

Hybrid Methodology for Structural Health Monitoring Based on Immune Algorithms and Symbolic Time Series Analysis

Rongshuai Li, Akira Mita, Jin Zhou

Department of System Design Engineering, Keio University, Yokohama, Japan.
Email: lrs0809@gmail.com, mita@sd.keio.ac.jp, zhoujin5120@gmail.com

Received October 17th, 2012; revised November 28th, 2012; accepted December 4th, 2012

ABSTRACT

This hybrid methodology for structural health monitoring (SHM) is based on immune algorithms (IAs) and symbolic time series analysis (STSA). Real-valued negative selection (RNS) is used to detect damage detection and adaptive immune clonal selection algorithm (AICSA) is used to localize and quantify the damage. Data symbolization by using STSA alleviates the effects of harmful noise in raw acceleration data. This paper explains the mathematical basis of STSA and the procedure of the hybrid methodology. It also describes the results of an simulation experiment on a five-story shear frame structure that indicated the hybrid strategy can efficiently and precisely detect, localize and quantify damage to civil engineering structures in the presence of measurement noise.

Keywords: Structural Health Monitoring; Adaptive Immune Clonal Selection Algorithm; Symbolic Time Series Analysis; Real-Valued Negative Selection; Building Structures

1. Introduction

Structural health monitoring (SHM) is a vast, interdisciplinary research field whose literature spans several decades. The focus of SHM research is the detection, localization, and quantification of damage in a variety of structures. Broadly speaking, SHM techniques for detecting, localizing, and quantifying damage rely on measuring the structural response to ambient vibrations or forced excitations. Ambient vibrations can be caused by earthquakes, wind, or passing vehicles, and forced vibrations can be delivered by hydraulic or piezoelectric shakers. SHM techniques infer the existence, location and severity of damage by detecting differences in local or global structural responses before and after the damage occurs.

Some success has been achieved with various heuristic optimization algorithms. The annealing algorithm (SA) and genetic algorithm (GA) methods have been used to accurately describe the dynamic behavior of structures [1]. Cunha & Smith used GAs to identify the elastic constants of composite materials [2]. Particle swarm optimization (PSO) has been used to estimate the severity of damage and identify the parameters of shear frame building structures [3]. An improved clonal selection algorithm (CSA), called adaptive immune CSA (AICSA), has been used for structural damage localization and quanti-

fication [4,5]. Moreover, recently, a novel pattern identification technique, called symbolic time series analysis (STSA), was developed. The core concept of STSA is the identification of statistical patterns from symbol sequences generated by coarse-graining of time series data. STSA for anomaly detection in complex systems [6] has the potential to deal with noise. Several case studies [7-9] have shown that STSA is more effective at anomaly detection than pattern recognition techniques such as principal component analysis and neural networks. STSA has also been used for fault detection in electromechanical systems, e.g., three-phase induction motors [10].

In this paper, we propose a hybrid methodology combining immune algorithms (real valued negative selection (RNS) and AICSA) and symbolic time series analysis (STSA) for detection, localization and quantification of damage to structural systems. In this methodology, RNS detects damage, and AICSA localizes and quantifies it by minimizing the Euclidean distance between the state sequence histogram (SSH) that STSA gets by transforming the raw acceleration data. We mathematically show that STSA improves noise immunity, and our experimental results show that this hybrid strategy can efficiently and precisely detect, localize and quantify damage to civil engineering structures in the presence of measurement noise.

2. Immune Algorithms and STSA

2.1. Real-Valued Negative Selection and Adaptive Immune Clonal Selection Algorithm

The negative selection (NS) algorithm [11] was inspired by observation of the activity of the human immune system, in particular, the selection process that takes place inside the thymus. In this process, T-cells that recognize the body's own cells (self cells) are eliminated; this guarantees that the remaining T-cells will recognize only foreign molecules. Gonzalez *et al.* [12] proposed a new negative selection algorithm that uses a real-valued representation of the self/non-self space. RNS tries to alleviate some of the drawbacks of NS while using the higher-level-of-representation real space to speed up the detector generation process. Refer to [13] for the detailed procedure of RNS for damage detection.

Inspired by the clonal selection principle (CSP), the clonal selection algorithm (CSA) has been used to deal with optimization problems, thanks its superior search capability compared with classical optimization techniques [14]. CSP explains how an immune response is mounted when a non-self antigenic pattern is recognized by B cells. In natural immune systems, only the antibodies that can recognize the intruding antigens are selected to proliferate by cloning [15]. Hence, the main idea of CSA is that those cells (antibodies) capable of recognizing the non-self cells (antigens) will proliferate. Although CSA has great advantages over GA, it is still difficult to solve complex problems with it. In order to solve complex problems, AICSA embodies three strategies: secondary response, adaptive mutation regulation, and vaccination to speed up CSA's convergence and ability to find the global optimum. For more information about AICSA, please refer to [16].

