Dongxiu OU
Maojie TANG
Rui XUE
Hongjing YAO

HYBRID FAULT DIAGNOSIS OF RAILWAY SWITCHES BASED ON THE SEGMENTATION OF MONITORING CURVES

HYBRYDOWA DIAGNOSTYKA USZKODZEŃ ZWROTNIC KOLEJOWYCH W OPARCIU O SEGMENTACJĘ KRZYWYCH PRĄDOWYCH

Switches are one of the most important pieces of infrastructure in railway signal systems, and they significantly influence the efficiency and safety of train operation. Currently, the identification of switch failures mainly depends on the experience of railway staff and the use of simple thresholding methods. However, these basic methods are highly inaccurate and frequently result in false and missing alarms. This paper aims to develop a hybrid fault diagnosis (HFD) method for railway switches. The method is an intelligent diagnosis method that uses massive current curves collected by microcomputer monitoring systems. We first divide the switch operation current curves into three segments based on the three mechanical processes that occur during switch operation. Then, a standard curve is selected from the fault-free curves, and common typical faults are ascertained through a microcomputer monitoring system. Finally, derivative dynamic time warping and a quartile scheme are employed to identify fault curves. An experiment based on current curves collected from the Guangzhou Railway Bureau in China demonstrates that the HFD method is extremely accurate and has low false and missing alarm rates. HFD performs better than the studied support vector machine (SVM) and dynamic time warping (DTW) methods, which are widely used for fault diagnosis.

Keywords: *switch system, fault detection and diagnosis, intelligent method.*

Zwrotnice stanowią jeden z najważniejszych elementów infrastruktury systemów sygnalizacji kolejowej i mają znaczący wpływ na wydajność i bezpieczeństwo eksploatacji pociągów. Obecnie, identyfikacja awarii zwrotnic zależy głównie od doświadczenia personelu kolejowego i opiera się na stosowaniu prostych metod progowania. Jednakże te elementarne metody są wysoce niedokładne i często skutkują fałszywymi alarmami lub brakiem alarmu. Niniejszy artykuł ma na celu opracowanie hybrydowej metody diagnostyki błędów (HFD) dla zwrotnic kolejowych. Metoda ta jest inteligentną metodą diagnostyczną, która wykorzystuje wykresy przebiegu prądowego zebrane przez mikrokomputerowe systemy monitorowania. Najpierw krzywe prądowe działania zwrotnicy dzieli się na trzy segmenty w oparciu o trzy procesy mechaniczne, które zachodzą podczas jej działania. Następnie, spośród krzywych opisujących działanie bezusterkowe, wybiera się przebieg standardowy, a w dalszej kolejności ustala się, z wykorzystaniem mikrokomputerowego systemu monitorowania, najczęściej występujące, typowe błędy działania zwrotnicy. Wreszcie, do identyfikacji krzywych błędów stosuje się schemat kwartylowy oraz metodę derivative dynamic time warping wykorzystującą pochodne do klasyfikacji szeregów czasowych. Eksperyment oparty na krzywych prądowych zebranych przez Guangzhou Railway Bureau w Chinach pokazuje, że metoda HFD jest wyjątkowo dokładna i skutkuje niską liczbą fałszywych i brakujących alarmów. HFD daje lepsze wyniki niż szeroko stosowane do diagnozowania błędów metody maszyny wektorów nośnych (SVM) i dynamic time warping (DTW).

Słowa kluczowe: *układ zwrotnicowy, wykrywanie i diagnozowanie usterek, metoda inteligentna.*

1. Introduction

With the rapid development of high-speed rail (HSR) around the world, the current world speed record for a commercial train has reached to 574.8 km/h. Increased train speed is convenient but also causes safety and reliability problems. Track circuits, railway annunciators and switches are generally the three main components that contribute to the operational safety of HSR. Of these three components, switches (Fig. 1), which connect equipment that supports train transit from one track to another, are mainly responsible for the efficiency and safety of HSR. However, switch failures have recently caused several major railway accidents [28]. According to a statistical report by the Jinan Railway Bureau in 2015-2016, 191 switch faults accounted for approximately 60% of signal faults. Thus, early diagnosis of issues with switch systems is critical for the operational safety of HSR.

To achieve the safe operation of HSR, microcomputer monitoring systems (MMSs) have been widely introduced to timely monitor switch states in China [27]. MMSs collect switch operation current and power curves that allow maintenance staff to identify the state of switches and make diagnoses based on their experience. However, a lack of experience can lead to missing or false alarms, both of which pose serious security risks. Furthermore, the number of switch operation curves is relatively large, and many financial and human resources are involved in such work.

Domestic and foreign experts have conducted several studies on fault diagnosis. Early attempts employed simple thresholding methods [16, 21] to detect faults, but frequent false and missing alarms limit the extensive application of these methods. A more recent study [6] summarized three primary approaches in the literature for switch diagnostics: feature, model and empirical methods.

Fig. 1. Railway track switch

For feature-based methods, special features that can be rapidly identified are extracted. Data collection, feature extraction, and feature selection form three subsections of this model. Marquez et al. [17] used data from tests conducted on a commonly found point mechanism and discussed the benefits of adopting a Kalman filter for preprocessing data collected during tests. Eker et al. [6] proposed a support vector machine (SVM) operated through principle component analysis (PCA) for dimensionality reduction to diagnose faults in switches. Six different features were selected, and four remained following a T-test. Asada et al. [2] developed a new approach to fault detection and diagnosis that involved utilizing parameters collected from low-cost and accessible sensors; they focused on fault detection and diagnosis for 'overdriving' and 'underdriving.' Lee et al. [13] introduced a data mining solution that employs audio data to detect and diagnose switch faults. Zhou et al. [27] proposed an improved SVM that accommodates fault detection, and the authors optimized the geometric parameter feature extraction method developed by He [10].

