Research paper

Damage identification of bridge structure model based on empirical mode decomposition algorithm and Autoregressive Integrated Moving Average procedure

Weijia Lu¹, Jiafan Dong², Yuheng Pan³, Guoya Li⁴, Jinpeng Guo⁵

Abstract: Time series models have been used to extract damage features in the measured structural response. In order to better extract the sensitive features in the signal and detect structural damage, this paper proposes a damage identification method that combines empirical mode decomposition (EMD) and Autoregressive Integrated Moving Average (ARIMA) models. EMD decomposes nonlinear and non-stationary signals into different intrinsic mode functions (IMFs) according to frequency. IMF reduces the complexity of the signal and makes it easier to extract damage-sensitive features (DSF). The ARIMA model is used to extract damage sensitive features in IMF signals. The damage sensitive characteristic value of each node is used to analyze the location and damage degree of the damaged structure of the bridge. Considering that there are usually multiple failures in the actual engineering structure, this paper focuses on analysing the location and damage degree of multi-damaged bridge structures. A 6-meter-long multi-destructive steel-whole vibration experiment proved the state of the method. Meanwhile, the other two damage identification methods are compared. The results demonstrate that the DSF can effectively identify the damage location of the structure, and the accuracy rate has increased by 22.98% and 18.4% on average respectively.

Keywords: ARIMA model, bridge, damage identification, EMD algorithm, multiple damage, structural health monitoring

¹Assistant Professor, MSc., Eng., Tianjin Chengjian University, Computer and Information Engineering Department, Tianjin, China, e-mail: luweijia@tcu.edu.cn, ORCID: 0000-0002-0234-9412
²MSc., Tianjin Chengjian University, Computer and Information Engineering Department, Tianjin, China, e-mail: jiafan_dong99@163.com, ORCID: 0000-0002-6416-6910
³Prof., PhD., Eng., Tianjin Chengjian University, Computer and Information Engineering Department, Tianjin, China, e-mail: panyuheng@tcu.edu.cn, ORCID: 0000-0003-3773-8810
⁴Prof., PhD., Eng., Tianjin Chengjian University, Computer and Information Engineering Department, Tianjin, China, e-mail: ligy@tcu.edu.cn@tcu.edu.cn, ORCID: 0000-0003-3224-2824
⁵Assistant Professor, MSc., Tianjin Chengjian University, Computer and Information Engineering Department, Tianjin, China, e-mail: 1571919251@qq.com, ORCID: 0000-0003-1705-2504
1. Introduction

According to the literature [1] there are about 1 million highway bridges in Europe, and 35% of these bridges are over 100 years old. Therefore, it is an urgent task to monitor the operational status of bridge. Structural health monitoring (SHM) system can monitor the development of structural damage in real time and effectively prevent sudden disasters [2]. The methods of damage identification, which is the key component of the SHM, are divided into two categories: modal parameter method and non-modal parameter method [3]. Modal parameter method employs structural modal parameters to reflect the structural damage, such as the natural frequency [4–6], mode shape [7, 8], modal curvature [9–11], modal flexibility [12], and modal strain energy [13]. However, structural parameters tend to be random and uncertain due to the complex and changeable application environment of bridge. In addition, the modal-based damage detection method is only sensitive to high strength damage and identifies false damage events affected by the noise. So many researchers turn their researches into the non-modal parameter identification method.

In the current researches, the non-modal damage identification method mainly includes the following contents: intelligent algorithm [14–16], Bayesian theory [17, 18], statistical methods [19–21] and signal processing techniques [22–28].

In the above methods, the statistical-based damage identification method is widely used to track and identify the data differences between the undamaged and damage, because this method dispenses with establishing an accurate finite element model, and has high anti-noise capability. A statistical method for damage detection has been presented [19], its application in a small wind turbine proves the availability and reliability of this method. Xin et al. [20] proposed a new bridge structure deformation prediction method based on Kalman filter, Autoregressive Integrated Moving Average (ARIMA), and generalized autoregressive conditional heteroscedasticity (GARCH), which can be applied to analyze and predict the bridge deformation of structure.