2.2. Symbolic Time Series Analysis

It may be appropriate to say that, while classical data analysis focuses on individuals, symbolic data analysis deals with concepts, a less specific type of information. The original time series signals are converted into sequences of discrete symbols, and the statistical features of the symbols can be used to describe the dynamic statuses of a system.

Consider a structural system Σ . The raw acceleration data can be recorded by using sensors. A section of this data, denoted as $\{\ddot{x}_0, \ddot{x}_1, \dots, \ddot{x}_{T-1}\}$, can be obtained by sliding a rectangular window of length T along the time series of the raw acceleration data. The first step is to transform the raw acceleration data into a binary symbol series $\{\sigma_0, \sigma_1, \dots, \sigma_{T-1}\}$. $\sigma_i (i \in [0, T-1])$ equals

“0” or “1” by setting a partition line. After that, we select an integer $r \geq 1$ (word length) and define the symbolic state at time t as the vector s_t containing the subsequent r output symbols, namely,

$$s_t = [\sigma_t, \sigma_{t+1}, \dots, \sigma_{t+r-1}], t \in [0, T-r+1] \quad (1)$$

s_t defines a state series $\{s_0, s_1, \dots, s_{T-r+1}\}$. A binary coded s_t should be transformed into the decimal domain, and note that s_t can take $Q = 2^r$ possible values (called “states”), which can be listed as a finite set $S = \{0, 1, \dots, Q-1\}$. We can then derive the statistics of the symbolic state, *i.e.*, compute the vector of the observed state frequencies $D = [d^0, d^1, \dots, d^{Q-1}]$, where d^i (integer $i \in [0, Q-1]$) is the number of occurrences of $S = i$. Also, since there are $T-r+1$ states in the state series in total, D can be normalized as $\frac{D}{T-r+1}$. The

histogram of S and D is called the state sequence histogram (SSH).

3. Proposed Method

3.1. Introduction

In our methodology, the self (non-self) element is an SSH gotten by STSA from raw acceleration data from a healthy (damaged) structure. The detector is represented as a redistribution of the states of self elements.

We introduce an index, the relative state sequence histogram error (RSSHe), to measure the distance between two histograms:

$$\begin{cases} SSH_a = [d_a^0, d_a^1, \dots, d_a^{Q-1}] \\ SSH_b = [d_b^0, d_b^1, \dots, d_b^{Q-1}] \end{cases} \quad Q = 2^r$$

$$RSSHe = \sqrt{\frac{\sum_{i=0}^{Q-1} (d_b^i - d_a^i)^2}{\sum_{i=0}^{Q-1} (d_a^i)^2}} \quad (2)$$

where $d_{a/b}^i$ is the frequency of state i in SSH_a or SSH_b .

3.2. Procedure

Our methodology has two stages: damage detection using RNS and damage localization and quantification using AICSA. The procedure is as follows (see **Figure 1**):

Stage 1: Damage detection.

1) Training phase

a) To create a subseries of raw acceleration data, a rectangular window of length T is chosen and slid along the raw acceleration data series (length L). A set of subseries numbering $j = 1, 2, \dots, L-T+1$ is created in this way. STSA (Section 2.1) is applied to each sub-series

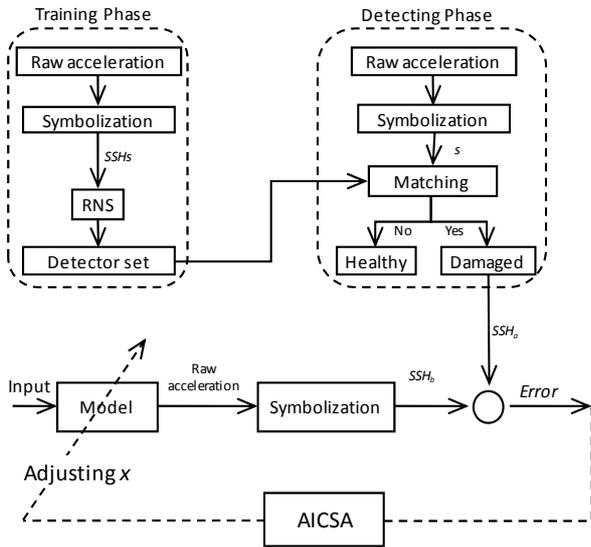


Figure 1. Procedure of IA and STSA for structural health monitoring.

to create a corresponding SSH_j ($j = 1, 2, \dots, L - T + 1$).

b) To create the self set, the distances ($RSSH_e$) between the j th SSH and all previous SSH s in the self set is calculated. If one of the distances is less than the predefined self radius r_s , the j th SSH is discarded. Otherwise, it is stored in the self set as a new self element. This procedure is repeated until the predefined criteria is reached.

c) RNS is applied to the self set to generate detectors.

2) Detection phase

a) New raw acceleration data of a structure is symbolized into a corresponding SSH_j by using the same method as in step a) of the training phase.

b) When matching SSH_j with the detectors generated in the training phase, if any detectors are activated (the distance between them is less than a certain value), a signal indicating the occurrence of the abnormal structural state will be given.