In model-based methods, a model is defined to characterize a system. Deviation from the model is defined as a failure and is identified as the difference between the model outcome and actual data. Eker et al. [7] presented a simple state-based prognostic (SSBP) method for fault detection and forecasting in electromechanical systems. Ardakani et al. [1] established a strategy and technical architecture for the prognostic and health management (PHM) of electromechanical point machines. Zhang et al. [23, 24] proposed a switch fault detection algorithm based on a probabilistic neural network and back propagation neural network. Letot et al. [14] proposed a model for degradation trend assessment and a methodology that updates degradation paths and reliability data to accurately estimate the remaining useful life. Wang et al. [22] proposed a failure prediction model based on a Bayesian network to evaluate the effects of weather patterns on railway switches.

In empirically based methods, a fault-free sample is used as a reference signal, and failures are identified based on the resemblance of a given signal to a reference signal. Atamuradov et al. [3] introduced an expert system based on an economic analysis method that identifies the best maintenance policy for a failure mode and/or system component. Zhao and Lu [26] presented a switch fault diagnosis method based on gray correlation analysis. The authors related the basis of the variations in the power curve to the typical faults of a switch machine. Kim et al. [12] proposed a diagnosis method that involves applying DTW to manage variations in the duration of railway point machine use; this model manages only phase-shifted shape faults, and the parameter δ of DTW chosen by maintenance staff serves as a threshold.

However, the abovementioned methods do not adequately address the problem at hand. For example, the Kalman filter method can achieve success only for a portion of a dataset (reverse to normal). SVM-based methods are sensitive to feature selection, and few authors can explain how to select such features. Artificial neural networks are not suitable for this task, as lacking a sufficient number of fault samples can result in underfitting. In addition, an expert system functions according to large amounts of priori knowledge, thereby requiring a considerable amount of manpower from experienced railway staff. Although DTW performs effectively for shape faults, it cannot detect faults over shorter or longer durations. To overcome these limitations, this paper trains HFD using a small sample dataset, i.e., with a small amount of priori knowledge. In addition, HFD is used to detect and diagnose eleven typical faults summarized by the maintenance staff of the Guangzhou HSR. Moreover, HFD identifies fault curves automatically from a computer and can reduce the quantities of manpower and resources required.

The remainder of this paper is organized as follows. Section 2 introduces switch operation current curves and explains why these curves must be divided into three segments before fault diagnosis. The mathematical principles and calculation processes of HFD are explained in Section 3. Section 4 presents a numerical experiment using real switch operation current curves for fault diagnosis, followed by a discussion and concluding remarks in Section 5.

2. Analysis of switch operation current curves

2.1. Basic analysis of current curves

Although MMSs can collect current and power curves, only current curves have been widely used for fault diagnosis because current values provide an enormous amount of information regarding switches, such as their electrical and mechanical characteristics [25]. Therefore, experienced maintenance staff can identify switch faults by observing various characteristics of current curves. Current curves can be divided into the following three segments based on three mechanical processes: the start stage, action stage and release stage. For example, Fig. 2 shows the fault-free curves of a railway switch. The start stage $(0 - T_1)$ exhibits a peak current when the machine begins to operate; the action stage $(T_1 - T_2)$ is relatively smooth, and it corresponds to the working process of the switch; and the release stage $(T_2 - T_3)$, which is typically called the "small step", indicates that the switch has finished switching and has connected the relevant circuit.

2.2. Fault types and segmented current curves

Through long-term observation and analysis, the maintenance staff of the Guangzhou HSR summarized the fault current curves for the track. Eleven types of faults occurred on the track: abnormal fluctuation, poor contact in the action circuit, abnormal impedance in the action circuit, start failure, conversion failure, release failure, open start-up circuit, electric relay 2DQJ switch failure, blocking in the gap, machine idling, and overlong release time of the starting relay. These faults, referred to as M1-M11, are described in Table 1.

In Table 1, the faults include shape and duration faults. The data associated with these fault modes are different from fault-free data in shape or duration. The fault stage indicates the stage in which a
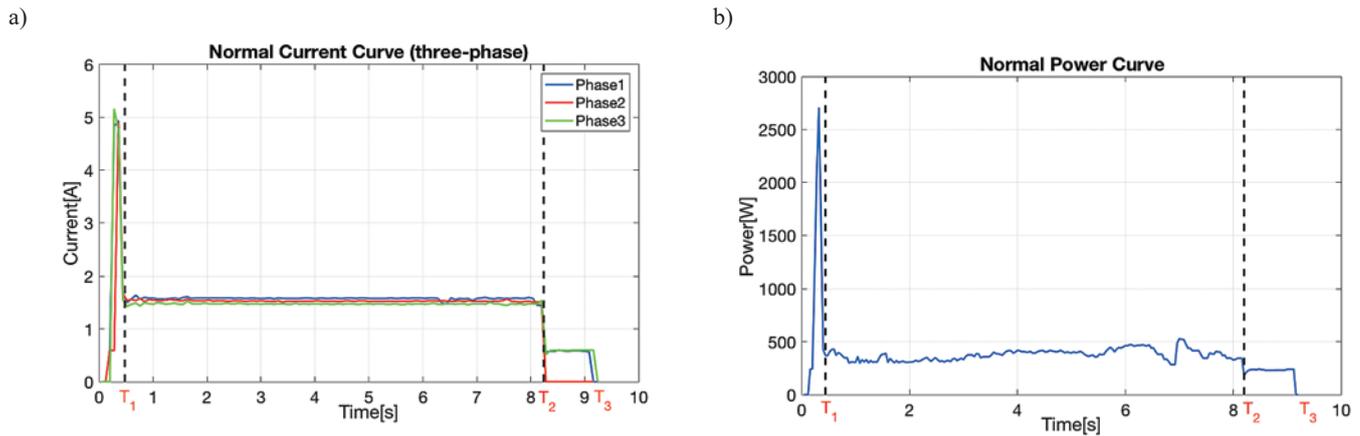
a)

b)

Fig. 2. Switch operation curves

fault occurs. Therefore, the maintenance staff can make rapid fault-solving decisions when the anomalous stage is known.

