one class and as far as possible

from the other. The concept of solving two smaller-sized QPPs replaces a single larger-sized QPP in SVM, which makes the learning of the TWSVM four times faster than the traditional SVM, whereas the Least Squares Twin Support Vector Machine (LSTSVM) uses two non-parallel hyper-planes by solving two linear equations instead of two QPPs as in TWSVM, and shows better generalization performance while being faster than the conventional TWSVM. In this paper, the nonlinear LSTSVM is used to identify the different working conditions.

Consider a data set D containing matrix $A(m_1$ data points that belong to class $+1)$ and matrix $B(m_2$ data points that belong to class $-1)$. Thus, the size of matrix A is $m_1 \times n$ and matrix B is $m_2 \times n$, where n is the dimension of feature space.

The two hyper-planes are obtained by solving a pair of LSTSVM-type QPPs as follows:

$$\min(\mu_1, \gamma_1, \xi)\ \frac{1}{2}\left\|K(A, D^{\mathrm{T}})\mu_1 + e\gamma_1\right\|^2 + \frac{c_1}{2}\xi^{\mathrm{T}}\xi$$
$$s.t.\ -(K(B, D^{\mathrm{T}})\mu_1 + e\gamma_1) = e - \xi \tag{14}$$

and

$$\min(\mu_2, \gamma_2, \eta)\ \frac{1}{2}\left\|K(B, D^{\mathrm{T}})\mu_2 + e\gamma_2\right\|^2 + \frac{c_2}{2}\eta^{\mathrm{T}}\eta$$
$$s.t.\ -(K(A, D^{\mathrm{T}})\mu_2 + e\gamma_2) = e - \eta \tag{15}$$

where $D = [A\ B]$; $c_1, c_2 > 0$; e is a vector of ones of appropriate dimensions; ξ and η are slack variables; and K is an arbitrary kernel function.

The solution of QPPs (14) and (15) can be derived as follows:

$$\begin{bmatrix}\mu_1\\ \gamma_1\end{bmatrix} = -\left(Q^{\mathrm{T}}Q + \frac{1}{c_1}P^{\mathrm{T}}P\right)^{-1}Q^{\mathrm{T}}e \tag{16}$$

$$\begin{bmatrix}\mu_2\\ \gamma_2\end{bmatrix} = -\left(P^{\mathrm{T}}P + \frac{1}{c_2}Q^{\mathrm{T}}Q\right)^{-1}P^{\mathrm{T}}e \tag{17}$$

where $P = [K(A, D^T)\ e]$, $Q = [K(B, D^T)\ e]$. After obtaining μ_1, μ_2, γ_1 and γ_2, two non-parallel hyper-planes can be calculated as follows:

$$K(C, D^{\mathrm{T}})\mu_1 + \gamma_1 = 0 \text{ and } K(C, D^{\mathrm{T}})\mu_2 + \gamma_2 = 0 \tag{18}$$

A new data point $C \in \Re^n$ is assigned to class i using the following decision function.

$$class(i) = \arg\min_{j=1,2} |C\mu_j + \gamma_j| / \|\mu_j\| \tag{19}$$

K is the RBF kernel function $(K(C, D^{\mathrm{T}}) = \exp\left(-\frac{\|C-D^{\mathrm{T}}\|^2}{2\sigma^2}\right))$, which is selected in this paper, where σ is the kernel parameter.

Different working conditions identification of a pipeline is a typical multi-classification problem. In various multi-classification methods, more popular methods include OAA (One-Against-All) and OAO (One-Against-One). Because training data is balanced and the performance is stable in OAA, we chose the OAA method to recognize the different working conditions in a pipeline.

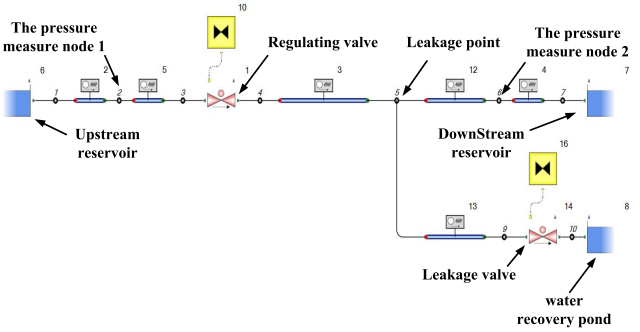

FIGURE 2. Leak model of the pipeline realized in Flowmaster.