Stage 2: Damage localization and quantification.

In the research field of structural parameter identification, the time response of the system is usually compared with that of a parameterized model using a norm or some performance criterion to give us a measure of how well the model explains the system.

Suppose that damage to a structure has been detected by the above procedure. Moreover suppose that there is a parameterized model able to capture the behavior of the physical system, and this model depends on a set of n parameters, *i.e.*, $x = (x_1, x_2, \dots, x_n)^T \in R^n$. Given a candidate parameter value x and a guess \hat{X}_0 of the initial state, $\hat{y}(t_i)$ ($i = 1, \dots, T$), the value of the parameterized model, *i.e.*, the identified system at the i th discrete time step, can be obtained. Hence, the problem of system

identification boils down to finding a set of parameters that minimize the prediction error between the system output $y(t_i)$, which is the measured data, and the model output $\hat{y}(x, t_i)$, which is calculated at each time instant t_i .

Usually, our interest would lie in minimizing the predefined error norm of the time series outputs, *e.g.*, the following mean square error (MSE) function,

$$f(x) = \frac{1}{T} \sum_{i=1}^T \|y(t_i) - \hat{y}(x, t_i)\|^2 \quad (3)$$

where $\|\cdot\|^2$ represents the Euclidean norm of vectors. Formally, the optimization problem requires one to find a set of n parameters $x^* \in R^n$ so that a certain quality criterion is satisfied, namely, that the error norm $f(\bullet)$ is minimized. The parameters of the model are taken to be equivalent to the parameters of the structure.

In our methodology, instead of comparing raw acceleration data directly, the Euclidean distance of SSH s from the structure and model is used as a measure of the distance between the system output SSH_a and model output SSH_b .

3.3. Generation of Detectors

If window length is T and word length is r , the number of states in SSH is 2^r and the minimum changeable unit is $\frac{1}{T-r+1}$, then the total number of possible

distributions N_{SSH} of SSH is equal to one classic combination problem, which is 'put $T-r+1$ same balls in 2^r different boxes, and the combinatorial number is:

$$N_{SSH} = C_{T-r+2^r}^{2^r-1}. \text{ Supposing that the } SSH \text{ can capture}$$

the dynamic features of a structure, the normal state of the structure is represented by SSH s (self), the damage to that structure can be expressed by other distributions of states in the SSH s (non-self). Instead of randomly generating candidate detectors (which may result in creating an impossible distribution of states in the SSH), a novel procedure to create candidate detectors on the basis of self elements is proposed. The procedure is described below (N_s is the number of self elements).

1) Redistribution: Randomly choose n (integer $n \in [2, 2^r]$) states from SSH_j ($j \in [1, N_s]$). Randomly changing the frequency values of the n chosen states and keep their sum unchanged. Note that the minimum unit of change is $\frac{1}{T-r+1}$. The newly created SSH

by this procedure is used as the candidate detector.

2) Self recognition: Calculate the distance between candidate detector and all self elements. If the minimum distance $dist_s < r_s$ (r_s is self radius), delete it. If not, keep it for the next step.

3) Recognition of existing detectors: Calculate the distance between the candidate detector and the existing detectors. If the minimum distance $dist_d < r_d * (1 - r_o)$ (r_d is the radius of existing detectors; r_o is the overlap rate, defined in 3.4), delete it. If not, store it as a new detector. Set $dist_d$ as the radius of the new detector.

Repeat Steps (1) - (3) until the predefined stopping criterion is reached.

3.4. Noise Immunity

Usually the raw acceleration data is chosen as input, and the Euclidean distance of raw acceleration data is used as the damage index or objective function; the problem is that the possible number of representations of raw acceleration data is infinite for any Euclidean distance. In other words, the raw acceleration data is easily affected by noise. STSA is a coarse graining process that is robust against measurement noise. It is expected that small changes in time series data do not affect the symbolized data. Therefore, it can be assumed that a certain band of states represents similar dynamic status of the dynamic structural system, and instead of having an infinite number of representations of the Euclidean distance in the case of using raw acceleration data as input, the *SSH* representation is finite.

Taking a SDOF (single degree of freedom) system as an example, the dynamic equation can be represented as:

$$m\ddot{x} + c\dot{x} + kx = F(t) \quad (4)$$

where m , c , and k are respectively the mass, damping ratio, and stiffness of the system, F is the force linked to the ground acceleration. By dividing both sides by the mass m , the equation of motion becomes

$$\ddot{x} + 2\xi\omega_0\dot{x} + \omega_0^2x = \frac{F}{m} \quad (5)$$

where

$\omega_0 = \sqrt{\frac{k}{m}}$: undamped natural frequency of the oscillator

$\xi = \frac{c}{c_c} = \frac{c}{2\sqrt{mk}}$: damping ratio

The value

$$c_c = 2\sqrt{mk} \quad (6)$$

is called the critical damping.