Currently, the segmenting methods mainly depend on two fixed points to divide current curves into three stages. However, the two fixed points may not apply to all switches. Fig. 3 shows the cumulative switch current curves of Station #1 (Fig. 3a) and Station #2 (Fig. 3b) for January (taking single-phase current data as an example). In Fig. 3, the duration of the current data is approximately 5.5 s for Station #1, and it is 9 s for Station #2. The durations are typically different at all stations, which can be referred to as a "different durations" problem. Therefore, only two adaptive points can divide all current curves into three stages with high accuracy rather than using two fixed points.

3. Model and algorithm for railway switch hybrid fault diagnosis

The proposed HFD method involves the following three steps: fault-free dataset selection, standard curve selection and fault detection and diagnosis. The first step involves dividing samples (current

curves) into three segments and constructing a fault-free dataset; the second step involves selecting the best sample, referred to as the "standard curve," from the fault-free dataset; and the third step involves comparing test samples with the standard curve and other fault types for fault detection and diagnosis. The details of HFD are presented below.

3.1. Fault-free dataset selection

3.1.1. Curve segmentation

In this section, an adaptive mean-shift (AMS) algorithm is used for segmentation [5, 8]. This algorithm iterates by pointing in the direction of the maximum increase in density and involves the following six steps.

- Step 1: Collect a current curve from MMSs, and start with an input $X = [x_1, x_2 \ldots x_n]$.
- Step 2: Choose an arbitrary point as the initial center y_0 from X, a bandwidth h and a kernel function $K(x)$. In AMS, the bandwidth equals σ_X (the standard deviation of X), and the

Table 1. Fault types and corresponding attributes

| Fault Types	Corresponding Curve Characteristics	Fault Modes	Abnormal Stages	Symbols
Abnormal fluctuation	Abnormal fluctuations in the action current	Shape Fault	$T_1 - T_2$	M_1
Poor contact in the action circuit	Abrupt change in the action current	Shape Fault	$T_1 - T_2$	M_2
Abnormal impedance in the action circuit	Conversion current that exceeds the limit	Shape Fault	$T_1 - T_2$	M_3
Start failure	Small step in the action stage	Shape Fault	$T_1 - T_2$	M_4
Conversion failure	Rising current in the action stage	Shape Fault	$T_1 - T_2$	M_5
Release failure	Two peaks exist in the action stage	Shape Fault	$T_1 - T_2$	M_6
Open start-up circuit	Zero value curve	Shape Fault	$T_1 - T_2$	M_7
Electric relay 2DQJ switch failure	A "small steps" curve	Shape Fault	$T_2 - T_3$	M_8
Blocking in the gap	Missing "small steps"	Shape Fault	$T_2 - T_3$	M_9
Machine idling	Overly long conversion time	Duration Fault	$T_1 - T_2$	M_{10}
Overlong release time of the starting relay	Overly long "small steps"	Duration Fault	$T_2 - T_3$	M_{11}

a)

b)
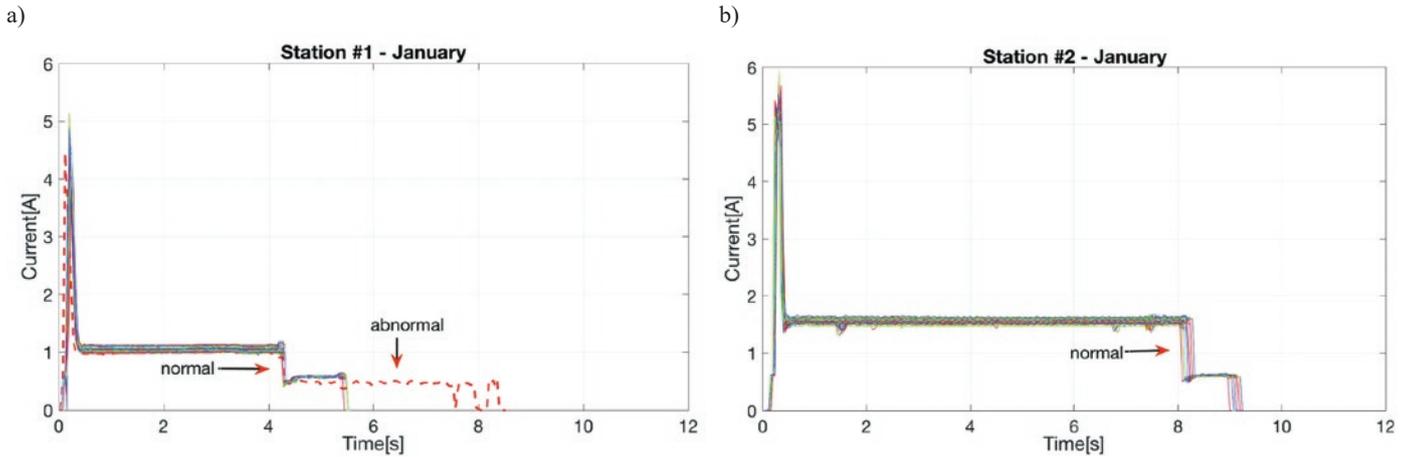
Fig. 3. Cumulative Current Curves

spherical normal kernel [8] function $\tilde{K}(x)$ is coordinated with the bandwidth. The multivariate kernel density estimate $f(x)$ obtained from $\tilde{K}(x)$ and σ_X is:

$$f(x) = \frac{1}{n\sigma_X} \sum_{i=1}^{n} \tilde{K}\left(\frac{x - x_i}{\sigma_X}\right). \tag{1}$$