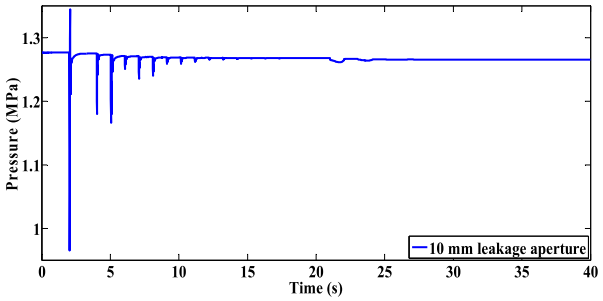

FIGURE 3. Pressure head under valve regulating and 10 mm leakage aperture at upstream.

IV. SIMULATION STUDY

All the methods are implemented in MATLAB R2014a and Flowmaster V7 environments on a PC with an Intel Pentium processor (2.90 GHz) and 6 GB RAM.

A. DATA GENERATION BY FLOWMASTER SOFTWARE

Leakage of the pipeline and the valve action are simulated by Flowmaster software [35], [36]; the establishment of the pipeline model is shown in Fig. 2. When the pipeline model is established, according to the real pipeline environment, the elastic pipeline is used to establish the pipeline model. The length of the pipeline is 1,510 m with a diameter of 0.05 m, the roughness of the inside pipeline wall is 0.025 mm, and the reservoir height of constant head upstream and downstream are 120 m and 0 m, respectively. The pressure wave speed is 1,000 m/s, and the external temperature is 20 degrees Celsius. To emulate leaks, a flow ball valve is positioned at 1,010 m, and the leak aperture can be selected as 10 mm, 5 mm, 3 mm or 1 mm. Two pressure measuring nodes are placed at node 1 and node 2. The sampling rate is 100 Hz, and the simulation time is 40 s. When valve regulating occurs at 2 s and the leakage occurs at 20 s, the pressure changes under different working conditions, as shown in Fig. 3 and Fig. 4. It can be seen from Fig. 3 and Fig. 4 that the amplitude of pressure is significant decreased and then increased rapidly at upstream and the amplitude of pressure is significant increased and decreased rapidly at downstream, when the regulating valve is opened. In 15 seconds, the amplitude of pressures fluctuation are stable at the pipeline ends.

The pressure difference of different leakage apertures is shown in Fig. 5 and Fig. 6. It can be seen from Fig. 5 and

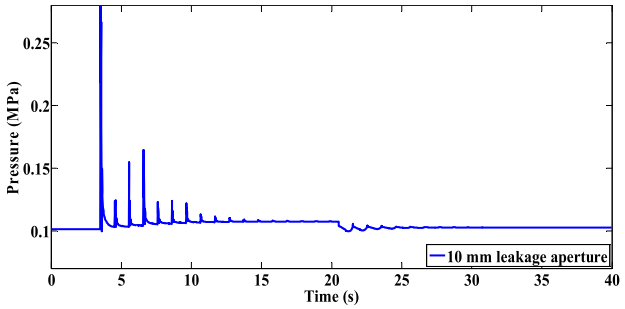

FIGURE 4. Pressure head under valve regulating and 10 mm leakage aperture at downstream.

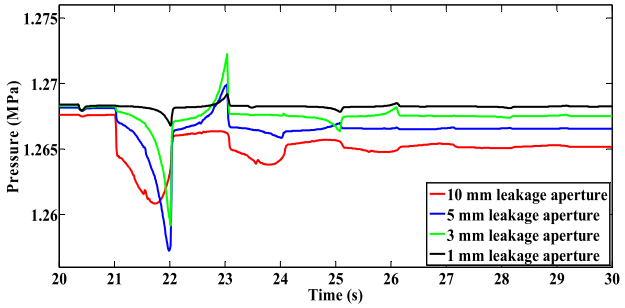

FIGURE 5. Comparison of different leakage apertures at the inlet pressure signal.

