# Research on the Large Data Intelligent Classification Method for Long-Term Health Monitoring of Bridge

Xiaojiang Hong(✉) and Mingdong Yu

Department of Civil and Hydraulic Engineering Institute, Xichang University, Xichang, China
xcxyymd@163.com

**Abstract.** In order to improve the intelligent management and information scheduling ability of bridge long-term health monitoring, the real-time data monitoring and automatic collection design of bridge long-term health monitoring are carried out with big data analysis method. A classification method of bridge long-term health monitoring data based on fuzzy correlation feature detection and grid area clustering is proposed. The information fusion and fuzzy chromatography analysis method are used to realize the information fusion of the real-time data of bridge long-term health monitoring, and the adaptive feature extraction of related data is carried out. Excavate the positive correlation characteristic quantity of bridge long-term health monitoring real-time monitoring data flow, carry on the fuzzy clustering and information prediction of bridge long-term health monitoring data flow, and improve the accuracy of bridge long-term health monitoring real-time data monitoring. The simulation results show that the intelligent classification of bridge long-term health monitoring based on this method has high accuracy and low error rate, which improves the real-time performance of bridge monitoring.

**Keywords:** Bridge · Long-term health monitoring · Big data classification

## 1 Introduction

With the development of big data's information processing technology, artificial intelligent information management technology is used to carry out bridge long-term health monitoring information management, and pattern recognition technology is used to carry out bridge health monitoring functional information monitoring. Improve the initiative of bridge health monitoring, study the information management technology of bridge long-term health monitoring, construct the bridge long-term health monitoring under the information construction environment, and improve the bridge health quality [1]. The intelligent information construction of bridge long-term health monitoring is based on the automatic monitoring of bridge long-term health monitoring data stream, and the automatic real-time data flow monitoring design is carried out by using artificial intelligence scheduling technology in all kinds of bridge long-term health monitoring. To improve the operation efficiency of bridge long-term health monitoring, the real-time data flow monitoring of bridge long-term health monitoring has attracted great attention [2].

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2019
Published by Springer Nature Switzerland AG 2019. All Rights Reserved
G. Gui and L. Yun (Eds.): ADHIP 2019, LNICST 302, pp. 1–10, 2019.
https://doi.org/10.1007/978-3-030-36405-2_1

At present, the research on intelligent real-time data flow monitoring and scheduling method for bridge long-term health monitoring mainly adopts traffic statistics method, and the main data mining methods are fuzzy mining method. C-means mining method and dynamic ARMA model mining method [3]. The information flow of intelligent real-time data flow monitoring for bridge long-term health monitoring is used for balanced scheduling and global equilibrium design. Improve the real-time data flow monitoring ability of bridge health monitoring center. In reference [4], an intelligent real-time data flow monitoring and equilibrium control method for bridge long-term health monitoring based on artificial intelligence control technology is proposed. The intelligent real-time data flow monitoring information transmission model of bridge long-term health monitoring is constructed by using fuzzy constraint control method, and the intelligent real-time data flow monitoring and classification scheduling of bridge long-term health monitoring is carried out by using Gaussian aggregation classification method. It has good data recall performance, but the online scheduling ability of this method in bridge long-term health monitoring real-time data monitoring is not good [5].

In view of the above problems, a long-term health monitoring data classification method based on fuzzy correlation feature detection and mesh region clustering is presented in this paper. establishing a grid distribution structure model of long-term health monitoring of the bridge and carrying out long-term health monitoring data flow monitoring statistical analysis of the bridge, and realizing the information fusion by adopting the information fusion and the fuzzy chromatography analysis method on the long-term health monitoring real-time data of the bridge, And the positive correlation characteristic quantity of the information flow of the real-time monitoring data of the long-term health monitoring of the bridge is excavated. Based on the hierarchical clustering algorithm, the fuzzy clustering and information prediction of long-term health monitoring data flow of the bridge are carried out by self-regression analysis, and finally, the simulation experiment analysis is carried out, and the superiority of the method in improving the automatic monitoring capability of the long-term health monitoring data flow of the bridge is shown.