For radially symmetric kernels, the profile of the kernel $k(x)$ is determined to satisfy:

$$\tilde{K}(x) = c_k k\left(x^2\right), \tag{2}$$

where c_k is a normalization constant that ensures that $K(x)$ satisfies:

$$\int_R \tilde{K}(x) = 1 \tag{3}$$

• Step 3: Calculate the gradient of the density estimate as follows:

$$\nabla f(x) = \frac{2c_k}{n\sigma_X^3} \sum_{i=1}^{n} (x_i - x) g\left(\left|\frac{y_t - x_i}{\sigma_X}\right|^2\right)$$

$$= \frac{2c_k}{n\sigma_X^3} \sum_{i=1}^{n} g\left(\left|\frac{y_t - x_i}{\sigma_X}\right|^2\right) \left(\frac{\sum_{i=1}^{n} x_i g\left(\left|\frac{y_t - x_i}{\sigma_X}\right|^2\right)}{\sum_{i=1}^{n} g\left(\left|\frac{y_t - x_i}{\sigma_X}\right|^2\right)} - x\right) \tag{4}$$

where $g(s)$ is equal to $-k'(s)$ and y_t is the center of the current iteration (t starts at index 0). The first term is proportional to the density estimate at x computed from kernel $G(x) = c_g g\left(x^2\right)$, and the second term is the mean-shift.

$$m_{\sigma_X}(x) = \frac{\sum_{i=1}^{n} x_i g\left(\left|\frac{y_t - x_i}{\sigma_X}\right|^2\right)}{\sum_{i=1}^{n} g\left(\left|\frac{y_t - x_i}{\sigma_X}\right|^2\right)} - x \tag{5}$$

• Step 4: Iterate the mean-shift procedure until convergence is achieved, including the successive computation of the mean-shift vector $m_{\sigma_X}(x^t)$ and the translation of the center $y_{t+1} = y_t + m_{\sigma_X}(x^t)$. This iteration is guaranteed to converge to a point where the gradient of the density function is zero [4].

• Step 5: Divide the points in X that satisfy Equation (6) into one cluster and remove them from X.

$$|x_i - y_t| \leq \sigma_X \quad 1 \leq i < n \tag{6}$$

• Step 6: Return to Step 2 until there are no points in X.

AMS can divide input X into several clusters. The cluster with the largest number of elements is defined as the action cluster. Furthermore, X can be grouped into three segments based on the two elements with the minimum subscript i and maximum subscript j of the action cluster. The segmentation result is shown in Fig. 4.

Due to the electromechanical properties of railway switches [20], the action cluster always corresponds to the action stage; therefore, the three parts of X correspond to the three stages of switch operation.

3.1.2. Fault-free dataset extraction

In this section, the K-means method is used to obtain a fault-free dataset. In the "different durations" problem, several features
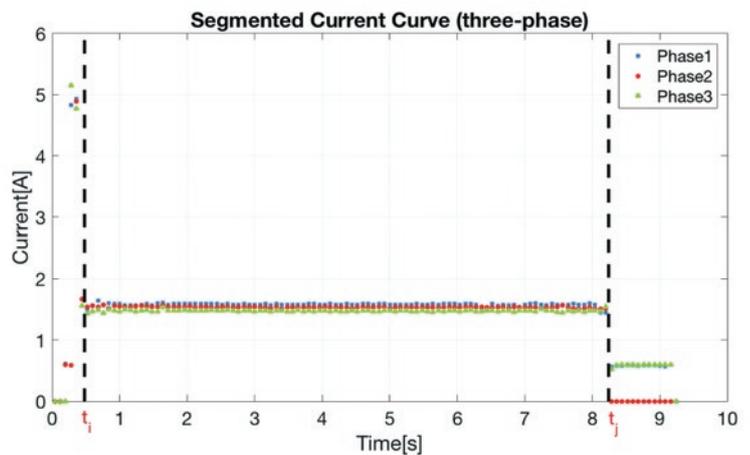

Fig. 4. Segmented Current Curve

are extracted based on previous research [27, 6, 15] to unite dimensions, as shown in Table 2. In the start stage, no fault type exists, and relatively few features have been chosen. In the action and release stages, duration and shape faults both exist; thus, the time span has been selected for duration faults, and other metrics are associated with shape faults.