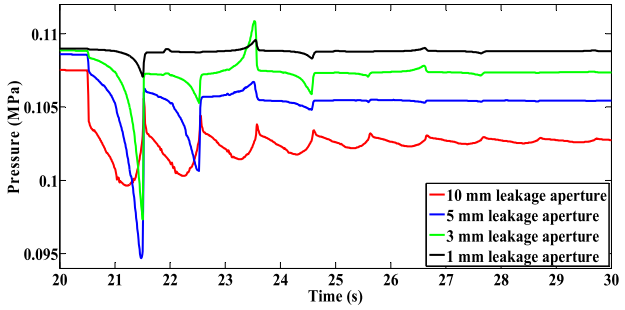

FIGURE 6. Comparison of different leakage apertures at the outlet pressure signal.

Fig. 6 that the amplitude of pressure fluctuation is not obvious with smaller leakage diameters.

B. FEATURE EXTRACTION

To simulate the real working condition, the standard normal distribution random number is added to the pressure data collected at the ends of the pipeline, and thus multiple experimental data is generated. In this paper, 20 groups of experimental data are selected. The Flowmaster software generates noiseless data, but in real conditions, there exists some noise, such as measurement noise and environment noise. Therefore, a normal noise distribution with a zero mean and unit standard deviation is imposed on the data.

The measured pressure signals are decomposed by LMD, and then the reconstructed signals are de-noised based on wavelet analysis (the wavelet basis function is selected by db4, and a 7-layer signal is decomposed). On this basis, the fourth layers are selected for low-frequency reconstruction.

TABLE 1. Features of time domain and waveform.

Pipeline operation condition	Absolute mean value	Effective value	kurtosis	Plus factor	Peak value factor
Normal	0.2892	0.3014	0.3347	3.4583	0.3014
	0.1369	0.1461	0.2337	3.8394	0.1461
	0.1815	0.1973	0.2753	5.5103	0.1973
	0.2299	0.2375	0.2941	4.3502	0.2375
	0.1591	0.1697	0.2437	5.3804	0.1926
Valve regulating	0.8525	0.8551	0.8577	1.1731	0.8551
	0.837	0.8392	0.8418	1.1948	0.8392
	0.8915	0.8952	0.899	1.1217	0.8952
	0.8923	0.894	0.8954	1.1207	0.894
	0.8552	0.8577	0.8601	1.1693	0.8577
10 mm leakage	0.5572	0.5805	0.616	1.7948	0.5805
aperture	0.5545	0.5785	0.6152	1.8035	0.5785
	0.5518	0.5752	0.6114	1.8122	0.5752
	0.5693	0.5877	0.6163	1.7567	0.5877
	0.5558	0.5614	0.5722	1.7992	0.5614
5 mm leakage	0.7328	0.7414	0.7534	1.3647	0.7414
aperture	0.6925	0.701	0.7121	1.4441	0.701
	0.7176	0.7259	0.7368	1.3936	0.7259
	0.763	0.7709	0.7814	1.3106	0.7709
	0.6842	0.6938	0.7063	1.4615	0.6938
3 mm leakage	0.6514	0.6522	0.6534	1.5351	0.6522
aperture	0.6286	0.6378	0.6505	1.5909	0.6378
	0.6157	0.6231	0.6346	1.6241	0.6231
	0.6613	0.6698	0.6824	1.5121	0.6698
	0.6129	0.6208	0.6335	1.6315	0.6208
1 mm leakage	0.574	0.595	0.6289	1.7422	0.595
aperture	0.5499	0.5692	0.6018	1.8184	0.5692
	0.5302	0.5582	0.604	1.8862	0.5582
	0.5259	0.5537	0.5995	1.9016	0.5537
	0.5023	0.5212	0.5537	1.991	0.5212

In this way, the waveform change is beneficial for extracting the feature and locating the leak point. Thus, 20 groups of pressure signals are selected in each working condition, with 5 sets of data among 20 groups, as shown in Table 1.

Based on analyzing six kinds of pressure waveforms with time domain and waveform features, one group of the feature values is calculated, and as shown in Fig. 7.