The damage detection method based on signal processing can display the broken physical characteristics and identifies the structural damage according to the change of the signal. Therefore, a powerful signal processing technology is crucial in identifying damage. In recent years, many signal processing approaches for structural damage identification have been proposed, such as fast Fourier transform (FFT) and its variant frequency response function (FRF), wavelet transform (WT), and blind source separation (BSS). Sulaiman et al. [23] presented a damage detection method based on FRF and model parameters, the experimental results indicated that the reduction of structural stiffness will affect the FRF value. Zenzen et al. [24] proposed an advanced method to identify the location and level of structural damage, using FRF data and optimization technology. Dilena et al. [25] utilized amplitude irregularity to detect and locate damage, combining interpolation damage detection (IDD) and FRF. These methods can realize the damage identification of the structure, but they are inefficient to analyze nonlinear and nonstationary signals. Besides, these methods cannot process the effect of noise, so potential errors are generated during identification. To solve nonlinear and nonstationary problems, continuous wavelet transform (CWT) is presented to decompose the difference of signals between the undamaged
and damaged [26], the change value of the signal is used to locate and detect damage. Patel et al. [27] used complex CWT to extract and analyze the discontinuity of the acceleration response signal, and utilized the wavelet coefficient to detect structural damage. Although wavelet transform can process nonlinear and nonstationary signals, the selection of wavelet generating function and the uncertainty of the decomposition level can easily affect the consequences of damage identification. Morovati et al. [28] proposes a damage identification method combining BSS and time-field analysis technology. Although BSS can explain the mixed characteristics in the measured response signal, this method still has limitations in analyzing signals with non-stationary characteristics.

From the above researches, the damage detection method based on time-domain signal processing cannot meet high-precision identification requirement, especially for nonlinear and nonstationary signals. As a result, Hilbert-Huang transform (HHT) is proposed to analyze signal. Qu et al. [29] analyzed the seismic damage of soil using edge spectrum identification theory based on HHT, experimental results show the HHT method is efficient to nonlinear and non-stationary signal. Also, Han et al. [30] proved that marginal spectrum in HHT can accurately identify the damage location, and has a strong anti-noise interference ability. Kelareh et al. [31] researched the influence of measurement noise and load types on the system, the result of simulations and experiments show that HHT technique can reduce the effect of noises. Through the improvement of HHT algorithm, a method is proposed based on empirical mode decomposition (EMD) and Shannon entropy index (SEI) [32,33], and the results showed that this method has good sensitivity to small damage.

In order to better extract the damage-sensitive features from the signals, a structural damage identification method combining EMD algorithm and ARIMA model is presented to locate and quantify structural damage. EMD decomposes nonlinear and non-stationary signals into different IMF according to frequency, which reduces the complexity of signals and makes it easier to extract damage sensitive features. The ARIMA model is used to extract the damage-sensitive features from the IMF signals, and the damage-sensitive features are defined as the variance difference of the residual sequence before and after the damage. The qualitative and quantitative analysis of the bridge damage structure is realized through the damage-sensitive characteristic values of each node. Considering that there are multiple damages in the actual engineering structure, a vibration experiment of a 6-meter steel-concrete composite bridge with multiple damages was carried out to prove the effectiveness of the proposed method.

2. Theoretical background

2.1. Empirical model decomposition (EMD)

EMD algorithm is a signal processing method established by Huang [34] based on HHT, which is not only suitable for nonlinear and non-stationary signal analysis but also suitable for linear and stationary signal analysis. According to the frequency information, EMD algorithm decomposes the original measured time series into trend term and multiple sub-
sequences of different frequency (i.e., IMFs). The decomposed IMF components contain local characteristic signals at different time scales of the original signal. EMD can retain the characteristics of the original signal to the greatest extent, the inherent fluctuation characteristics of the sequence are revealed by analysing the fluctuation information of each frequency component. The process of the EMD algorithm is as follows:

1. Find out all the extreme points (local maximum and minimum) of the original time series \( X(t) \). According to the extreme points and using the cubic spline difference function, the upper envelope \( U(t) \) and lower envelope \( L(t) \) of the sequence are fitted, and then the mean envelope \( m(t) \) is obtained (Eq. 2.1):

\[
m(t) = \frac{U(t) + L(t)}{2}
\]

where: \( U(t) \) – upper envelope sequence, \( L(t) \) – lower envelope sequence, \( m(t) \) – mean envelope sequence.