## 2 Real-Time Data Flow Monitoring Information Flow Model and Preprocessing for Long-Term Health Monitoring of Bridges

### 2.1 Bridge Long-Term Health Monitoring Real-Time Data Flow Monitoring Information Flow Model

In order to realize real-time data monitoring and balanced management control for long-term health monitoring of bridges, fuzzy correlation feature detection and analytic hierarchy process (AHP) are used for data monitoring. Firstly, the grid distribution structure model of bridge long-term health monitoring is established and the statistical characteristics of bridge long-term health monitoring data flow monitoring are analyzed. In the distributed structure model of bridge long-term health monitoring real-time monitoring data, if directed graph $G_1 = \left(M_1^{\alpha}, M_1^{\beta}, Y_1\right)$, $G_2 = \left(M_2^{\alpha}, M_2^{\beta}, Y_2\right)$ are used as the candidate set of bridge long-term health monitoring real-time monitoring data, then $G_1 \subseteq G_2 \Leftrightarrow Y_1 \subseteq Y_2$, so that $A = \{a_1, a_2, \ldots a_n\}$ can generate candidate set for the model of bridge long-term health monitoring real-time monitoring data acquisition model, and the edge $(u, v) \in E$, of directed graph of bridge long-term health monitoring real-time monitoring data distribution [6]. Under the information management environment, the statistical feature analysis model of bridge long-term health monitoring real-time monitoring is constructed, which is represented by five-tuple $O = (C, I, P, Hc, R, A^0)$, in which $C$ is the utility threshold item set of real-time monitoring data sampling time series of bridge long-term health monitoring. S is an example set of data flow, so the efficient item set for bridge long-term health monitoring is:

$$x(t) = \sum_{i=0}^{p} a(\theta_i) s_i(t) + n(t) \tag{1}$$

Wherein, p is the number of the minimum utility threshold, the $n(t)$ is the output characteristic item which is directly mined, the $s_i(t)$ is the data fuzzy clustering measure value of the relative utility threshold value, and the $a(\theta_i)$ is the long-term health monitoring real-time monitoring data for each transaction set, The data update rules for long-term health monitoring of the bridge are as follows:

$$R_s^{(0)} = \sum_{n=0}^{k} \left\langle R_s^{(n)}, d_{\gamma n} \right\rangle d_{\gamma n} + R_s^{(k+1)} \tag{2}$$

Where, $R_s^{(n)}$ represents the structural scale characteristics of bridge long-term health monitoring real-time monitoring data, $d_{\gamma n}$ is the variation operation of bridge long-term health monitoring real-time monitoring data, and $R_s^{(k+1)}$ is the joint distribution coefficient of internal utility and external utility. According to the above analysis, the grid distribution structure model of bridge long-term health monitoring is constructed, and the minimum relative utility threshold analysis method and analytic hierarchy process (AHP) are used for data information fusion and adaptive feature extraction preprocessing [7].

### 2.2 Analysis of Real-Time Data Attribute Set for Long-Term Health Monitoring of Bridges

The long-term health monitoring real-time data of the bridge is fused with the information fusion and the association mining method information, a grid distribution structure model of the long-term health monitoring of the bridge is established, and the statistical characteristics of the long-term health monitoring data flow of the bridge are carried out, the vector quantization characteristic decomposition of the real-time data attribute set of long-term health monitoring of the bridge is carried out, the relevant data self-adaptive feature extraction is realized, the utility information is divided into the internal utility and the external utility in the long-term health monitoring real-time data clustering space of the bridge, and the main frequency characteristic quantity is obtained as follows:

$$X_p(u) = s_c(t)e^{j2\pi f_0 t} = \frac{1}{\sqrt{T}} rect(\frac{t}{T})e^{j2\pi(f_0 t + Kt^2)/2} \tag{3}$$