Acceleration \ddot{x} is:

$$\ddot{x} = -\omega_0^2x - 2\xi\omega_0\dot{x} + \frac{1}{m}F \quad (7)$$

At time t , \ddot{x}_t can be obtained as

$$\ddot{x}_t = -\omega_0^2x_{t-1} - 2\xi\omega_0\dot{x}_{t-1} + \frac{1}{m}F_{t-1} \quad (8)$$

Since the window length is T and word length is r , all the raw acceleration data in the window is $\ddot{x} = \{\ddot{x}_0, \ddot{x}_1, \dots, \ddot{x}_{T-1}\}$. The mean value of \ddot{x} is:

$$\begin{aligned} \bar{\ddot{x}} &= \frac{\sum_{t=0}^{T-1} \ddot{x}_t}{T} \\ &= \frac{\sum_{t=0}^{T-1} \left(-\omega_0^2x_{t-1} - 2\xi\omega_0\dot{x}_{t-1} + \frac{1}{m}F_{t-1} \right)}{T} \end{aligned} \quad (9)$$

This mean value of \ddot{x} is for the noise-free case. In case of raw acceleration data polluted by noise, the acceleration data at time t is

$$\ddot{x}'_t = \ddot{x}_t + \sigma_t \quad (10)$$

where σ_t is the value of noise at time t . The partition line (the mean value of \ddot{x}'_t) is:

$$\bar{\ddot{x}}' = \sum_{t=0}^{T-1} (\ddot{x}_t + \sigma_t) = \bar{\ddot{x}} + \sum_{t=0}^{T-1} \sigma_t \quad (11)$$

For a sufficiently long window length T and word length r , $\sum_{t=0}^{T-1} \sigma_t \approx 0$, which means that noise will not affect the partition line.

A random sample of *SSH* is defined as SSH_a . From Equation (2), any other *SSH* (i.e., SSH_b) will be some distance away from SSH_a .

$$\begin{aligned} \Delta SSH &= SSH_b - SSH_a \\ &= (\Delta SSH_0, \Delta SSH_1, \dots, \Delta SSH_{2^r-1}) \\ \Delta SSH_i &= SSH_{b,i} - SSH_{a,i} \quad (i \in [0, 2^r - 1]) \end{aligned} \quad (12)$$

where

$$\sum_{i=0}^{i=2^r-1} \Delta SSH_i = 0 \quad \text{and} \quad \sum_{i=0}^{i=2^r-1} (\Delta SSH_i)^2 = \varepsilon \quad (13)$$

where $i \in [0, 2^r - 1]$ and $\varepsilon \geq 2$ is an even number.

From reference [18], since the minimum unit of change in *SSH* is $\frac{1}{T-r+1}$ and the dimension of *SSH* is

2^r , there is a possibility that not only one but many different *SSHs* will have the same *RSSHe* as SSH_a . This possibility occurs when Equation (13) is satisfied.

The minimum number of SSH_b s that have the same *RSSHe* as SSH_a occurs when the difference between SSH_b and SSH_a is only one minimum unit; in that case, $\varepsilon = 2$, the minimum number is: $N_{\min} = 2^r(2^r - 1)$. If there are two different minimum units between SSH_b and SSH_a , ε equals 4, 6 or 8.

$\varepsilon = [2 * \theta, 2 * \theta + 2, \dots, 2 * \theta^2]$, where θ is the number of different units between SSH_b and SSH_a . The number of different values of ε is $N_\varepsilon = \theta^2 - \theta + 1$. For every certain value of ε can be called as a representative band, then the number of representative band is N_ε . Also, we can see that for a certain combination of $\Delta SSHs$, when α

is the number of elements for which

$\Delta SSH_i \neq 0, (i \in [0, 2^r - 1])$, the possible combinatorial number is $C_{2^r}^\alpha$. For example, when $r=9$ and α is the minimum 2, the combinatorial number is 130816, which is a big number, and we should note this is only the minimum one. The conclusion is that STSA has very good fault tolerance.

In addition, since the total number of units in SSH is $T-r+1$, $\theta \leq T-r+1$. For each representative band ε , there will be many representations of SSH_b ; in other words, a small change in the raw acceleration data (say, due to noise) will not affect the $RSSHe$ between SSH_b and SSH_a .

3.5. Indexes Used in the Performance Evaluation

Two indexes are used to evaluate the classification accuracy: the detection rate (DR) and the false alarm rate (FAR). DR is the ratio of correctly classified negative elements to the total negative elements, and the FAR is the ratio of incorrectly classified positive elements to the total negative elements. Four values are needed to calculate DR and FAR: the number of true positives (TP, positive elements identified as positive), true negatives (TN, negative elements identified as negative), false positives (FP, negative elements identified as positive), and false negatives (FN, positive elements identified as negative). DR and FAR are calculated as

$$DR = \frac{TN}{TN + FP} \quad (14)$$

$$FAR = \frac{FN}{TP + FN} \quad (15)$$

In the damage localization and quantification stage, the severity of damage is defined as

$$Severity = \frac{|k_i - k_r|}{k_r} \times 100\% \quad (16)$$

where k_i and k_r are the identified stiffness and real stiffness of the i th story, respectively.