2. Subtract \( m(t) \) from original time series \( X(t) \) to get a new sequence \( h(t) \).

\[
h(t) = X(t) - m(t)
\]

where: \( h(t) \) – a new sequence, \( X(t) \) – original time series.

3. Determine whether the \( h(t) \) sequence satisfies two conditions: a) In the entire data range, the number of local extremums and zero crossings must be equal, or the number of differences must be at most 1. b) At any time, the average value of the envelope of the local maximum (upper envelope) and the envelope of the local minimum (lower envelope) must be zero. If a) and b) are satisfied, the sequence \( h(t) \) is the first IMF of the original sequence \( X(t) \), namely the IMF1; otherwise, then \( h(t) \) as a new \( X(t) \) continues to perform 1), 2) operations until both conditions are satisfied.

4. By subtracting the first mode component IMF1 from the original sequence \( X(t) \), a new sequence \( r_1 \) is obtained. And \( r_1 \) is decomposed by 1), 2) and 3) to obtain the second mode component IMF2 is obtained. If time series cannot be decomposed, the sequence that cannot be decomposed is called trend term \( r_n \).

\[
r_1 = X(t) - IMF_1
\]

where: \( r_i \) – trend term sequence, IMF1 – mode component.

In the process of measurement, the accuracy of the experimental data will be affected by noise. However, after EMD decomposition, the noise will be evenly distributed on each IMF. Through the analysis of IMFs, the influence of noise on the series is reduced and the accuracy of the experiment can be improved.

### 2.2. Autoregressive integrated moving average model (ARIMA)

The ARMA model was proposed by mathematician Walker in 1931 and has been widely used in various fields. Because the ARMA model only relies on the output data of the system. It does not depend on the accurate finite element model and need to consider the
relevant information of the load input. The model is widely used in the time series analysis of structural damage identification.

In the ARMA model, the value of time is not only related to the past time value, but also to the previous error term. The sequence value of time can be expressed as:

\[
(2.4) \quad x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \ldots + \varphi_p x_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q}
\]

where: \( \varphi_i \) \( (i = 0, 1, \ldots, p) \) \( - \) the coefficient of the order autoregressive term, \( \beta_j \) \( (j = 0, 1, 2, \ldots, q) \) \( - \) the order moving average term coefficient, \( \varepsilon_t \) \( - \) a white noise sequence with zero mean and variance.

The ARMA model has the following properties: (1) A random event sequence can be represented by an autoregressive moving average model. The sequence can be explained by its own past or lag values and random disturbance terms. (2) If the series is stationary, i.e. it does not change over time. It can be seen from the above formula, the modeling idea of ARMA is that the system output of the current state or a certain time in the future is only related to the output and error term of the past time. The correlation between the inherent attribute information of the system and the output time series information is integrated into the model. in the coefficient.

The lag operator \( B \) is introduced here. For the time series value \( x_t \) at any time, the lag operator satisfies:

\[
(2.5) \quad B(x_t) = x_{t-1}
\]

where: \( B(x_t) \) \( - \) lag operator, \( x_{t-1} \) \( - \) time series at time \( t - 1 \).