C. RECOGNITION METHOD BASED ON LSTSVM

The features are different under normal condition, valve regulating, and different leakage apertures. Based on the above analysis, the LSTSVM recognizes different working conditions and the size of leakage apertures of the pipeline. In this simulation, feature vectors (absolute mean value, effective value, kurtosis, plus factor and peak value factor) are the inputs of the LSTSVM. The feature vectors of 84 groups (fourteen of each condition) are chosen as training samples and the others as testing samples. We employ libsvm to implement the multi-classifier of SVM, where the slack

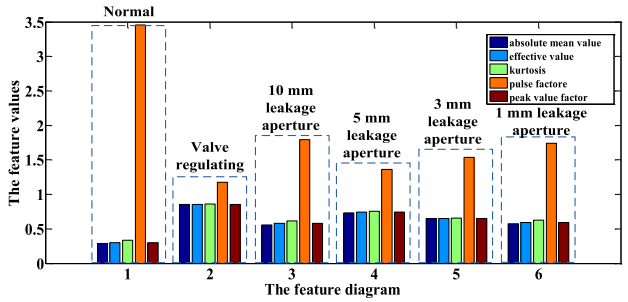

FIGURE 7. Feature values of different working conditions.

TABLE 2. Comparison of the different models' identification results.

Model	Time(s)	Accuracy(%)
SVM	0.95	0.9048
LS-SVM	0.51	0.9048
LSTSVM	0.42	0.9444

variable c is 10 and the kernel parameter σ is 1. We employ LS-SVMlab to implement the multi-classifier of LSSVM, where the slack variable c is 10 and the kernel parameter σ is 1. We set the slack variable $c_1 = c_2$ to 10, and the kernel parameter $\sigma_1 = \sigma_2$ to 1. The OAA method is used to accomplish multi- classification. A comparison of the testing results with SVM and LSSVM are shown in Table 2, where Time is the program's run time.

The classification results of theses 36 testing samples are shown in Table 2. The LSTSVM classification accuracy is 94.44% compared with the SVM or LSSVM classification accuracy of 90.48%. Additionally, less time is needed to run this program.

D. LEAK LOCATION

The LMD method is used to process the different working condition pressure signals at the ends of pipeline, and thus a series of PF components can be obtained. Several PF components are chosen that contain useful leak information by cross correlation as the reconstructed signal, and the reconstructed signal is de-noised by wavelet transform. Figs. 8-15 give the reconstructed results of the pressure signal at the end of the pipeline under 10 mm leakage aperture, 5 mm leakage aperture, 3 mm leakage aperture, and 1 mm leakage aperture, where the blue line represents that the signal is processed by improved LMD and wavelet analysis, and the red dashed line represents that the signal is processed by LMD and wavelet analysis.

It can be seen that the amplitude of the waveform changes gradually with a decrease in the leakage aperture, as shown in Figs. 8-15.

When the leakage aperture reaches 1mm, the inflection point of the waveform is not easily distinguished, and the location error increases accordingly. The signal which is processed by the LMD and wavelet, is smoother than one

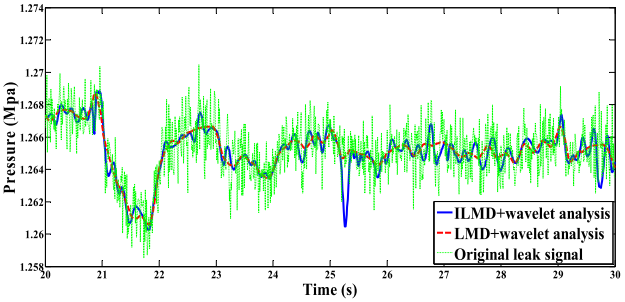

FIGURE 8. Comparison of inlet pressure reconstructed signal of 10 mm leakage aperture.

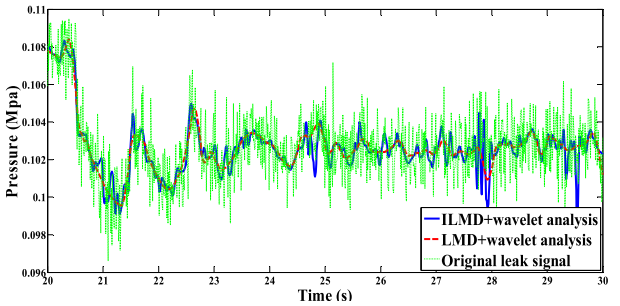

FIGURE 9. Comparison of outlet pressure reconstructed signal of 10 mm leakage aperture.