Where, $s_c(t)$ represents the feature attribute, $e^{j2\pi f_0 t}$ represents the structural similarity measure of the finite data set, $f_0$ is the main frequency of the feature, and $T$ represents the time delay of the global feature extraction. After extracting the global time feature, the finite dataset is obtained:

$$X = \{x_1, x_2, \cdots, x_n\} \subset R^s \tag{4}$$

Wherein, the real-time data set of bridge long-term health monitoring contains n samples, sample $x_i$, $i = 1, 2, \cdots, n$. The relationship between other related parameters extracted from real-time information is expressed as follows:

$$h(t) = \sum_i a_i(t)e^{j\theta_i(t)}\delta(t - iT_S) \tag{5}$$

Through the mining of association rules by the upper bound of itemset utility, the fuzzy index set of real-time data of bridge long-term health monitoring is obtained by using the basis function $d_{\gamma_0}$ for feature transformation and template matching.

$$\Lambda_0 = \left\{ \beta \in \Gamma : \left| \langle f, d_{\gamma_0} \rangle \right| \geq a. \sup_{\gamma \in \Gamma} \left| \langle f, d_\gamma \rangle \right| \right\} \tag{6}$$

The fuzzy hierarchical matching of bridge long-term health monitoring real-time data transmission is used for adaptive optimization, and the decentralized subspace method is used to obtain the statistical characteristic information flow of bridge long-term health monitoring real-time data:

$$X = F_\alpha \cdot x \tag{7}$$

The N subspace is searched, and the data clustering results of N subspace mining are fuzzy predictive control, and the output subsequence is obtained as:

$$X = [X_\alpha(0), X_\alpha(1), \cdots, X_\alpha(N-1)]^T \tag{8}$$

With the above feature extraction and processing of associated data, big data fusion of real-time data flow monitoring sequence is carried out by cloud computing, which reduces the overhead of data prediction and improves the accuracy of data monitoring and fuzzy clustering [8].

# 3 Automatic Detection and Optimization of Data Flow

## 3.1 Feature Extraction and Grid Region Clustering

On the basis of the above-mentioned grid distribution structure model for establishing the long-term health monitoring of the bridge and carrying out long-term health monitoring data flow monitoring statistical analysis of the bridge, an automatic monitoring and optimization design of the long-term health monitoring data flow of the bridge is carried out [9], a long-term health monitoring data classification method based on fuzzy correlation feature detection and mesh region clustering is presented in this paper. when the type of the interference vector $r_j$ in the long-term health monitoring intelligent real-time data flow monitoring system of the bridge is H, the maximum independent set $P(n_i) = \{p_k|pr_{kj} = 1, k = 1, 2, \cdots, m\}$ exists for all the node sets, the fuzzy correlation characteristic detection and the grid region clustering design are carried out, and the fuzzy membership function of the extracted data is as follows:

$$P_F = \sum_{j=k}^{N} \sum_{\sum u_i = j} \prod_{i=1}^{N} (P_{fi})^{u_i} (1 - P_{fi})^{1-u_i} \tag{9}$$

$$P_D = \sum_{j=k}^{N} \sum_{\sum u_i = j} \prod_{i=1}^{N} (P_{di})^{u_i} (1 - P_{di})^{1-u_i} \tag{10}$$

Wherein, $P_{fi}$ represents the clustering center of bridge long-term health monitoring real-time data fusion. $P_{di}$ is the sampling frequency of bridge long-term health monitoring real-time data. The average value of positive correlation characteristic components of bridge long-term health monitoring real-time data is obtained as:

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{11}$$

The real-time data matching detection of bridge long-term health monitoring is carried out in the decentralized subspace, and the variance of the relevant statistical features of the data monitoring output in the subspace is obtained:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} |x_i - \overline{x}|^2 \tag{12}$$

The variance reflects the oscillatory amplitude of the real-time data of bridge long-term health monitoring, and the information fusion and fuzzy chromatography analysis method are used to realize the information fusion of the real-time data of bridge long-term health monitoring [10]. The dynamic copy of the real-time data of bridge long-term health monitoring is obtained, and the results of AHP mining are as follows:

$$x_i^{'} = \frac{x_i}{\|x_i\|} = (\frac{x_{i1}}{\|x_i\|}, \frac{x_{i2}}{\|x_i\|}, \cdots, \frac{x_{iN}}{\|x_i\|}) \tag{13}$$

Wherein, the end value of real-time data of bridge long-term health monitoring $x_N$, is fuzzy scheduling according to prior knowledge and itemsets, which improves the accuracy and real-time performance of data monitoring [11].