4. Numerical Simulation

4.1. Description of Model

For simplicity and generality, we used a five-story shear frame structure as a representative case to verify the performance of our methodology, and we modeled it as a five degree-of-freedom lumped mass system (**Figure 2**). **Table 1** lists the structural and modal parameters of the structure.

The dynamic equation is [17]:

$$M\ddot{X}_i + C\dot{X}_i + KX_i = f(t) \quad (17)$$

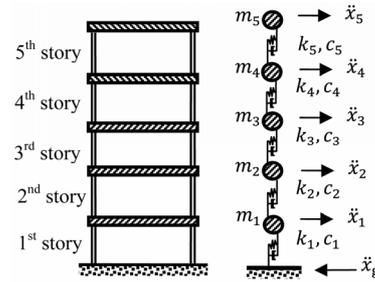


Figure 2. 5-DOF shear frame structure.

Table 1. Structural and modal parameters of the structure.

DOF	1	2	3	4	5
Mass (kg)	1000	1000	1000	1000	1000
Stiffness (MN/m)	2.000	2.000	2.000	2.000	2.000
Frequency (Hz)	2.03	5.91	9.32	11.98	13.66

where M , C , and K are respectively the mass matrix, damping matrix, and stiffness matrix. $f(t)$ is the force vector linked to the ground acceleration. \ddot{X} , \dot{X} , and X are respectively relative acceleration, velocity, and displacement response. The sampling frequency was 100 Hz. In the simulation, the input signal was Gaussian white noise (**Figure 3**). To test the noise immunity of our method, noise at levels of 5% and 10% was added to the raw acceleration data.

4.2. Results of Damage Detection

The training data sets that were to generate detectors were acceleration time histories from the top story of the healthy shear structure under ground motion following a pattern of randomly generated Gaussian white noise. In the test phase, the test data sets (self and non-self) were obtained from the top story of the healthy and damaged shear structure under ground motion following another pattern of randomly generated Gaussian white noise. The sampling frequency was 100 Hz, and all time histories were normalized.

As is shown in [18], the word length and window length will greatly affect the performance of the methodology. Larger word and window lengths yield better performance. The reason is that, a longer word or window can symbolize the raw acceleration data much more accurately than a shorter one. As the word length and window length increase, much more dynamic information about the system is captured, and the representation space of the problem becomes more accurate. In the simulation, the word length was set as 9 and window length was set as 3.0E+03.

Several tests were performed by simulating stiffness reductions occurring in different locations and to different degrees. The simulated cases included stiffness re-

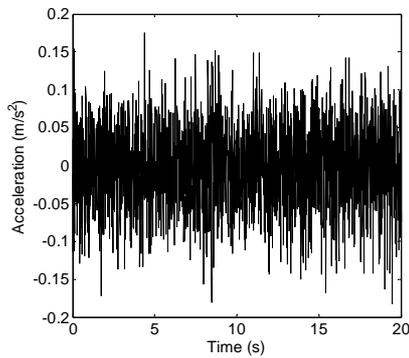


Figure 3. Example input signal (Gaussian white noise).

ductions on a single story (2nd story or 4th story) and on two stories (2nd story and 4th story). The degree of stiffness reduction was 10% or 20%. Table 2 lists the most pertinent results, and Figure 4 compares the DR and FAR values of the different damage cases.

The results show that, by employing our methodology, perfect results can be obtained for the noise-free case, DR is 100% and FAR is 0%, which means all the damage can be detected and no SSH from the healthy structure is misclassified. For the noise polluted cases as well, a high DR and low FAR can be obtained no matter the location or degree of stiffness reduction. Even as the noise level increases and the detection rate decreases, the extent of misclassification remains very small. This means that noise is not problematic in our methodology. Indeed, the good results probably reflect the fact that the RSSHe is much more sensitive to changes in the structure itself than to the environment.

5. Numerical Simulation of Damage Localization and Quantification

5.1. Description of Procedure

In this simulation, the system to be identified was the same five-story shear frame structure as stated before; the mass distribution and damping parameters were assumed to be known, and the stiffness of each story was set as the objective parameters that needed to be identified. First, the parameters of the healthy structure were identified as a reference. AICSA was performed when the abnormal SSH of the damaged structure was detected. The abnormal SSH was the output of the damaged structure. The input ground acceleration data corresponding to the abnormal SSH was used as the input of the candidate model, and the output of the candidate model was symbolized using the procedure in Section 2.2. The Euclidean distance of system and candidate SSHs was used as the objective function to be minimized. After the parameters of the damaged structure were identified, the severity (Equation (6)) showing the location and degree of the

Table 2. DR and FAR of different damage cases.

