



THE HONG KONG
POLYTECHNIC UNIVERSITY

香港理工大學

Pao Yue-kong Library

包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**FAULT DIAGNOSIS OF SENSOR
NETWORKED COMPLEX
STRUCTURES: A VIRTUAL BEAM-
LIKE STRUCTURE APPROACH**

WANG HONG

Ph.D

The Hong Kong Polytechnic University

2017



The Hong Kong Polytechnic University
Department of Mechanical Engineering

**Fault Diagnosis of Sensor
Networked Complex Structures: A
Virtual Beam-Like Structure
Approach**

Wang Hong

A thesis submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September, 2016

Certificate of Originality

I hereby declare that this thesis entitled “*Fault Diagnosis of Sensor Networked Complex Structures: A Virtual Beam-Like Structure Approach*” is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

(Signed)

WANG HONG

(Name of student)

Abstract

Fault detection and localization techniques are among the prominent technical processes involved in ensuring reliability. This is particularly true with regard to safety related processes or systems such as trains, power plants, and aircrafts which may also be characterized by limited datasets compiled through real measurements (e.g. pre-launching testing). Sensor-network based fault diagnosis in complex structures with limited prior knowledge is an interesting but difficult topic in structure health monitoring.

In this study, a novel method for the fault diagnosis of sensor networked complex structures based on a novel optimized virtual beam-like structure (VBLS) approach is developed to fault diagnosis of the pre-launching systems. The statistical methods together with adaptive threshold technique are presented for fault detection to overcome the challenge of obtaining an appropriate threshold value with only limited experimental data. A complex structure is regarded as a combination of numerous virtual beam-like structures considering the vibration transmission paths from vibration sources to each sensor. The “virtual beam” consisting of a sensor chain automatically represents a candidate vibration transmission path, which can be obtained automatically by an improved optimization algorithm. The dynamic response of the structures in this vibration transmission path can demonstrate obvious fault features if there is a fault (e.g., cracks in connecting rods or around

bolts, and bolt-loosening etc.). These fault features can be effectively characterized and efficiently captured and utilized for fault localization using the optimization method based on the virtual beam-like structure concept. This novel virtual beam-like-structure approach is applicable to fault diagnosis of complex structures without too much prior knowledge of the faults, and with only limited prior knowledge of normal operational conditions. Neither does it require stationary response data, nor is it confined to a specific structure design. It is easy to implement within a sensor network attached to the monitored structure.

The effectiveness of the proposed virtual beam-like structure approach is well validated in experiments for both single-fault and multi-fault diagnosis including loosening bolts or cracks around bolts of complex structures such as bolted-base hanging structures in satellites. Compared with the classical subtract on negative add on positive (SNAP) localization methods, the proposed virtual beam-like structure (VBLS) approach has proved to be more accurate for fault localization. The computational complexity of the virtual beam-like structure approach has been evaluated, which indicates that the proposed method provides a promising solution to on-line fault diagnosis of the sensor networked complex structures with limited prior knowledge.

Publications

1. **Hong Wang**, XingJian Jing: A Sensor Network based Virtual Beam-like Structure Method for Fault Diagnosis and Monitoring of Complex Structures with Improved Bacterial Optimization. *Mechanical Systems and Signal Processing*, vol. 84, part A, pp. 15-38, 2017.
2. **Hong Wang**, XingJian Jing: Vibration signal–based fault diagnosis in complex structures: A beam-like structure approach. *Structural Health Monitoring*, 2017.
<http://journals.sagepub.com/doi/full/10.1177/1475921717704383>
3. **Hong Wang**, XingJian Jing: Fault Diagnosis of the Sensor Networked Structure with Multiple Faults using Virtual Beam-like Structure Method. *Journal of Sound and Vibration*, 2017. <http://doi.org/10.1016/j.jsv.2017.03.020>
4. **Hong Wang**, XingJian Jing: A Virtual Beam based Event-oriented Algorithm for Multiple Fault Localization in Complex Structures with an Optimal Sensor Network. *Nonlinear Dynamics*, 2017. (Submitted)
5. **Hong Wang**, Xingjian Jing: A Virtual Beam-Like Structure Approach for Fault Diagnosis of Complex Structures. *Proceedings of the 35th Chinese Control Conference (CCC 2016)*. Chendu, China. 26-30 July, 2016. Doi: [10.1109/ChiCC.2016.7554365](http://doi.org/10.1109/ChiCC.2016.7554365)
6. **Hong Wang**, Xingjian Jing, Ben Niu: Bacterial-Inspired Feature Selection Algorithm and Its Application in Fault Diagnosis of Complex Structures. *Proceedings of 2016 IEEE World Congress on Computational Intelligence/IEEE Congress on Evolutionary Computation (WCCI 2016 /CEC2016)*. Vancouver, Canada. 24-29 July, pp. 3809-3816, 2016.
7. **Hong Wang**, Xingjing Jing, Ben Niu: A Weighted Bacterial Colony Optimization for Feature Selection. *Proceedings of 2014 International Conference on Intelligent Computing (ICIC2014)*. Taiyuan, China, 3-6 August, 2014, vol. 3, pp. 379-389

