Simulation Research on High-Speed Railway Dropper Fault Detection and Location Based on Time-Frequency Analysis

Jin Li¹, Xuewu Zhang¹, Cheng Zhang¹ and Tiantian Tian²

¹China Railway Frist Survey and Design Institute Group CO., Ltd, Xi’an, China
²School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an, China
Email: 757340551@qq.com

Abstract. In this paper, a machine learning detection method, namely, SVM-ICA, which aims to solve the fault identification of droppers in high-speed railway, was proposed based on time-frequency analysis. The proposed SVM-ICA method can be utilized to detect and locate the faulty droppers. In detail, the time-frequency statistical features of the sensor data are firstly extracted and the significant features are selected to construct the training set. Secondly, based on the training set, the support vector machine (SVM) fault detection model is trained. Finally, the trained model is used for detecting faulty droppers, and further the position of the break string can be located by using the independent component analysis (ICA) method. Simulation results show that the proposed fault detection method of droppers based on time-frequency analysis can accurately identify if the string breaks, and additionally can locate the position where the failure occurred.

1. Introduction
As an important part of the OCS (overhead contact system), droppers are installed between messengers and contact lines, whose main function is to control the height of the contact line and guarantee the good current-receiving quality of the pantograph-catenary system [1-2]. The catenary of high-speed railway is exposed in the natural environment for a long time. Due to the severe weather and the impact of the pantograph during operation, droppers may break, threatening the safety of driving [3-4]. According to statistics, the OCS has become a weak link of the railway system, and has been highly valued by the Safety Monitoring Departments [5]. At present, the fault detection of droppers is generally carried out by manual inspection, but some hidden problems, such as the breakage of the thread at the adjusting bolt, can be only found when the OCS is inspected at high altitude. On the other hand, due to the heavy workload of human inspection and the subjective influence, it is difficult to ensure the quality of inspection [6-8]. Therefore, the automatic detection technology of broken droppers has become a research hotspot that railway departments pay more attention to.
In this paper, we obtain the mechanical performance of the OCS when the dropper is broken by simulation experiments, calculate the accelerations of control lines and messengers at specific points, and extract the characteristics of the accelerations based the time-frequency analysis. Finally, a SVM-ICA method for the dropper fault detection is proposed.

2. Methodology
2.1. Support Vector Machines (SVMs)
Support vector machines [9] (SVMs) is a classification method based on the VC (Vapnik-Chervonenkis)
dimension and structural risk minimization, which obtains the classification hyperplane by finding the maximum class spacing and minimizing the structural risk. Assume the training data of D with a set of n points of calculating by equation (1):

\[ D = \{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{+1, -1\}\}_{i=1}^N \] (1)

where \( x_i \) is the \( i \)-th sample, N is the sample size, and p is the dimension of the sample. We want to find the maximum-margin hyperplane that divides the points having \( y_i = 1 \) from those having \( y_i = -1 \). To find the optimal classification hyperplane \( \omega \cdot x + b = 0 \) of SVM is equivalent to minimizing the objective function

\[
\min J(\omega) = \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^{m} \xi_i \\
\text{s.t. } y_i (\omega^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \ldots, m
\] (2)

Here \( \omega \) denotes the normal vector of hyper plane; \( b \) is the offset; \( \frac{2}{||\omega||} \) is the margin; the \( \xi_i \) measures the degree of misclassification of the data point \( i \), the nonzero \( C \) is a cost parameter, which is a trade-off between a large margin and a small error. Lagrange multipliers method is utilized to obtain the dual problem:

\[
\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j x_i^T x_j \\
\text{s.t. } \sum_{i=1}^{m} \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, i = 1, 2, \ldots, m
\] (3)

where \( \alpha_i \geq 0 \) is the Lagrange multiplier. Then, we can get the optimal solution:

\[
\omega = \sum_{i=1}^{m} \alpha_i y_i x_i, b = y - \omega x
\] (4)

In equation (4), \((x, y)\) are any points that satisfy \( 0 < \alpha_i < C \), so the decision function of the linear support vector machine is

\[ f(x) = \text{sgn}(\omega \cdot x + b) \] (5)

Here, \( \text{sgn}(x) \) is a sign function. In practice, the above-mentioned dual problem (3) is generally solved by the SMO algorithm [10].