Similarly, a higher-order lag operator can be defined. For any time series \( x_t \), the higher-order lag operator satisfies:

\[
(2.6) \quad B^k(x_t) = x_{t-k}
\]

where: \( B^k(x_t) \) \( - k \) order lag operator

Therefore, the Eq. (2.4) can be written as:

\[
(2.7) \quad \left(1 - \sum_{i=1}^{p} \varphi_i B^i\right) x_t = \left(1 - \sum_{j=1}^{q} \beta_j B^j\right) \varepsilon_t
\]

When \( \varphi(B) = \left(1 - \sum_{i=1}^{p} \varphi_i B^i\right) \), \( \beta(B) = \left(1 - \sum_{j=1}^{q} \beta_j B^j\right) \), the Eq. (2.7) can be converted into:

\[
(2.8) \quad \varphi(B)x_t = \beta(B)\varepsilon_t
\]

After shifting the formula, we get:

\[
(2.9) \quad x_t = \frac{\beta(B)}{\varphi(B)} \varepsilon_t
\]
It can be seen from the above formula, the ARMA model can be converted into the form of a transfer function, which is represented by white noise as the environmental excitation, and has an equivalent relationship with the actual system in time series.

Therefore, in the same structural system, if the input of the two systems is the same, the two ARMA models will only produce smaller residuals. On the contrary, if the two structural systems are different, the ARMA models will produce larger residuals.

2.3. Damage index

EMD algorithm is a signal processing method established by Huang based on HHT, which is suitable for nonlinear and non-stationary signal analysis. According to the frequency information, EMD algorithm decomposes the original measured time series into trend term and multiple sub-sequences of different frequency (i.e., IMFs). The decomposed IMF components contain local characteristic signals at different time scales of the original signal. EMD can retain the characteristics of the original signal to the greatest extent, the inherent fluctuation characteristics of the sequence are revealed by analysing the fluctuation information of each frequency component.

The bridge vibration signal usually shows the characteristics of non-stationary, non-linear, complex spectrum components and so on. At any point, these data contain multiple fluctuation modes, decomposing the original data can retain the characteristics of the original signal to the greatest extent. The noise data in the signal is decomposed to IMFs at the same time, it is convenient for noise reduction. In the experiment, the EMD algorithm is used to process the time-series, signals at different nodes in different states are decomposed into IMFs respectively, the inherent fluctuation characteristics of the sequence are revealed by analyzing the fluctuation information at different scales. Figure 1 shows the result of EMD decomposition.

After the EMD algorithm, the time series signal is decomposed into various modal components of different frequencies. The ARMA model is used to predict each modal component, and the residual of each modal component signal is obtained. At this time, the error between the predicted value of the residual ARMA model and the actual value is calculated by using the obtained residual to solve its variance to construct the damage-sensitive characteristic factor. At this time, the variance of the residual is used to indicate the degree of agreement between the estimated value of the model and the real data. When the variance is smaller, the degree of agreement is considered to be better, and the degree of damage is considered to be smaller.

In order to accurately identify and quantify the damage site of the structure, the damage index is defined as:

\[ DSF = |\sigma^2(\varepsilon_d) - \sigma^2(\varepsilon_u)| \]

where: \( \sigma^2(\varepsilon_d) \) – the conditional variance of the residuals of the structural model under the damage condition, \( \sigma^2(\varepsilon_u) \) – the conditional variance of the residuals of the structural model under the intact condition.
When the DSF value is greater than 0, the structure is damaged, and when the DSF value is equal to 0 or close to 0 within the allowable range of error ($\pm (0.1 \cdot \max(\text{DSF}))$), the structure is not damaged. When identifying damage, qualitatively judge the damage degree of the structure according to the size of the DSF mutation.

### 2.4. Index evaluation

*Precision*, *Recall* and valuation value $F$ are used to determine the accuracy of damage identification results.

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{2.11}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{2.12}
\]

\[
F = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{(\text{Precision} + \text{Recall})} \tag{2.13}
\]

where: $TP$ – the number of true positive samples, $FP$ – the number of false-positive samples, $FN$ – the number of false-negative samples.