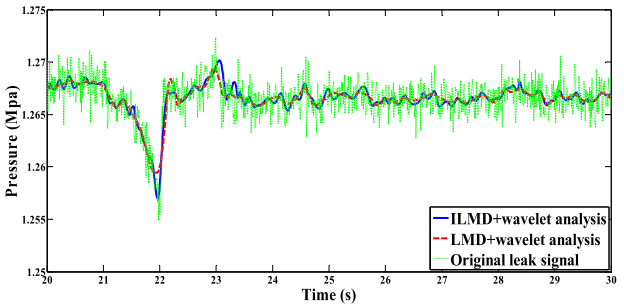

FIGURE 10. Comparison of inlet pressure reconstructed signal of 5 mm leakage aperture.

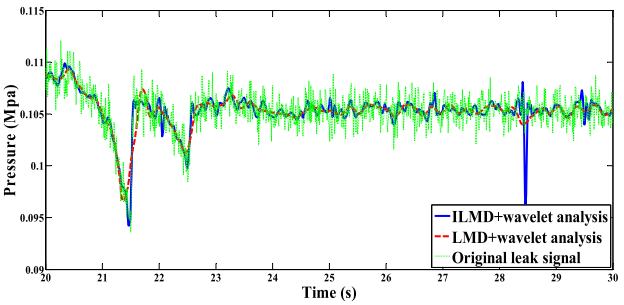

FIGURE 11. Comparison of outlet pressure reconstructed signal of 5 mm leakage aperture.

processed by improved LMD and wavelet analysis, but the signal inflection point cannot be clearly distinguished. Therefore, the method in this paper, can more accurately acquire the leak information than the LMD and wavelet method, which is beneficial for leak recognition and localization.

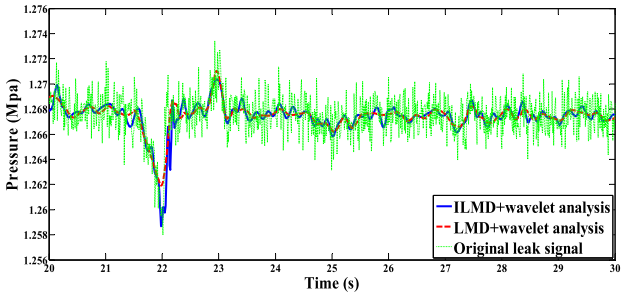

FIGURE 12. Comparison of inlet pressure reconstructed signal of 3 mm leakage aperture.

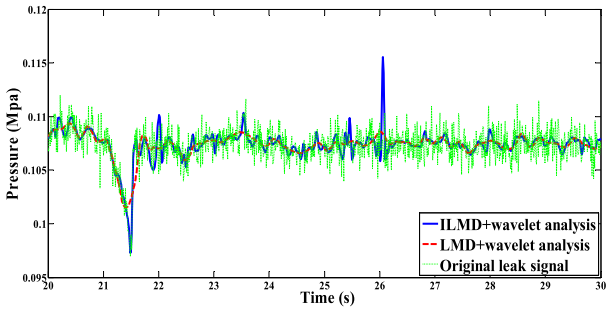

FIGURE 13. Comparison of outlet pressure reconstructed signal of 3 mm leakage aperture.

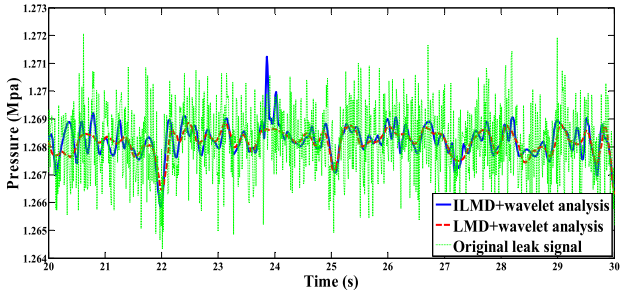

FIGURE 14. Comparison of inlet pressure reconstructed signal of 1 mm leakage aperture.

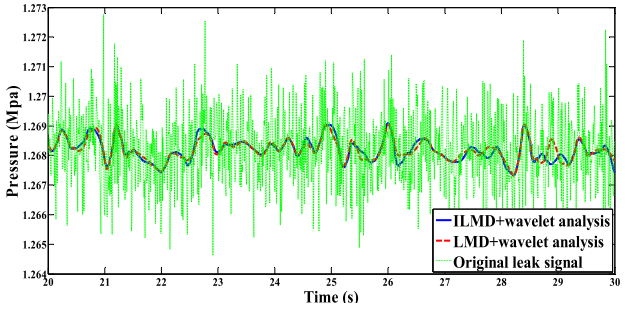

FIGURE 15. Comparison of outlet pressure reconstructed signal of 1 mm leakage aperture.