2.2. Independent Component Analysis (ICA)

The main components of the catenary, such as droppers, messenger wires, and contact lines, form a flexible system. The vibration of these components overlaps each other, therefore, the accelerations and vibration signals of the messenger wires and contact lines are actually composed of multiple source signals. In this paper, ICA method is applied to separate the source signals so as to locate the broken position of droppers.

Independent component analysis (ICA) is a machine learning method, whose goal is to estimate underlying component signals and unknown mixed channel parameters only through mixed observed signals [11]. The linear transient mixture model of ICA is expressed as:

\[ x(t) = As(t) \] (6)

where \( x(t) \) denotes a m-dimensional observed signals, \( A \) is a \( m \times n \) mixing matrix, \( s(t) \) means unknown source signals.

The mixture decomposition of ICA is to find the unmixed matrix \( W \), through which unknown source signals \( \tilde{s}(t) \) can be obtained only from the observed signals \( x(t) \), the unmixed matrix can be estimated by MLE (Maximum Likelihood Estimation) or information maximization algorithm. The mixture decomposition process of ICA is shown in figure 1.
Figure 1. The mixture decomposition process of ICA.

In the monitoring process based on ICA [12-13], the independent statistics are introduced, in which the statistics $SPE$ represents the change that can't be explained by the independent component model at the $i$-th. The statistics $I^2$ is the main statistics, which reflects the change of multi-variables through the fluctuations of the independent component vector inside the independent model. And $I^2_e$ is the auxiliary statistics, which is an additional monitoring tool, it can compensate the error caused by the improper choice of main independent components.

2.3. Fault Detection and Location: SVM-ICA Algorithm

Using SVM and ICA to detect and locate the faulty droppers. In detail, the time-frequency statistical features of the accelerations at specific points are firstly extracted and the significant features are selected to train the SVM fault detection model. Then, the signals in normal state are unmixed by ICA, the thresholds of detection statistics are calculated. By the range of the over-thresholds, the location of dropper breakage finally be found. The algorithm flow is shown in figure 2, the steps are as follows figure 2.

Training and detection process of SVM-ICA model, which is utilized to detect and locate the broken droppers as follows:

- Step 1. Extract time-frequency statistical features and select significant features of accelerations;
- Step 2. Train SVM model by using training set and construct the dropper fault detection model;
- Step 3. Decompose the positive sample in test set by ICA, and calculate the thresholds of detection statistics;
- Step 4. Detect the test set by using the model in Step 2. If the result is broken, turn to Step 5, otherwise the output is no fault;
- Step 5. Decompose the signals to be detected, calculate the time periods when the statistics exceed the thresholds, and locate the fault position of dropper according to the train speed.

Figure 2. SVM-ICA algorithm flow.
3. Experiment of Catenary Dropper Fault Detection and Location Algorithm

3.1. Dropper Fracture Simulation

Although there are many breakdowns of droppers on the main railway line every year, the corresponding state data cannot be obtained due to the absence of sensors on the catenary. If a large number of sensors are installed in catenary contact wires and messenger lines, it will inevitably consume too much manpower and financial resources, and the collected data of the dropper fracture will not reach the sample size required for data modeling. In this paper, based on the simulation model in Ref. [14, 15], droppers of catenary are simulated for many times when pantograph and pulsating wind work together, and the data of normal and abnormal state of catenary are both obtained.

3.1.1. Catenary Simulation Parameters. The Xinjiang section of Lanxin high-speed railway is often attacked by strong wind, so there is a high possibility of dropper fracture failure. In this paper, the driving state of two-semi overlap span and catenary in single catenary system with the mileage of K3066 + 568.795 ~ K3065 + 588.795 are simulated. The design parameters are shown in table 1.

| Parameter                  | Value       |
|---------------------------|-------------|
| Span (m)                  | 50          |
| Messenger wire            | JTMH120     |
| Contact line              | CTMH150     |
| Dropper interval (m)      | 5           |
| Number of droppers        | 9           |
| Quality per unit messenger wire (kg/m) | 1.065 |
| Quality per unit contact line (kg/m) | 1.350 |
| Analysis interval of droppers fault (m) | 750-800 |

3.1.2. Pantograph Simulation Model. In conjunction with this OCS is the pantograph DSA250. The simulation model of pantograph is a multi-body system with mass, stiffness and damping, as shown in figure 3, and the parameters are shown in table 2. The static contact stress $f_0$ is 70N, and the aerodynamic stress is $0.00097v^2$ according to the IEC 62486, where $v$ is the speed of pantograph, taking 250 km/h and single pantograph.