According to the $F$ criterion, the larger of $F$, the higher the accuracy of damage index identification.
3. Experimentation

3.1. Experimental setup

A 6-meter steel-concrete bridge model is designed in the laboratory. The cross-section size of the model is $900 \times 90$ mm, this model is divided into 19 units and 18 nodes. The elastic modulus of the model material are $2.0795 \times 10^8$ kN/m$^2$ and $2.10 \times 10^8$ kN/m$^2$, and the Poisson’s ratio is $\mu = 0.3$. There are 18 removable shear connectors embedded in the model, and all shear connectors are tightened to 120 N-m through the torque wrench. The nut is $12 \times 1.7$ mm, and the bolts length is 50 mm. The concrete slab and the steel beam are closely connected, and the beam slab is poured separately. Each shear connector is composed of the screw rod of the beam arch and it is tied in the embedded nut cap. The width of the bridge model support is 0.5 m, the height is 0.8 m, and the width of the bridge is 0.9 m. The designed bridge model are shown in Figs. 2 and 3.
Firstly, the vibration data are obtained through the vibration experiment in the designed bridge model. The NI acceleration sensors are fixed at the corresponding position of the steel beam structure through the magnetic bearing, as shown in Figs. 4 and 5. The overall frame design of the bridge is shown in Figs. 5 and 6. The influence of sensor quality on the experiment was not considered. The data is acquired by a dynamic acquisition instrument INV306U with the maximum sampling rate 1 MHz and recorded by a computer.

3.2. Study cases

A bridge has a large number of nodes, different damage factors will cause different damages to these nodes of bridge. In order to better simulate the damage of the bridge structure, the shear connectors at different nodes are removed simultaneously in this experiment. Since the removed shear connectors have a great influence on the surrounding adjacent nodes, the damage degree is set as the ratio of the number of removed connectors to the total number of adjacent nodes. Five different damage states were set up in the test,
the vibration acceleration under health and damage conditions was tested by the hammering method. Experimental damage conditions are shown in Table 1.

Table 1. Data labels of structural state conditions

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
<th>Damage</th>
</tr>
</thead>
<tbody>
<tr>
<td>OM1</td>
<td>Remove a row of shear connectors of S8, S12 and S15</td>
<td>10%</td>
</tr>
<tr>
<td>OM2</td>
<td>Remove two row of shear connectors of S8, S12 and S15</td>
<td>20%</td>
</tr>
<tr>
<td>OM3</td>
<td>Remove a row of shear connectors of S8 and S12</td>
<td>8%</td>
</tr>
<tr>
<td>OM4</td>
<td>Remove four row of shear connectors of S8, S12 and S15</td>
<td>40%</td>
</tr>
<tr>
<td>OM5</td>
<td>Remove two row of shear connectors of S8 and S12</td>
<td>15%</td>
</tr>
</tbody>
</table>

3.3. Results

The Figure 7 shows the identification of bridge damage structure using the ARIMA model, Kalman-ARIMA model, and EMD-ARIMA model. When the structure is damaged,
the DSF values of the above three methods are mutated, and the DSF value of damaged nodes are significantly greater than 0, which fully reflects that our method can effectively identify the damage of the structure. It can be seen from Fig. 7 that the DSF values of nodes with different damage degrees are different, and with the increase of damage degree, the DSF value increases, which can preliminarily reflect that there is a linear correlation between the damage degree of the structure and the DSF. EMD-ARIMA model based on structural damage identification can effectively reduce the probability of false recognition, from the $F$ value can be obtained (Table 2). Compared with ARIMA model and Kalman-ARIMA model, the EMD-ARIMA model used in this paper has the best effect, the accuracy of recognition is improved obviously under each condition.