The generalized cross-correlation function of the processed signals at the ends of pipeline are shown in Fig. 16.

The time difference is calculated by the generalized cross-correlation analysis, and the leak location is calculated as by

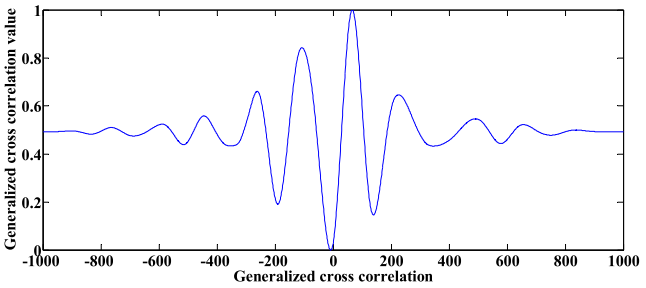

FIGURE 16. Generalized cross-correlation function of the processed signals at the ends of pipeline.

TABLE 3. Comparison of the different methods leakage localization results.

Localization method	Leak aperture	Calculated Localization	Absolute error
LMD and wavelet analysis method	10 mm	1157 m	147 m
	5 mm	1186.25 m	186.25 m
	3 mm	1271 m	261 m
	1 mm	1411.5 m	401.5 m
The method proposed in	10 mm	1051.5 m	41.5 m
this paper	5 mm	1084.5 m	74.5 m
	3 mm	1095 m	85 m
	1 mm	1260.25 m	150.25 m

formula (20). The formula can be obtained as:

$$Lx = (L + v \times \Delta t)/2 \tag{20}$$

where L is the length of pipeline, m; Lx is the distance from the leak point to the upstream sensor, m; v is the negative pressure wave propagation velocity, m/s; and Δt is the time delay estimation, s.

When the leakage position occurs at 1,010 m, a comparison of the location results obtained with the proposed method and the LMD and wavelet method is shown in Table 3, where the calculated localization is the mean of the 20 groups calculated.

Based on the above analysis, the proposed method has a better inflection point than the conventional method. From Table 3, the proposed method can self-adaptively decompose the collected signal, extract the significant feature, and remove the noise. It can effectively locate pipeline leaks with higher accuracy of the time difference than the LMD and wavelet method.

V. CONCLUSIONS

In pipeline leak detection, the non-stationary and non-linear characteristics of the collected signals are generated by different working conditions and influenced by noise. It is difficult to recognize the different leak apertures and locate the leak position. The proposed method can obtain better feature extraction results than LMD after wavelet analysis. The proposed approach is applied to extract the features with time domain characteristic and waveform characteristic under the different working conditions. The chosen feature vectors,

absolute mean value, effective value, kurtosis, plus factor and peak value factor are input into the working condition classifier, and the LSTSVM recognizes different working conditions of the pipeline.

For the purpose of improving the accuracy of time-delay estimation and location, the proposed method is based on LMD and wavelet analysis in order to remove each frequency band of noise and improve the accuracy of the time delay estimation. The simulation results showed that the proposed method can obtain higher accuracy of the time delay, and reduce the location error.

REFERENCES

[1] A. Abdulshaheed, F. Mustapha, and A. Ghavamian, "A pressure-based method for monitoring leaks in a pipe distribution system: A Review," *Renew. Sustain. Energy Rev.*, vol. 69, pp. 902–911, Mar. 2017.

[2] S. Datta and S. Sarkar, "A review on different pipeline fault detection methods," *J. Loss Prevention Process Ind.*, vol. 41, pp. 97–106, May 2016.

[3] P.-S. Murvay and I. Silea, "A survey on gas leak detection and localization techniques," *J. Loss Prevention Process Ind.*, vol. 25, pp. 966–973, Nov. 2012.

[4] Z. Qu, H. Feng, Z. Zeng, J. Zhuge, and S. Jin, "A SVM-based pipeline leakage detection and pre-warning system," *Measurement*, vol. 43, no. 4, pp. 513–519, May 2010.

[5] A. Shui, W. Chen, and Y. Shui, "Improved flexible sensing cable system for leak detection and localization," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4287–4295, Nov. 2013.