Table 2. Features of Different Stages

Stages	No.	Features
Start stage	1	Time span
	2	Maximum value
	3	Mean current value
	4	Median current value
Action stage	5	Time span
	6	Max current value
	7	Minimum current value
	8	Mean
	9	Median
	10	Standard deviation
	11	Peak factor
	12	Fluctuation factor
Slow release stage	13	Time span
	14	Max current value
	15	Minimum current value
	16	Mean
	17	Median
	18	Standard deviation
	19	Peak factor
	20	Fluctuation factor

The K-means method has been widely used in clustering for simplicity, and the algorithmic details have been summarized in previous research [9]. The inputs of this method consist of two parameters: the feature matrix $\bar{M}$ and number of clusters K.

$$Idx^K = K\_means\left(\bar{M}, K\right) \tag{7}$$

where Idx^K is an array and the superscript of Idx represents the number of clusters in the array. The feature matrix $\bar{M}$ is defined by the twenty features shown in Table 2 (e.g., m sequences $\left(n_1, n_2, \ldots, n_m\right)$ can generate a feature matrix with m rows and twenty columns). The number of clusters K is determined by assuming that more than half of the samples are fault-free for regular switches. The optimal K^* can be determined from the following optimization problem:

$$\max K$$

$$\begin{cases} \sum_{i=1}^{m} 1\left(Idx_i^{K+1} = mode\left(Idx^{K+1}\right)\right) < m/2 \\ \sum_{i=1}^{m} 1\left(Idx_i^K = mode\left(Idx^K\right)\right) \geq m/2 \end{cases} \tag{8}$$

where:

1 = indicator function

$mode(x) =$ value that appears most often in array x

m = number of samples

The above integer programming problem can be solved by the enumeration method. As a result, the fault-free dataset N^* with K^* satisfies:

$$N^* = \left\{ n_j^* \in \{n_1, n_2, \ldots, n_m\} \mid n_j^* * 1\left(Idx_j^{K^*} = mode\left(Idx^{K^*}\right)\right) = n_j^* \right\}_{j=1}^{m}. \tag{9}$$

3.2. Standard curve selection

3.2.1. Derivative dynamic time warping

Derivative dynamic time warping (DDTW) is a modified DTW method [11]. The approach involves obtaining similarities between two arbitrary trajectories, and it achieves better alignment by "warping" the time axis of one sequence or both sequences. The algorithm details can be summarized as follows.

Assume two arbitrary switch current sequences X^1 and X^2 of lengths n_1 and n_2, respectively, where:

$$X^1 = \left\{ x_1^1, x_2^1, \ldots x_i^1 \ldots, x_{n_1}^1 \right\} \tag{10}$$

$$X^2 = \left\{ x_1^2, x_2^2, \ldots x_j^2 \ldots, x_{n_2}^2 \right\}. \tag{11}$$

To align the two sequences, an n-by-m matrix is considered, where the $\left(i^{th}, j^{th}\right)$ element represents the distance $d\left(x_i^1, x_j^2\right)$ between points x_i^1 and x_j^2. With DDTW, the distance measure $d\left(x_i^1, x_j^2\right)$ is the square of the difference of the estimated derivatives [18] of x_i^1 and x_j^2. Each matrix element (i, j) corresponds to the alignment between points x_i^1 and x_j^2. Therefore, a warping path W is used to define the mapping between X^1 and X^2. The lth element of W is defined as $w_1 = (i, j)_1$; thus, we obtain the following relation:

$$W = \{w_1, w_2, \ldots, w_l, \ldots, w_L\}, \quad max(n_1, n_2) \leq L < n_1 + n_2 - 1. \tag{12}$$

Previous research [18] on DTW has demonstrated that W can be efficiently found by dynamic programming. To formulate a dynamic programming problem, a distance measure must be used between two elements. In this paper, the 1-norm $\left(\|\cdot\|\right)$ is chosen as the distance function δ:

$$\delta(i, j) = x_i^1, x_{j1}^2 \tag{13}$$

After a distance measure is defined, the DTW problem can be formally defined as a minimization over potential warping paths based on the cumulative distance of each path, where δ is a distance measure between two elements. As a result, the similarity between two sequences is defined by Equation (14).

$$DDTW\left(X^1, X^2\right) = min_W\left[\sum_{l=1}^{L}\delta\left(w_l\right)\right] \qquad (14)$$

3.2.2. Standard curve selection

In this section, the "best" option from N^* is selected as the standard curve. For the fault-free dataset N^* with l cardinality, an l-rank square matrix D can be constructed for which the $\left(i^{th}, j^{th}\right)$ element represents the similarity between the ith sequence and j^{th} sequence in N^* (the "similarity" is typically defined as $D_{ij} = DDTW\left(n_i^*, n_j^*\right)$).

In this paper, the sth sequence is defined as the standard curve if its index satisfies:

$$s = arg_i min\left(max_j\left(D_{ij}\right)\right). \qquad (15)$$

3.3. Fault detection and diagnosis

3.3.1. Duration fault detection and diagnosis

In this section, an arbitrary sequence can be detected using a quartile scheme to determine whether a duration fault has occurred. The three steps of the quartile scheme are as follows:
- Step 1: Assume that a dataset with m samples (current curves) has been segmented into three stages (set the start stage F^{sta} as an example):

$$F^{sta} = \left\{F_1^{sta}, F_2^{sta} \ldots, F_m^{sta}\right\} \qquad (16)$$

where F_i^{sta} is the start stage of the ith sample. In addition, an array C^{sta} is set for when the jth element c_j^{sta} equals the cardinality of F_j^{sta}.
- Step 2: Calculate the interquartile range of C^{sta} as:

$$IQR^{sta} = Q_3^{sta} - Q_1^{sta} \qquad (17)$$

where Q_1^{sta} and Q_3^{sta} are the first and third quartiles of the start stage, respectively.
- Step 3: Define a decision function $P(i)$.

$$P(i) = 1\left(c_i^{sta}\left\langle IQR^{sta} - 1.5Q_1^{sta} \vee c_i^{sta}\right\rangle IQR^{sta} + 1.5Q_3^{sta}\right) \; 1 \le i \le m \quad (18)$$

where:

$\vee = $ logical OR

As a result, the ith sample can be identified as a duration fault when $P(i)$ equals one.