Acknowledgements

I wish to express my deep and sincere gratitude to my Chief Supervisor Dr. Jing Xingjian whose guidance, encouragement, suggestions and constructive criticism have contributed immensely to the inspiration of my ideas on the research studies.

I am very much thankful to the NSSP (Name of our research group) members, especially to Mr. Li Quankun and Mr. Sun Bo, for their valuable assistances in experimental works, and providing me with useful suggestions to improve the quality of related project work. I also thank to my friends in office and roommates in Hong Kong Polytechnic University for the precious friendship and rendering me with enjoyable life in Hong Kong.

Above all, I would like to express my special gratitude to my parents whose kindness, understanding and countless love assist me at various stages of my life.

Many thanks to my friends in Shen Zhen who have helped me a lot in getting through the blue days and throwing into the research studies.

Finally, I gratefully acknowledge the support from the CAST-funded project, a NSFC project of China.

Contents

Certificate of Originality	I
Abstract.....	II
Publications.....	IV
Acknowledgements	V
Contents	VI
List of Figures	XI
List of Tables	XV
Nomenclature	XVI
Chapter 1. Introduction	1
1.1. Background.....	1
1.2. Objectives and Methodology	5
1.3. Outline of the Thesis.....	8
Chapter 2. Literature Review	12
2.1. Typical Features for Pattern Recognition.....	12
2.2.1. Features in Statistical Techniques	13
2.2.2. Features in Structural Techniques	14
2.2.3. Features in Hybrid Techniques.....	15
2.2. Feature Extraction Methods.....	16
2.2.1. Time Domain Features based Techniques	17

2.2.2.	Frequency Domain Features based Techniques	18
2.2.3.	Time-Frequency Domain Features based Techniques	19
2.3.	Feature Selection Methods	22
2.3.1.	Traditional Feature Selection Algorithms	22
2.3.2.	Evolutionary Feature Selection Algorithms	25
2.3.3.	Bacterial based Feature Selection Algorithms	27
2.4.	Fault Diagnosis Techniques using Limited Prior Knowledge.....	29
2.4.1.	Model-based Fault Diagnosis Methods	29
2.4.2.	Model-free based Fault Diagnosis Techniques	32
2.5.	Conclusion.....	36
Chapter 3.	Bacterial Optimization Algorithm for Feature Selection.....	38
3.1.	Definition of Feature Selection	38
3.2.	Bacterial Algorithms	39
3.2.1.	Typical Bacterial Algorithms	39
3.2.2.	Bacterial Colony Optimization	40
3.3.	The Proposed Bacterial Algorithm based Feature Selection (BAFS)	43
3.3.1.	Weighting Mechanism and ‘Archive’ Strategy	45
3.3.2.	Discriminatory Variables Selection	47
3.3.3.	Randomness Control Mechanisms.....	47
3.3.4.	Modified Reproduction and Elimination Strategies.....	49
3.3.5.	Bacterial Algorithm based Feature Selection Method	51

3.4.	Parameters Estimation and Decision	53
3.4.1.	Parameter Settings and Benchmarks	53
3.4.2.	Experimental Results on Parameters P_{te} , P_{re} , and P_{el}	54
3.5.	Benchmark Experiments	59
3.5.1.	Comparison Techniques and Datasets	59
3.5.2.	Experimental Results and Analysis	61
3.6.	Discussion and Conclusion	64
Chapter 4.	Feature Characterization and Fault Detection.....	65
4.1.	Signal Preprocessing.....	65
4.1.1.	Data Acquisition.....	65
4.1.2.	Feature Characterization	67
4.1.3.	Selection of Features for Signal Analysis	68
4.2.	Feature Based Fault Indicators	76
4.3.	Fault Detection	78
4.3.1.	Fault Indicator based Adaptive Threshold.....	78
4.3.2.	Statistical Methods.....	79
4.3.3.	Combined Method for Fault Detection	83
4.4.	Conclusion	84
Chapter 5.	Virtual Beam-like Structure Approach.....	85
5.1.	Concept of Virtual Beam-like Structure.....	85
5.2.	Sensor Networks Distribution and Selection	87