[6] M. Afzal and S. Udpa, "Advanced signal processing of magnetic flux leakage data obtained from seamless gas pipeline," *NDT E Int.*, vol. 35, no. 7, pp. 449–457, 2002.

[7] P. Ostapkowicz, "Leak detection in liquid transmission pipelines using simplified pressure analysis techniques employing a minimum of standard and non-standard measuring devices," *Eng. Struct.*, vol. 113, pp. 194–205, Apr. 2016.

[8] J. A. Delgado-Aguiñaga, O. Begovich, and G. Besançon, "Exact-differentiation-based leak detection and isolation in a plastic pipeline under temperature variations," *J. Process Control*, vol. 42, pp. 114–124, Jun. 2016.

[9] J. A. Delgado-Aguiñaga, G. Besançon, O. Begovich, and J. E. Carvajal, "Multi-leak diagnosis in pipelines based on extended Kalman filter," *Control Eng. Pract.*, vol. 49, pp. 139–148, Apr. 2016.

[10] M. M. Gamboa-Medina, L. F. R. Reis, and R. C. Guido, "Feature extraction in pressure signals for leak detection in water networks," *Procedia Eng.*, vol. 70, pp. 688–697, Feb. 2014.

[11] H. Jin, L. Zhang, W. Liang, and Q. Ding, "Integrated leakage detection and localization model for gas pipelines based on the acoustic wave method," *J. Loss Prevention Process Ind.*, vol. 27, pp. 74–88, Jun. 2014.

[12] M. Zadkarami, M. Shahbazian, and K. Salahshoor, "Pipeline leakage detection and isolation: An integrated approach of statistical and wavelet feature extraction with multi-layer perceptron neural network (MLPNN)," *J. Loss Prevention Process Ind.*, vol. 43, pp. 479–487, Sep. 2016.

[13] C. Guo, Y. Wen, P. Li, and J. Wen, "Adaptive noise cancellation based on EMD in water-supply pipeline leak detection," *Measurement*, vol. 79, pp. 188–197, Feb. 2016.

[14] H. Liu and M. Han, "A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings," *Mech. Mach. Theory*, vol. 75, pp. 67–78, May 2014.

[15] M. Han and J. Pan, "A fault diagnosis method combined with LMD, sample entropy and energy ratio for roller bearings," *Measurement*, vol. 76, pp. 7–19, Dec. 2015.

[16] J. Sun, Q. Xiao, J. Wen, and Y. Zhang, "Natural gas pipeline leak aperture identification and location based on local mean decomposition analysis," *Measurement*, vol. 79, pp. 147–157, Feb. 2016.

[17] Z. Liu, Z. He, W. Guo, and Z. Tang, "A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery," *ISA Trans.*, vol. 61, pp. 211–220, Mar. 2016.

[18] L. Ni, J. Jiang, and Y. Pan, "Leak location of pipelines based on transient model and PSO-SVM," *J. Loss Prevention Process Ind.*, vol. 26, no. 6, pp. 1085–1093, Nov. 2013.

[19] D. Tomar and S. Agarwal, "A comparison on multi-class classification methods based on least squares twin support vector machine," *Knowl.-Based Syst.*, vol. 81, pp. 131–147, Jun. 2015.

[20] L. Bai, Z. Wang, Y.-H. Shao, and N.-Y. Deng, "A novel feature selection method for twin support vector machine," *Knowl.-Based Syst.*, vol. 59, pp. 1–8, Mar. 2014.

[21] D. Tomar and S. Agarwal, "Twin support vector machine: A review from 2007 to 2014," *Egyptian Inform. J.*, vol. 16, no. 1, pp. 55–69, Mar. 2015.

[22] J. A. Nasiri, N. Moghadam, and S. Jalili, "Least squares twin multi-class classification support vector machine," *Pattern Recognit.*, vol. 48, pp. 984–992, Mar. 2015.

[23] W.-J. Chen, Y.-H. Shao, C.-N. Li, and N.-Y. Deng, "MLTSVM: A novel twin support vector machine to multi-label learning," *Pattern Recognit.*, vol. 52, pp. 61–74, Apr. 2016.

[24] F. Alamdar, S. Ghane, and A. Amiri, "On-line twin independent support vector machines," *Neurocomputing*, vol. 186, pp. 8–21, Apr. 2016.