3.3.2. Shape fault detection and diagnosis

In this section, three steps are used to diagnose an arbitrary sequence F.

- Step 1: Divide F into three segments with the curve segmentation method:

$$F = \left\{F^{sta}, F^{act}, F^{rel}\right\} \qquad (19)$$

- Step 2: Define a diagnosis dataset M that includes the standard current curves of three stages and their corresponding shape faults, as follows:

$$M = \left\{[S_1], \begin{bmatrix} S_2 \\ M_{1(2)} \\ M_{2(2)} \\ M_{3(2)} \\ M_{4(2)} \\ M_{5(2)} \\ M_{6(2)} \\ M_{7(2)} \end{bmatrix}, \begin{bmatrix} S_3 \\ M_{8(3)} \\ M_{9(3)} \end{bmatrix}\right\} \qquad (20)$$

where $M_{i(j)}$ denotes j^{th} stage data of the ith fault (S_j is the standard curve of the jth stage). In the first stage (start stage), there is no fault type, which means that only S_1 exists in the first column of M. $M_1 - M_7$ occur in the second stage (action stage); thus, $M_{1(2)} - M_{7(2)}$ and S_2 are grouped together in the second column of M. Furthermore, M_8 and M_9 occur in the third stage (release stage); therefore, the third column of M consists of $M_{8(3)}$, $M_{9(3)}$ and S_3.
- Step 3: DDTW is employed to calculate the similarities between one stage in F and the corresponding stage in M. Each stage of F can be evaluated with Equation (21) and diagnosed with Table 3.

Table 3. Diagnostic Results for Shape Faults

Equation (20) Outputs / Diagnostic Results	$Label^{sta}$	$Label^{act}$	$Label^{rel}$
Fault-free	1	1	1
M1	1	2	1
M2	1	3	1
M3	1	4	1
M4	1	5	1
M5	1	6	1
M6	1	7	1
M7	1	8	1
M8	1	1	2
M9	1	1	3

$$\begin{cases} Label^{sta} = arg_i\ min\Big(DDTW\big(F_{sta}, M_{i,1}\big)\Big) \\ Label^{act} = arg_i\ min\Big(DDTW\big(F_{act}, M_{i,2}\big)\Big) \\ Label^{rel} = arg_i\ min\Big(DDTW\big(F_{rel}, M_{i,3}\big)\Big) \end{cases} \qquad (21)$$

where $Label^{sta}$, $Label^{act}$ and $Label^{rel}$ respectively denote the classification results of the three stages.

4. Experiment and results

In this study, 1,964 fault-free curves and 115 fault curves were collected from the Guangzhou-Shaoguan Railway in China. The dataset was randomly split into two subsets (training and testing sets) that account for 70% and 30% of the entire dataset. For HFD, all training data are used to generate the standard curve. Then, 70% fault curves of the training set and the standard curve are combined to form the diagnosis dataset. The diagnostic results of 10 current curves are shown in Table 4.

In Table 4, M_{10} and M_{11} are determined by the quartile scheme, and the other faults are determined by DDTW. Test samples can be classified only as M_{10} and M_{11} when the corresponding decision function equals one. Without considering duration faults (M_{10} and M_{11}), the minimum of each row is found, which indicates that the ith test sample is highly similar to the reference template; therefore, the samples can be classified in the same class.

Additionally, the DTW method [8] with the quartile scheme and the SVM method based on twenty features (Table 2) are compared with HFD. For the SVM, a Gaussian kernel is used as the kernel function, and the penalty factor and kernel parameter are determined by a 10-fold cross-validation method [19]. A quantitative comparison of the three methods is provided in Table 5. Two indicators, the false alarm rate (FAR) and missing alarm rate (MAR), are introduced in the table. FAR denotes the probability of classifying the fault-free data as faulty, and MAR denotes the probability of classifying fault data as fault free.

The following conclusions can be drawn from Table 5 regarding the experimental results.

- The HFD method is the best of the three methods due to its high accuracy, low FAR and low MAR.
- Compared to HFD, the DTW method exhibits classification results and cannot be used for fault diagnosis because of its high MAR. HFD performs better than DTW for two reasons. First, drawbacks such as "singularities" [26] prevent DTW from producing the best warping results. Second, DTW is focused on current values, but HFD focuses on both current values and data fluctuations.
- Compared to HFD, the SVM method offers a generally acceptable level of classification quality, but it still makes incorrect classifications and generates a relatively high MAR, which prevents the application of the SVM method in practical applications. As shown in Table 5 the HFD method performs better than the SVM method because HFD makes full use of all available information, whereas SVM disregards certain information when applying the feature extraction method.

5. Conclusions

In this paper, an intelligent fault diagnosis method is proposed based on the segmentation of railway switches. Through previous analysis, this paper illustrates how to divide current curves based on three mechanical processes for all railway switches and how to determine the similarities between them.

The experimental results show that the HFD method can detect faults with 99.43% accuracy and can diagnose faults with 98.67% accuracy. This approach is superior to the other two methods introduced above. Furthermore, the lower FAR and MAR of the HFD method demonstrate that HFD is the most robust tool for fault detection and diagnosis.

Future work will strive to integrate power curves with the proposed HFD method to achieve more accurate results. Furthermore, undefined switch faults will be examined for broader applicability and operability of the method. The final future objective is to more intelligently detect railway switch faults and eventually improve the safety and efficiency levels for passenger and cargo transport.