Noise level (%)	Single				Double	
	Location & Degree	Rate (%)		Location & Degree	Rate (%)	
		DR	FAR		DR	DR
0		100	0		100	0
5	2nd & 10%	98.64	0.20	2nd & 10% 4th & 10%	98.87	0.19
10		97.76	0.41		98.15	0.24
0		100	0		100	0
5	2nd & 20%	99.06	0.13	2nd & 20% 4th & 10%	99.28	0.14
10		98.91	0.17		98.69	0.19
0		100	0		100	0
5	4th & 10%	98.09	0.23	2nd & 10% 4th & 20%	99.04	0.13
10		97.56	0.44		98.26	0.21
0		100	0		100	0
5	4th & 20%	98.96	0.16	2nd & 20% 4th & 20%	99.91	0.09
10		98.21	0.19		99.54	0.11

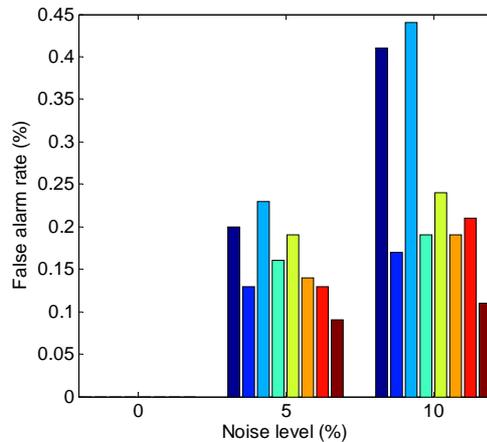
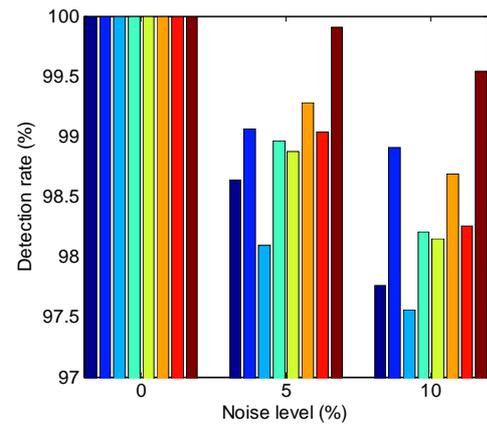
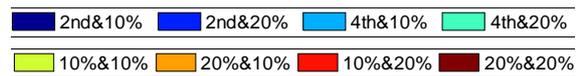


Figure 4. DR and FAR for damage detection.

damage was calculated.

5.2. Identification Results for Healthy and Damaged Structures

The input signal was Gaussian white noise, as in the previous simulation. The window and word lengths were the same as before, and the full output of the structure was used. **Table 3** lists the identified stiffness of each story of the healthy and damaged structures. The damage cases are the same as those in the damage detection stage. Severity of damage was calculated for each damage case using the results in **Table 3**, and **Figures 5, 6** and **7** plot the severity for the noise-free, 5% noise and 10% noise cases, respectively.

Table 3. Estimated parameters for healthy and damaged models with a word length of 9 and window length of 3000.

Case State	Stiffness	True Value (MN/m)	Calculated Value(MN/m)		
			No noise	5% noise	10% noise
0 Healthy	k1	2.000	2.000	2.001	2.000
	k2	2.000	2.000	1.999	2.003
	k3	2.000	2.000	2.021	1.987
	k4	2.000	2.000	1.997	2.031
	k5	2.000	2.000	2.000	1.993
1 One damage 2nd & 10%	k1	2.000	2.000	2.003	2.010
	k2	1.800	1.800	1.796	1.805
	k3	2.000	2.000	2.014	1.996
	k4	2.000	2.000	2.000	1.978
	k5	2.000	2.000	1.989	1.984
2 One damage 4th & 20%	k1	2.000	2.000	2.008	2.030
	k2	2.000	2.000	2.017	2.017
	k3	2.000	2.000	2.010	1.997
	k4	1.600	1.600	1.598	1.601
	k5	2.000	2.000	1.998	2.024
3 Two damage 2nd & 10% 4th & 20%	k1	2.000	2.000	1.994	2.000
	k2	1.800	1.800	1.796	1.815
	k3	2.000	2.000	2.019	1.987
	k4	1.600	1.600	1.608	1.582
	k5	2.000	2.000	2.001	1.993
4 Two damage 2nd & 20% 4th & 20%	k1	2.000	2.000	1.992	1.979
	k2	1.600	1.600	1.598	1.624
	k3	2.000	2.000	2.031	2.024
	k4	1.600	1.600	1.619	1.631
	k5	2.000	2.000	2.006	1.989