Table 2. Damage identification result analysis

<table>
<thead>
<tr>
<th>State</th>
<th>ARIMA</th>
<th></th>
<th></th>
<th>Kalman-ARIMA</th>
<th></th>
<th></th>
<th>EMD-ARIMA</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision/%</td>
<td>F/%</td>
<td></td>
<td>Precision/%</td>
<td>F/%</td>
<td></td>
<td>Precision/%</td>
<td>F/%</td>
<td></td>
</tr>
<tr>
<td>OM1</td>
<td>60.0</td>
<td>75.0</td>
<td></td>
<td>75.0</td>
<td>85.7</td>
<td></td>
<td>75.0</td>
<td>87.5</td>
<td></td>
</tr>
<tr>
<td>OM2</td>
<td>60.0</td>
<td>75.0</td>
<td></td>
<td>60.0</td>
<td>75.5</td>
<td></td>
<td>75.0</td>
<td>87.5</td>
<td></td>
</tr>
<tr>
<td>OM3</td>
<td>40.0</td>
<td>57.0</td>
<td></td>
<td>40.0</td>
<td>57.1</td>
<td></td>
<td>66.7</td>
<td>80.2</td>
<td></td>
</tr>
<tr>
<td>OM4</td>
<td>37.5</td>
<td>54.5</td>
<td></td>
<td>50.0</td>
<td>66.7</td>
<td></td>
<td>75.0</td>
<td>85.7</td>
<td></td>
</tr>
<tr>
<td>OM5</td>
<td>28.6</td>
<td>44.4</td>
<td></td>
<td>28.6</td>
<td>44.4</td>
<td></td>
<td>66.7</td>
<td>80.0</td>
<td></td>
</tr>
</tbody>
</table>

It can be seen from the test results that the error of damage identification results under working conditions 1, 2, and 4 is small. Therefore, the peak values of damage indexes under different damage conditions of each node in Fig. 7 are extracted, and the relationship curve between the peak value of indexes and the damage degree is fitted. The peak values of DSF indexes at nodes 5, 8, and 10 are selected for fitting. The fitting polynomial times are set to 3, and the fitting damage degree function equations are respectively:

\[
y_5 = 0.190427x^3 - 0.158638x^2 + 0.041907x - 0.001826
\]

(3.1)

\[
y_8 = 0.218747x^3 - 0.164436x^2 + 0.040985x - 0.001586
\]

(3.2)

\[
y_{10} = 0.100987x^3 - 0.067962x^2 + 0.015779x - 0.000088
\]

(3.3)

The fitting curve of the structural damage relationship is shown in Figure 8. The damage degree increases from 10% to 40%, the damage indexes of nodes 5, 8, and 10 increases gradually. The relationship between the structural damage degree and the maximum mutation value of the DSF index of the corresponding unit is fitted. As shown in Figure 8 the mutation value of the DSF index in the damage position increases linearly with the increase of the damage degree. The damage degree is greater than 40%, the increase of the mutation value of the index increases rapidly. It can be inferred that the structural components at the damage position have entered the plastic failure stage [35,36]. Due to the measurement error in the measurement process, the fitted DSF value has a negative value in a healthy
state. The cumulative effect of damage is considered in the curve, and the fitting curve of damage degree can qualitatively verify the correctness of the growth trend.

![Damage Degree Fitting Curve](image)

**Fig. 8. Damage degree fitting curve**

The Figure 9 is the comparison of the identification effect of the three methods on the damage nodes under the OM1, OM2 and OM3 states respectively. The DSF values of the three methods at the damage nodes show a sudden rise. However, interestingly,

![Comparison of Recognition Result](image)

*Fig. 9. Comparison of recognition result. (a) OM1. (b) OM2 and (c) OM4*
compared with the ARIMA method and the Kalman-ARIMA method, the DSF value of the EMD-ARIMA method is more obvious, and it is more sensitive to the identification of the damaged structure.

4. Conclusions

This paper presents a bridge structure damage identification method based on EMD and ARIMA model. In this work, these data are obtained from damaged structures under different working conditions. The EMD is applied to process non-stationary signal without presetting the basis function. The ARIMA model is employed to analyse data and construct damage indicators, and the indicators are used to verify damage of structure. The following conclusions can be obtained:

1. The damage identification method based on EMD-ARIMA is effective and correct. Besides, the proposed DSF can also reflect the state of the structure.
2. In the case of multiple injuries, this method can effectively identify the damage at different locations and has a high accuracy.
3. The damage identification index proposed by this method is more sensitive to damage.
4. However, the number of nodes in this article is limited, and the results of complex bridges need to be further verified.

References


Received: 2022-03-18, Revised: 2022-09-20