![Figure 3. Three equivalent mass of pantograph.](image-url)
Table 2. Parameters of DSA250 pantograph.

| Symbol | Value | Symbol | Value |
|--------|-------|--------|-------|
| $m_3$ (kg) | 7.51 | $c_3$ (Ns/m) | 0 |
| $m_2$ (kg) | 5.855 | $c_2$ (Ns/m) | 0 |
| $m_1$ (kg) | 4.645 | $c_1$ (Ns/m) | 70 |
| $k_3$ (N/m) | 8380 | $f_3$ (N) | 0.5 |
| $k_2$ (N/m) | 6200 | $f_2$ (N) | 3.5 |
| $k_1$ (N/m) | 80 | $f_1$ (N) | 3.5 |

3.1.3. Installation of Acceleration Sensors. Acceleration sensors are respectively set at the support point 1, the center and the support point 2 of the messenger, and at the mid-point and fix point 2 of the contact line, as shown in figure 4. The layout of sensors in each span of 18-25 is the same.

Figure 4. Sensor installation.

3.1.4. Simulation Experiment. The simulation experiment is that the train at a constant speed of 250km/h, and the pantograph and pulsating wind act on the two test sections of the catenary: case 1 is the normal OCS in the test section; case 2 is the OCS, where the dropper is broken in the center of 21-span. Based on the simulation method of [14], simulations were performed 200 times for each case, with a sampling frequency of 277 Hz. The horizontal and vertical accelerations of support points 1, points 2, mid-span of the messengers, and the contact lines fix point 2 and the center of span of 18 to 25 are obtained.

3.2. Detection and Location Experiment of Dropper Fault

3.2.1. Feature Extraction. For simulated sensor data, feature extraction is carried out in time and frequency respectively [16-18], and the significant statistical features include: the waveform factor, kurtosis factor, root mean square frequency and frequency standard deviation of the vertical acceleration in the center span of messenger; the maximum, minimum, peak, standard deviation, kurtosis, root mean square, waveform factor, peak factor, kurtosis factor, pulse factor, margin, center of gravity frequency, mean square frequency and frequency variance of horizontal acceleration in the center span of messenger; the maximum, minimum, mean, average of absolute, peak, standard deviation, kurtosis, root mean square, waveform factor, peak factor, kurtosis factor, pulse factor and margin of the vertical acceleration in the contact line span in the center span of contact line; the standard deviation, root mean square, waveform factor and kurtosis factor of horizontal acceleration in mid-span of contact line. A total of 38 characteristics. A total of 38 features.

3.2.2. Detection Model Construction. In this experiment, the data are divided into training set and test set, in which the training set accounts for 80%, that is, 320 samples, including 160 positive samples, 160 negative samples, the test set counts for 20%, that is, 80 samples, including 40 positive samples, 40 negative samples. Through 5-fold cross-validation, the detection model under the simultaneous action of wind and pantograph is trained. The test shows its accuracy reaches 95.17%.

3.2.3. Fault Location. The 10 acceleration signals of the positive sample (sample 1 as an example) in the test set are analyzed by ICA. The unmixed signals are obtained as figure 5, and the distribution of
detection statistics is gained as figure 6. Then, we can get thresholds of statistics by kernel density estimation.

For the detection statistics $I^2$, $I_e^2$ and $SPE$ in figure 6, nonparametric kernel density estimation method is utilized to estimate their probability function and thresholds. Then fault detection is carried out to determine the time periods when the fault exceeds the thresholds. From the test set, the 40 positive samples, can be applied to estimate the statistic thresholds, and then the fault location of the negative samples can be determined as shown in table 3.