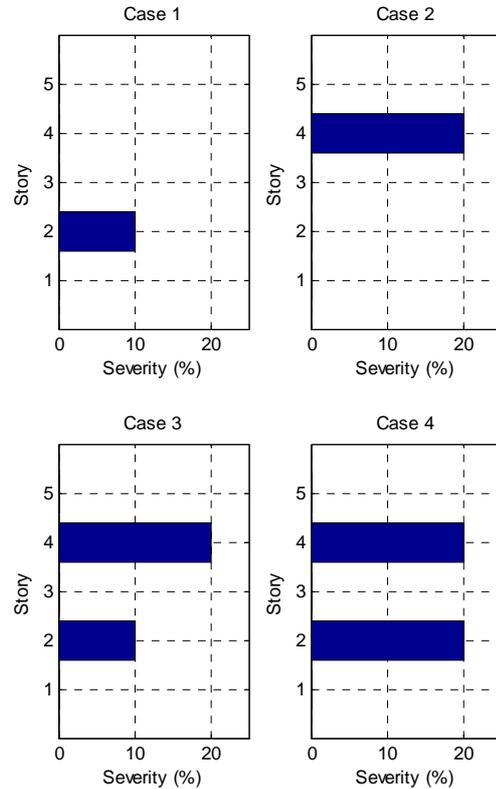


Figure 5. Identification results for noise-free case.

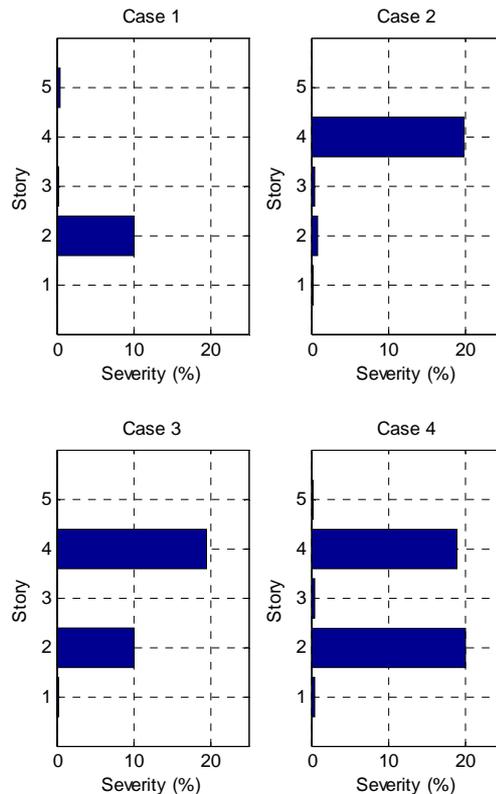


Figure 6. Identification results for 5% noise case.

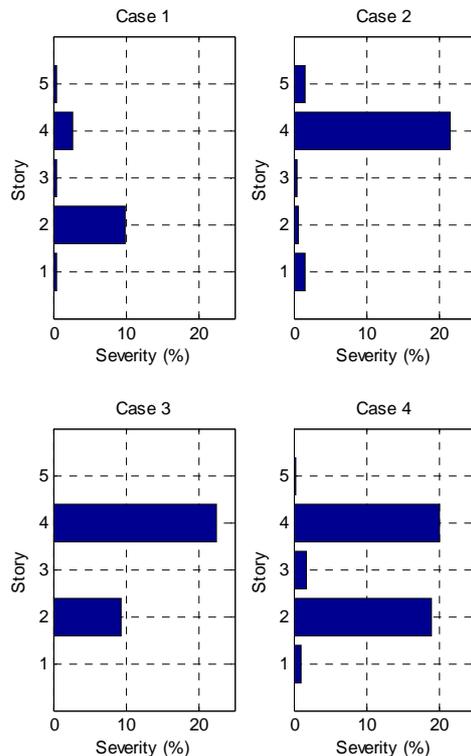


Figure 7. Identification results for 10% noise case.

The results show that, AICSA combining STSA can identify the parameters of a structure accurately regardless of whether the structure is healthy or damaged. Even error of severities of each story increase slightly as the noise level of the raw acceleration increase, the location and severity of the damage can be identified distinctly.

6. Conclusion

We proposed a hybrid methodology based on immune algorithms (IAs) and symbolic time series analysis (STSA) for structural health monitoring (SHM). In this methodology, RNS is used detect damage, and AICSA is used to localize and quantify the damage to the structure. Data symbolization by using STSA alleviates the effects of noise in the raw acceleration data. We explained the mathematical basis of STSA and described a simulation experiment on a five-story shear frame structure. The results showed that this hybrid methodology can efficiently and precisely detect, localize and quantify damage to civil engineering structures in the presence of measurement noise.

7. Acknowledgements

This work was supported in part by a Grant-in-Aid No. 22310103 (PI: A. Mita) and Grant-in-Aid of the Global Center of Excellence Program for the “Center for Edu-

cation and Research of Symbiotic, Safe and Secure System Design” from the Ministry of Education, Culture, Sport, Science and Technology of Japan.