### 3.2 Data Mining and Monitoring Output

The real-time data of long-term health monitoring of the bridge needs to be fused with the information fusion and the association mining method information, so that the relevant data self-adaptive feature extraction can be realized, and the prediction cost and the calculation complexity are reduced. the fuzzy clustering and information prediction of long-term health monitoring data flow of the bridge are carried out on the basis of the hierarchical clustering algorithm, and the fuzzy set quality of the real-time data attribute set of the long-term health monitoring of the bridge is adopted, $R_{u,v}$ is a cross-correlation function, the optimization index of the long-term health monitoring data flow monitoring of the bridge is $(RT_1, RT_2)$, the degree of fuzzy clustering is $RW$, the convergence constraint function of the long-term health monitoring data flow monitoring process of the bridge is as follows:

$$F_j = \sum_{k=1}^{n} X_{kj}, Q_j = \sum_{k=1}^{n} \left(X_{kj}\right)^2 \tag{14}$$

By adjusting the total utility ratio for many times, the time series of data flow for bridge long-term health monitoring is $\{x(t_0 + i\Delta t)\}$, $i = 0, 1, \cdots, N-1$. The length of sample set output from the automatic monitoring system of bridge long-term health monitoring is $N$. The fuzzy convergence control function of bridge long-term health monitoring is given as follows:

$$\begin{aligned} M_v = {} & w_1 \sum_{i=1}^{m \times n} (H_i - S_i) + M_h w_2 \sum_{i=1}^{m \times n} (S_i - V_i) \\ & + w_3 \sum_{i=1}^{m \times n} (V_i - H_i) \end{aligned} \tag{15}$$

In the above formula, the load of the long-term health monitoring real-time data prediction of the bridge is $V$, and after a group of cluster attribute characteristics $M_h$ is generated, the long-term health monitoring real-time data attribute set dispatching of the bridge is carried out, and the main component characteristic quantity of the long-term health monitoring data flow monitoring of the bridge is as follows:

$$s(t) = s_c(t)e^{j2\pi f_0 t} = \frac{1}{\sqrt{T}} rect(\frac{t}{T})e^{j2\pi(f_0 t + Kt^2)/2} \tag{16}$$

The vector quantitative feature decomposition function of the real-time data attribute set of bridge long-term health monitoring is as follows:

$$f(k) = \begin{cases} f(k-1) - \frac{1}{n}, 1 \leq k < n \\ 1, k = n \end{cases} \tag{17}$$

Wherein, $k$ represents the feature fusion center of real-time data for long-term health monitoring of distributed bridges, and the data queue is used as a Chunk to reorganize the time area. The information fusion set of predicting spatial vector in real-time data of bridge long-term health monitoring is expressed as follows:

$$P = \{p_1, p_2, \cdots p_m\}, m \in N \tag{18}$$

Wherein, m is the embedded dimension of the automatic monitoring system of bridge long-term health monitoring data stream. In the statistical regression analysis model, the training sample set of vector quantification feature decomposition of the attribute set of bridge long-term health monitoring real-time data is $X = [X_1, X_2, \cdots, X_k, \cdots, X_N]^T$. One of the training samples is $X_k = [x_{k1}, x_{k2}, \cdots, x_{km}, \cdots, x_{kM}]$, from the above feature decomposition results, it is known that the real-time data output of bridge long-term health monitoring meets the convergence condition, according to the above analysis, On the basis of hierarchical clustering algorithm, autoregression analysis is used to monitor bridge long-term health monitoring real-time data monitoring fuzzy clustering and information scheduling, so as to improve the accuracy of bridge long-term health monitoring real-time data monitoring. At the same time, the risk of long-term health monitoring data flow monitoring is reduced [9].