Table 4. Distance between the Test Samples and Reference Faults

i	S	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}	Predicted Label	Actual Label
1	0.31	0.42	49.92	2.26	223.38	8.74	172.04	181.12	1.79	17.51	0	0	S	S
2	3.21	1.04	2.57	11.30	49.43	37.62	33.28	20.94	31.24	16.25	0	0	M_1	M_1
3	4.48	1.49	2.34	3.91	29.21	0.59	43.15	31.23	1.67	17.21	0	0	M_5	M_5
4	0.41	0.60	1.04	2.26	31.24	8.76	49.43	37.61	1.84	17.22	0	0	S	S
5	2.31	1.41	3.10	2.28	32.10	6.91	21.13	0.83	1.91	16.83	0	0	M_7	M_7
6	0.36	0.49	0.53	2.27	29.89	8.73	51.62	39.69	1.75	17.24	0	0	S	S
7	0.36	0.66	0.48	2.25	28.30	8.79	51.21	40.04	1.72	17.23	0	0	S	S
8	5.40	0.93	1.02	2.21	19.93	6.97	50.83	30.21	1.73	17.28	1	0	M_{10}	M_{10}
9	4.31	2.10	0.43	2.31	20.08	9.31	53.21	32.25	1.79	17.30	0	0	M_2	M_2
10	0.40	0.64	0.37	2.28	24.78	8.76	52.67	40.76	1.77	17.23	0	0	S	S

Table 5.   Fault Detection and Diagnosis Results

DTW	M1%	M2%	M3%	M4%	M5%	M6%	M7%	M8%	M9%	M10%	M11%	S%
M1%	97	0	0	0	3	0	0	0	0	0	0	0
M2%	0	100	0	0	0	0	0	0	0	0	0	0
M3%	0	0	64	0	0	0	0	0	0	0	0	36
M4%	0	3	0	94	0	0	0	0	0	0	0	42
M5%	0	0	0	0	58	0	0	0	0	0	0	0
M6%	0	0	0	0	0	100	0	0	0	0	0	0
M7%	0	0	0	0	0	0	100	0	0	0	0	0
M8%	0	0	0	0	0	0	0	97	3	0	0	0
M9%	0	0	0	0	0	0	0	0	100	0	0	0
M10%	0	0	0	0	0	0	0	0	0	100	0	0
M11%	0	0	0	0	0	0	0	0	0	0	100	0
S%	0	0	0	0	0	0	0	0	0	0	0	100
Detection Accuracy(%)						84.08						
FAR						0.02						
MAR						0.44						
Diagnosis Accuracy(%)						68.32						

(a) DTW

SVM	M1%	M2%	M3%	M4%	M5%	M6%	M7%	M8%	M9%	M10%	M11%	S%
M1%	100	0	0	0	0	0	0	0	0	0	0	0
M2%	0	0	0	0	0	0	0	0	0	0	0	100
M3%	0	0	91	0	0	0	0	0	0	0	0	9
M4%	0	0	3	97	0	0	0	0	0	0	0	0
M5%	0	0	0	0	0	0	0	0	0	0	0	100
M6%	0	0	0	0	0	100	0	0	0	0	0	0
M7%	0	0	0	0	0	0	100	0	0	0	0	0
M8%	0	0	0	0	0	0	0	100	0	0	0	0
M9%	0	0	0	0	0	0	0	0	100	0	0	0
M10%	0	0	0	0	0	0	0	0	0	100	0	0
M11%	0	0	0	0	0	0	0	0	0	0	100	0
S%	0	0	0	0	0	0	0	0	0	0	0	100
Detection Accuracy(%)						90.29						
FAR						0.013						
MAR						0.182						
Diagnosis Accuracy(%)						82.11						

(b) SVM

HFD	M1%	M2%	M3%	M4%	M5%	M6%	M7%	M8%	M9%	M10%	M11%	S%
M1%	100	0	0	0	0	0	0	0	0	0	0	0
M2%	0	100	0	0	0	0	0	0	0	0	0	0
M3%	0	0	100	0	0	0	0	0	0	0	0	0
M4%	0	3	0	94	0	0	0	0	0	0	0	3
M5%	0	0	0	0	100	0	0	0	0	0	0	0
M6%	0	0	0	0	0	100	0	0	0	0	0	0
M7%	0	0	0	0	0	0	100	0	0	0	0	0
M8%	0	0	0	0	0	0	0	100	0	0	0	0
M9%	0	0	0	0	0	0	0	0	100	0	0	0
M10%	0	0	0	0	0	0	0	0	0	100	0	0
M11%	0	0	0	0	0	0	0	0	0	0	100	0
S%	0	0	0	0	0	0	0	0	0	0	0	100
Detection Accuracy(%)						99.43						
FAR						0.013						
MAR						0.013						
Diagnosis Accuracy(%)						98.67						