[25] Z.-M. Yang, H.-J. Wu, C.-N. Li, and Y.-H. Shao, "Least squares recursive projection twin support vector machine for multi-class classification," *Int. J. Mach. Learn. Cybern.*, vol. 7, no. 3, pp. 411–426, Jun. 2016.

[26] R. Khemchandani and S. Sharma, "Robust least squares twin support vector machine for human activity recognition," *Appl. Soft. Comput.*, vol. 47, pp. 33–46, Oct. 2016.

[27] L. Sun and N. Chang, "Integrated-signal-based leak location method for liquid pipelines," *J. Loss Prevention Process Ind.*, vol. 32, pp. 311–318, Nov. 2014.

[28] C.-H. Ge, G.-Z. Wang, H. Ye, and G.-S. Li, "Leak location based on generalized correlation analysis," *Inf. Control*, vol. 38, no. 2, pp. 194–198, Apr. 2009.

[29] C. Park, D. Looney, M. M. Van Hulle, and D. P. Mandic, "The complex local mean decomposition," *Neurocomputing*, vol. 74, no. 6, pp. 867–875, Feb. 2011.

[30] W. Guo, L. Huang, C. Chen, H. Zou, and Z. Liu, "Elimination of end effects in local mean decomposition using spectral coherence and applications for rotating machinery," *Digit. Signal Process.*, vol. 55, pp. 52–63, Aug. 2016.

[31] G.-M. Xian and B.-Q. Zeng, "An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines," *Expert Syst. Appl.*, vol. 36, no. 10, pp. 12131–12136, Dec. 2009.

[32] M. Han and Y. Liu, "Noise reduction method for chaotic signals based on dual-wavelet and spatial correlation," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 10060–10067, Aug. 2009.

[33] M. Zadkarami, M. Salahshoor, and S. Karim, "Pipeline leak diagnosis based on wavelet and statistical features using Dempster–Shafer classifier fusion technique," *Process Safety Environ. Protection*, vol. 105, pp. 156–163, Jan. 2017.

[34] J. Jiao, Y. Li and B. Wu, "Research on acoustic signal recognition method for pipeline leakage with BP neural network," *Chin. J. Sci. Instrum.*, no. 11, pp. 2588–2596, Nov. 2016.

[35] N. Dutta, C. Rouaud, M. Masera, F. Beuzelin, and C. A. V. Hughes, "Powertrain cooling concept selection process for hybrid electric vehicles," *Innov. Fuel Econ. Sustain. Road Transp.*, pp. 61–72, 2011.

[36] U. Jeong, Y. H. Kim, J.-M. Kim, T.-J. Kim, and S. J. Kim, "Experimental evaluation of permanent magnet probe flowmeter measuring high temperature liquid sodium flow in the ITSL," *Nucl. Eng. Des.*, vol. 265, pp. 566–575, Dec. 2013.

XIANMING LANG was born in Suihua, China, in 1984. He received the M.S. degree in control theory and control engineering from Liaoning Shihua University, Fushun, China, in 2010. He is currently pursuing the Ph.D. degree in control theory and control engineering with Northwestern Polytechnical University. His research interests include leak and location of long distance oil and gas pipeline.

PING LI received the Ph.D. degree in automatic control from Zhejiang University, Hangzhou, China, in 1995. He is currently a Full Professor with Liaoning Shihua University. He has published over 100 papers, successfully completed over 20 research projects supported by national and ministry funds and won more than ten awards of nation and ministry. His research interests include process control and automation, especially the advanced control and optimization of chemical industry process control systems.

ZHIYONG HU received the M.S. degree in oil and gas storage and transportation engineering from Liaoning Shihua University, Fushun, China, in 2009. He is currently a Senior Experimentalist with Liaoning Shihua University. His research interests include safe transportation technology of crude oil pipeline, especially the design and operation of the relevant experimental device.

HONG REN received the bachelor's degree in automatic control from the Fushun Petroleum Institute, Fushun, China, in 1986. He is currently an Assistant Chief Engineer with China Huanqiu Contracting and Engineering Corporation (Liaoning). His research interests include design of process instrumentation and project management.

YAN LI received the Ph.D. degree in automatic control from Nanyang Technological University, Singapore, in 2001. He is currently a Full Professor with Northwestern Polytechnical University. His research interests include robust control theory and fault tolerant control method.