REFERENCES

- [1] R. I. Levin and N. A. J. Lieven, “Dynamic Finite Element Model Updating Using Simulated Annealing and Genetic Algorithms,” *Mechanical System and Signal Process*, Vol. 12, No. 1, 1998, pp. 91-120. [doi:10.1006/mssp.1996.0136](https://doi.org/10.1006/mssp.1996.0136)
- [2] K. Cunha and V. V. Smith, “A Determination of the Solar Photospheric Boron Abundance,” *Astrophysical Journal*, Vol. 512, No. 2, 1999, pp. 1006-1013. [doi:10.1086/306796](https://doi.org/10.1086/306796)
- [3] S. T. Xue, H. S. Tang and J. Zhou, “Identification of Structural Systems Using Particle Swarm Optimization,” *Journal of Asian Architecture and Building Engineering*, Vol. 8, No. 2, 2009, pp. 517-524. [doi:10.3130/jaabe.8.517](https://doi.org/10.3130/jaabe.8.517)
- [4] R. Li and A. Mita, “Structural Damage Identification Using Adaptive Immune Clonal Selection Algorithm and Acceleration Data,” *SPIE Smart Structures/NDE2011*, San Diego, 6-10 March, 2011, 79815A.
- [5] R. Li and A. Mita, “Hybrid Immune Algorithm for Structural Health Monitoring Using Acceleration Data,” *8th International Workshop on Structural Health Monitoring*, Stanford University, 13-15 September 2011, pp. 1095-1102.
- [6] A. Ray, “Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection,” *Signal Processing*, Vol. 84, No. 7, 2004, pp. 1115-1130. [doi:10.1016/j.sigpro.2004.03.011](https://doi.org/10.1016/j.sigpro.2004.03.011)
- [7] S. Chin, A. Ray and V. Rajagopalan, “Symbolic Time Series Analysis for Anomaly Detection: A Comparative Evaluation,” *Signal Processing*, Vol. 85, No. 9, 2005, pp. 1859-1868. [doi:10.1016/j.sigpro.2005.03.014](https://doi.org/10.1016/j.sigpro.2005.03.014)
- [8] A. Khatkhat, A. Ray, S. Chin, V. Rajagopalan and E. Keller, “Detection of Fatigue Crack Anomaly: A Symbolic Dynamic Approach,” *Proceedings of American Control Conference*, Boston, 30 June-2 July 2004; pp. 3741-3746.
- [9] D. Tolani, M. Yasar, A. Ray and V. Yang, “Anomaly Detection in Aircraft Gas Turbine Engines,” *AIAA Journal of Aerospace Computing Information, and Communication*, Vol. 3, No. 2, 2006, pp. 44-51. [doi:10.2514/1.15768](https://doi.org/10.2514/1.15768)
- [10] S. Bhatnagar, V. Rajagopalan and A. Ray, “Incipient Fault Detection in Mechanical Power Transmission Systems,” *Proceedings of American Control Conference*, Portland, 8-10 June 2005, pp. 472-477.
- [11] S. Forrest, A. Perelson, L. Allen and R. Cherukuri, “Self-Nonself Discrimination in a Computer,” *Proceedings of IEEE Symposium on Research in Security and Privacy*, Los Alamitos, 16-18 May 1994, pp. 202-212.
- [12] F. Gonzalez, D. Dasgupta and R. Kozma, “Combining Negative Selection and Classification Techniques for Anomaly Detection,” *Proceedings of the 2002 Congress*

- on *Evolutionary Computation CEC2002*, 12-17 May Honolulu, pp. 705-710.
- [13] F. A. Gonzalez and D. Dasgupta, "Anomaly Detection Using Real-Valued Negative Selection," *Genetic Programming and Evolvable Machines*, Vol. 4, No. 4, 2003, pp. 383-403.
- [14] X. Wang, X. Z. Gao and S. J. Ovaska, "Artificial Immune Optimization Methods and Applications: A Survey," *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, The Hague, 10-13 October 2004, pp. 3415-3420.
- [15] J. Timmis, P. Andrews, N. Owens and E. Clark, "An Interdisciplinary Perspective on Artificial Immune Systems," *Evolutionary Intelligence*, Vol. 1, No. 1, 2008. pp. 5-26. [doi:10.1007/s12065-007-0004-2](https://doi.org/10.1007/s12065-007-0004-2)
- [16] L. Zhang, X. R. Meng, W. J. Wu and H. Zhou, "Network Fault Feature Selection Based on Adaptive Immune Clonal Selection Algorithm," 2009 *International Joint Conference on Computational Sciences and Optimization*, Sanya, 24-26 April 2009, pp. 969-973. [doi:10.1109/CSO.2009.342](https://doi.org/10.1109/CSO.2009.342)
- [17] A. Mita, "Structural Dynamics for Health Monitoring," Sankeisha Co., Ltd, Nagoya, 2003.
- [18] R. S. Li, A. Mita and J. Zhou, "Feasibility Study of Parameter Identification Method Based on Symbolic Time Series Analysis and Adaptive Immune Clonal Selection Algorithm," *Open Journal of Civil Engineering*, Vol. 2 No. 4, 2012, pp. 198-205.