(c) HFD

References

1. Ardakani H D, Lucas C, Siegel D, Chang S, Dersin P, Bonnet B, Lee J. PHM for railway system-a case study on the health assessment of the point machines. IEEE Conference on Prognostics and Health Management 2012: 1-5, https://doi.org/10.1109/ICPHM.2012.6299533.
2. Asada T, Roberts C, Koseki T. An algorithm for improved performance of railway condition monitoring equipment: alternating-current point machine case study. Transportation Research Part C: Emerging Technologies 2013; 30: 81-92, https://doi.org/10.1016/j.trc.2013.01.008.
3. Atamuradov V, Camci F, Baskan S, Sevkli M. Failure diagnostics for railway point machines using expert systems. IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives 2009: 1-5, https://doi.org/10.1109/DEMPED.2009.5292755.
4. Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002; 24: 603-619, https://doi.org/10.1109/34.1000236.
5. Comaniciu D, Ramesh V, Meer P. The variable bandwidth mean shift and data-driven scale selection. Proceedings 8th IEEE International Conference on Computer Vision ICCV 2001: 438-445, https://doi.org/10.1109/ICCV.2001.937550.
6. Eker O, Camci F, Kumar U. SVM based diagnostics on railway turnouts. International Journal of Performability Engineering 2012; 8: 289-298.
7. Eker O F, Camci F, Guclu A, Yilboga H, Sevkli M, Baskan S. A simple state-based prognostic model for railway turnout systems. IEEE Transactions on Industrial Electronics 2011; 58: 1718-1726, https://doi.org/10.1109/TIE.2010.2051399.
8. Fasshauer G E. Positive definite kernels: past, present and future. Dolomite Research Notes on Approximation 2011; 4: 21-63.
9. Hartigan J A, Wong M A. Algorithm AS 136: A K-means clustering algorithm. Journal of the Royal Statistical Society Series C 1979; 28: 100-108, https://doi.org/10.2307/2346830.
10. He Y. Research on fault diagnosis method of high-speed railway turnouts. Master's Thesis. Beijing: Beijing Jiaotong University, 2014.
11. Keogh E J, Pazzani M J. Derivative dynamic time warping. Proceedings of the 2001 SIAM International Conference on Data Mining 2001: 1-11, https://doi.org/10.1137/1.9781611972719.1.
12. Kim H, Sa J, Chung Y, Park D, Yoon S. Fault diagnosis of railway point machines using dynamic time warping. Electronics Letters 2016; 52: 818-819, https://doi.org/10.1049/el.2016.0206.
13. Lee J, Choi H, Park D, Chung Y, Kim H-Y, Yoon S. Fault detection and diagnosis of railway point machines by sound Analysis. Sensors 2016; 16: 549, https://doi.org/10.3390/s16040549.
14. Letot C, Dersin P, Pugnaloni M, Dehombreux P, Fleurquin G, Douziech C, La-Cascia P. A data driven degradation-based model for the maintenance of turnouts: a case study. IFAC-PapersOnLine 2015; 48: 958-963, https://doi.org/10.1016/j.ifacol.2015.09.650.
15. Lu Q. Method of turnout fault diagnosis based on grey correlation analysis. MSc Thesis. Beijing, China: Beijing Jiaotong University, 2015.
16. Márquez F P G, Roberts C, Tobias A M. Railway point mechanisms: condition monitoring and fault detection. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 2010; 224: 35-44, https://doi.org/10.1243/09544097JRRT289.
17. Marquez F P G, Schmid F, Collado J C. A reliability centered approach to remote condition monitoring. A railway points case study. Reliability Engineering & System Safety 2003; 80: 33-40, https://doi.org/10.1016/S0951-8320(02)00166-7.
18. Mills T C. Time series techniques for economists. Cambridge, MA: Cambridge University Press, 2012.
19. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V. Scikit-learn: machine learning in Python. Journal of Machine Learning Research 2011; 12: 2825-2830.
20. Ren Z, Sun S, Zhai W. Study on lateral dynamic characteristics of vehicle/turnout system. Vehicle System Dynamics 2005; 43: 285-303, https://doi.org/10.1080/00423110500083262.
21. Shaw D C. A universal approach to points condition monitoring. 4th IET International Conference on Railway Condition Monitoring 2008: 1-6, https://doi.org/10.1049/ic:20080315.
22. Wang G, Xu T, Tang T, Yuan T, Wang H. A Bayesian network model for prediction of weather-related failures in railway turnout systems. Expert Systems with Applications 2017; 69: 247-256, https://doi.org/10.1016/j.eswa.2016.10.011.
23. Zhang K. The railway turnout fault diagnosis algorithm based on BP neural network. IEEE International Conference on Control Science and Systems Engineering 2014: 135-138, https://doi.org/10.1109/CCSSE.2014.7224524.
24. Zhang K, Du K, Ju Y. Algorithm of railway turnout fault detection based on PNN neural network. 7th International Symposium on Computational Intelligence and Design 2014: 544-547, https://doi.org/10.1109/ISCID.2014.140.
25. Zhao H, Chen H, Dong W, Sun X, Ji Y. Fault diagnosis of rail turnout system based on case-based reasoning with compound distance methods. Chinese Control and Decision Conference (CCDC) 2017: 4205-4210, https://doi.org/10.1109/CCDC.2017.7979237.
26. Zhao L, Lu Q. Method of turnout fault diagnosis based on grey correlation analysis. Journal of the China Railway Society 2014; 36: 69-74.
27. Zhou F, Xia L, Dong W, Sun X, Yan X, Zhao Q. Fault diagnosis of high-speed railway turnout based on support vector machine. IEEE International Conference on Industrial Technology (ICIT) 2016: 1539-1544, https://doi.org/10.1109/ICIT.2016.7474989.
28. Zhou J-L, Lei Y. Paths between latent and active errors: analysis of 407 railway accidents/incidents' causes in China. Safety Science 2017, https://doi.org/10.1016/j.ssci.2017.12.027.

Dongxiu OU
Maojie TANG
Rui XUE
The Key Laboratory of Road and Traffic Engineering, Ministry of Education
China School of Transportation Engineering, Tongji University
Shanghai, 201804, P.R. China

Hongjing YAO
Jinan Railway Bureau, Jinan, Shangdong, China, 250000, P.R. China

Emails: ou.dongxiu@tongji.edu.cn, 1632486@tongji.edu.cn, xuerui@tongji.edu.cn, 13954123465@139.com