## 4 Analysis of Simulation Experiment

In order to test the application performance of this algorithm in realizing long-term health monitoring data flow monitoring of bridges, simulation experiments are carried out. Big data hierarchical cluster analysis of bridge long-term health monitoring real-time data monitoring is carried out by using C and Matlab 7 mixed programming. The sampling length of single group bridge long-term health monitoring real-time data time series is 800. The time interval of real-time data packet collection for bridge long-term health monitoring is 0.28 s, and characteristic sampling frequency $f_s = 4f_0 = 20\,\text{kHz}$, the maximum iteration number is 2000. According to the above simulation environment and parameters, the automatic monitoring and simulation analysis of bridge long-term health monitoring data flow is carried out, and the time-domain output of bridge long-term health monitoring data flow monitoring is shown in Fig. 1.

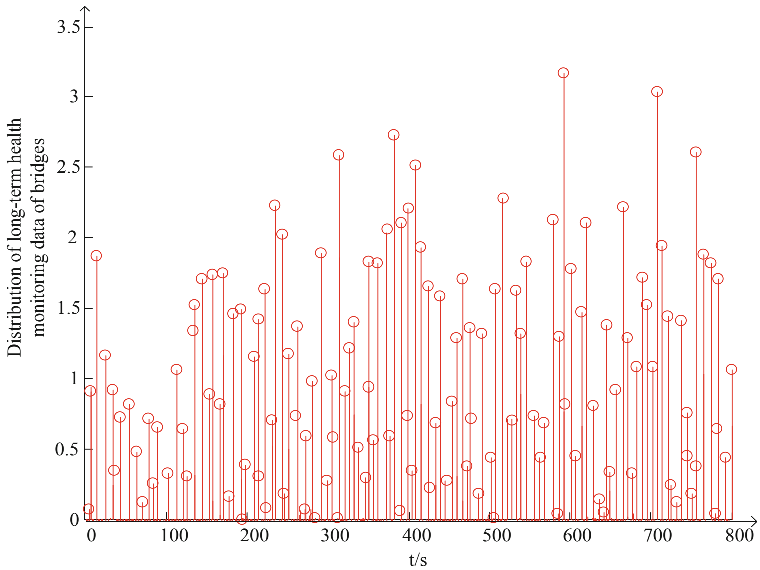


**Fig. 1.** Monitoring and output of long-term health monitoring data flow of bridge

Analysis Fig. 1 shows that the method can effectively realize the monitoring of long-term health monitoring data flow of the bridge, and analyze the convergence curve of the test data monitoring level, and the result is shown in Fig. 2.

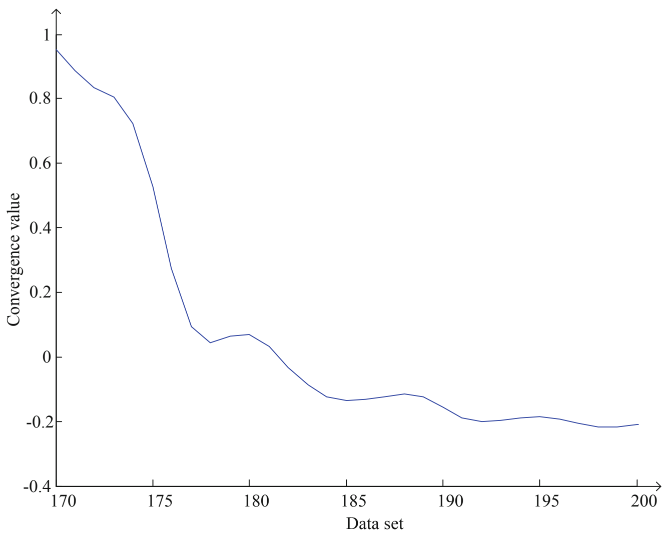


**Fig. 2.** Convergence test of data monitoring

Figure 2 shows that the convergence of this method for automatic monitoring of long-term health monitoring data flow of bridges is better, and the recall rate of data monitoring by different methods is tested. The comparative results are shown in Table 1, and Table 1 of the analysis shows that the method in this paper has good convergence for automatic monitoring of long-term health monitoring data streams of bridges. In this paper, the recall rate of bridge long-term health monitoring data flow automatic monitoring is high.