![Figure 5. Unmixed signals.](image1)

![Figure 6. Distribution of detection statistics.](image2)

### Table 3. Fault location.

| Statistics | State               | Wind 1  | Wind 2  | Wind 3-39 | Wind 40  |
|------------|---------------------|---------|---------|-----------|----------|
| $I^2$      | Normal (threshold)  | 62.11   | 70.04   | ......     | 64.71    |
|            | Failure period (s)  | 10.66-11.75 | 11.00-11.714 | ...... | 10.69-11.46 |
| $I_e^2$    | Normal (threshold)  | 21.216  | 18.49   | ......     | 15.20    |
|            | Failure period (s)  | 11.21-11.71 | 10.93-11.41 | ...... | 10.67-11.45 |
| $SPE$      | Normal (threshold)  | 265.48  | 275.61  | ......     | 187.48   |
|            | Failure period (s)  | 10.67-11.39 | 10.99-11.70 | ...... | 10.71-11.68 |

In the case of Wind 1, the thresholds determined by the kernel density estimation of the three statistics are 62.11, 21.216 and 265.48 respectively under normal conditions, and the time periods of failure data exceeding the thresholds are 10.66-11.75 s and 11.21-11.71 s, and 10.668-11.394 s. Moreover, because the value of the intermediate point is the largest, it can be considered that the intermediate point is the most likely to have fault. The average time of the three statistics is 11.2303s, and the train speed is 250 km/h. So at this time, the train is located at 29.884m from the beginning of the 21st span, while the length of the 21 span is 50m, so it can be determined that the fault is approximately near the middle of the 21 span, and other wind are similar.

### 4. Conclusion

In this paper, a machine learning detection method, namely, SVM-ICA, which aims to solve the fault identification of droppers in high-speed railway, was proposed based on time-frequency analysis. Experiment shows that the fault detection model based on statistical features of accelerations in time-frequency can identify broken droppers with 95.17% accuracy and locate the fault with less error. The next step is to install acceleration sensors in the wind test section of the Lanxin line to verify the validity of the model.
Acknowledgements
This paper is supported by Major Special Project (18-a02) of China Railway Construction Corporation Limited in 2018 and Science and Technology Program (201809164CX5JC6, 2019421315 KYPT 004JC006) of Xi’an.

References
[1] Zhang B 2016 Cause Analysis and countermeasures of breaking of integral suspension chord of catenary in high-speed railway in China Zhengzhou Railway Science & Technology (4) 2-10.
[2] Ma J 2019 Discussion on the running state of rigid monolithic hanging strings Shandong Industrial Technology (20) 97-97.
[3] Wang W 2014 Fatigue Characteristics Analysis of Catenary Droppers in High-Speed Railway (Southwest Jiaotong University).
[4] He C 2015 Research on The Dynamic of Catenary Dropper (Southwest Jiaotong University).
[5] Shi Y 2016 Research on Measurement System of Geometric Parameters for Railway Overhead Line System (Tianjin University).
[6] Yang G 2016 Fracture analysis of integral suspension chord of overhead contact system in high-speed railway and verification of improvement effect Journal of Railway Technology Supervision 44 (9) 21-23.
[7] Jiang X, Gu X and Deng H 2019 Research on damage mechanism and optimization of integral dropper string based on fretting theory Journal of the China Railway Society (6).
[8] Wu G 2016 Fault Detection Algorithm of Support and Suspension Device of Catenary in High-Speed Railway (Shijiazhuang Tiedao University).
[9] Yan W 2016 Support Vector Machines Optimization Model and Its Application (Hunan Normal University).
[10] Keerthi S S and Shevade S K 2003 SMO algorithm for least-squares SVM formulations Neural Computation 15 (2) 487-507.
[11] Wang G X, Zhang L M and Zhang Y 2010 Blind separation of frequency hopping signals based on independent component analysis Computer and Digital Engineering 10 72-74.
[12] Zhu F 2017 Research on Fault Diagnosis of Industrial Process Based on Data Driven (Harbin University of Science and Technology).
[13] Zhao Y 2012 Monitoring and Abnormalities Diagnosis for Multivariate Process (Tianjin University).
[14] Guan J and Wu J 2017 Building and confirmation for dynamic simulation model of pantograph and catenary Journal of Railway Science and Engineering 14 (11) 2444-2451.
[15] Jing L 2014 DSA250 pantograph principle and common failure analysis Heilongjiang Science and Technology Information (33) 165-165.
[16] Zhang S, Zhou Q and Zhang S 2004 Features extraction of TWACS signal based on time-frequency analysis Journal of Chinese Electrical Engineering Science (07) 89-93.
[17] Ding Y and Gao X 1994 Digital Signal Processing (Xidian University Press).
[18] Mei Z, Shi H and Chang S 2011 Fast time domain diagnosis of signal power spectrum characteristic change Journal of Huazhong University of Technology (3) 87-92.