**Table 1.** Comparison of recall of data flow monitoring

| Data sample set | Proposed method | Reference [3] | Reference [4] |
|---|---|---|---|
| 100 | 0.943 | 0.843 | 0.887 |
| 200 | 0.987 | 0.889 | 0.954 |
| 300 | 0.997 | 0.921 | 0.969 |

## 5 Conclusions

In this paper, a classification method of bridge long-term health monitoring data based on fuzzy correlation feature detection and grid area clustering is proposed. The information fusion and fuzzy chromatography analysis method are used to realize the information fusion of the real-time data of bridge long-term health monitoring, and the adaptive feature extraction of related data is carried out. Excavate the positive correlation characteristic quantity of bridge long-term health monitoring real-time monitoring data flow, carry on the fuzzy clustering and information prediction of bridge long-term health monitoring data flow, and improve the accuracy of bridge long-term health monitoring real-time data monitoring. The simulation results show that the intelligent classification of bridge long-term health monitoring based on this method has high accuracy and low error rate, which improves the real-time performance of bridge monitoring. This method has important application value in improving the health monitoring ability of bridges.

**Acknowledgement.** Two heights "project of xichang university (LGLZ201824): settlement characteristics analysis and deformation prediction research of xigeda high-rise building with soil layer in xichang.

## References

1. Liu, Y., Yang, H., Cai, S., Zhang, C.: Single image super-resolution reconstruction method based on improved convolutional neural network. J. Comput. Appl. **39**(5), 1440–1447 (2019)
2. Xu, R., Zhang, J.G., Huang, K.Q.: Image super-resolution using two-channel convolutional neural networks. J. Image Graphic **21**(5), 556–564 (2016)

3. Megha, G., Yashpal, L., Vivek, L.: Analytical relation & comparison of PSNR and SSIM on babbon image and human eye perception using Matlab. Int. J. Adv. Res. Eng. Appl. Sci. **4** (5), 108–119 (2015)
4. Li, Y.F., Fu, R.D., Jin, W., et al.: Image super-resolution using multi-channel convolution. J. Image Graphics **22**(12), 1690–1700 (2017)
5. Dong, X.W., Yao, S.M., Wang, Y.W., et al.: Virtual sample image set based multi manifold discriminant learning algorithm. Appl. Res. Comput. **35**(6), 1871–1878 (2018)
6. Zhu, Y., Zhu, X., Wang, J.: Time series motif discovery algorithm based on subsequence full join and maximum clique. J. Comput. Appl. **39**(2), 414–420 (2019)
7. Wang, Z., Huang, M., et al.: Integrated algorithm based on density peaks and density-based clustering. J. Comput. Appl. **39**(2), 398–402 (2019)
8. Gui, J., Sun, Z.N., Jia, W., et al.: Discriminant sparse neighborhood preserving embedding for face recognition. Pattern Recogn. **45**(8), 2884–2893 (2012)
9. Dong, B., Chen, D., Jing, W.: Salient region detection method based on symmetric region filtering. Comput. Eng. **45**(5), 216–221 (2019)
10. Zhou, S.B., Xu, W.X.: A novel clustering algorithm based on relative density and decision graph. Control Decis. **33**(11), 1921–1930 (2018)
11. Tu, B., Chuai, R., Xu, H.: Outlier detection based on K-mean distance outlier factor for gait signal. Inf. Control **48**(1), 16–21 (2019)