5.3.	Optimization Methodology for Virtual Beams	91
5.3.1.	Objective Function and Methodology	91
5.3.2.	BAFS for Optimal Selection of Sensors	94
5.3.3.	“Biased Running” based Optimization Methodology	95
5.4.	Validation of the “Virtual” Beam Method in a “Real” Beam-like Structure	99
5.4.1.	The Beam-Like Structure	99
5.4.2.	Benchmark Testing on the Beam-Like Structure	100
5.5.	Fault Localization based on Virtual Beams	102
5.5.1.	Fault Localization based on Single Virtual Beam	103
5.5.2.	Fault Localization based on Multiple Virtual Beams	105
5.6.	Conclusion.....	106
Chapter 6.	Fault Diagnosis of Complex Structures with Single Fault	108
6.1.	Experimental Platform	108
6.2.	Experimental Testing and Results	110
6.2.1.	Fault Diagnosis on Band Antenna Structure	112
6.2.2.	Fault Diagnosis on Solar Panel Structure	115
6.2.3.	Fault Diagnosis on Body Structure	119
6.3.	Discussion and Conclusion	125
Chapter 7.	Fault Diagnosis of Complex Structures with Multiple Faults	127
7.1.	Problem Formulation.....	127

7.2.	Experimental Testing and Results.....	130
7.3.1.	Experimental Platform	131
7.3.2.	Multi-Fault Diagnosis on Solar Panel	132
7.3.3.	Multi-Fault Diagnosis on Main Body	138
7.3.4.	Multi-Fault Diagnosis on both Solar Panel and Main Body	146
7.3.	Discussion.....	150
7.4.	Conclusion	153
Chapter 8.	Procedures and Computational Complexity Evaluation for Fault Diagnosis System.....	154
8.1.	Procedures for Fault Diagnosis	155
8.2.	Evaluation of Computational Complexity	162
8.3.	Application of the VBLS Approach.....	164
8.4.	Advantages of the Proposed Fault Diagnosis Method	165
Chapter 9.	Conclusion and Prospect.....	169
9.1.	Conclusion	169
9.2.	Prospect	171
References.	173

List of Figures

Figure 1-1 The overall framework of the proposed methodology	6
Figure 1-2 The overall structure of the thesis	9
Figure 3-1 The overall framework of BFO	39
Figure 3-2 The main principle of BCO	42
Figure 3-3 The overall framework of the Bacterial Algorithm based Feature Selection	44
Figure 3-4 Average classification accuracy achieved with varying P_{te} , P_{re} , and P_{el}	55
Figure 3-5 Average classification accuracy across 30 repeated runs for varying subset size on datasets	62
Figure 3-6 Computational time spent by feature selection methods across the 30 runs	63
Figure 4-1 The designed beam-like structure.....	66
Figure 4-2 Accuracy rate achieved by clustering method using the data provided by three sensors with individual time feature as signal representative.....	75
Figure 4-3 Classification accuracy achieved by three sensors (i.e. A1, A2 and A3)	75
Figure 4-4 The flowwork of the combined fault detection method	83
Figure 5-1 Two typical examples of vibration transmission path along the structures	86
Figure 5-2 Typical bolted-base hanging structures in a satellite-like model.....	89
Figure 5-3 Sensor networks on the satellite-like structure.....	90
Figure 5-4 Subnetwork separation based on region of influence	96
Figure 5-5 The mean and variance IDev of three sensors in different conditions	100
Figure 5-7 Typical examples of fault isolation according to the Re_Dev of the ‘virtual beam’	104

Figure 5-8 Virtual beams for fault localization based on time domain features..	106
Figure 6-1 The overall framework of virtual beam-like structure approach (VBLS) for single-fault diagnosis	108
Figure 6-2 Satellite-like structure and testing platform	109
Figure 6-3 Sensor networked structure of band antenna	112
Figure 6-4 Comparison between the normal system and fault system according to IDev of sensors.....	113
Figure 6-5 The changes of Re_Dev along the transmission path when the signals are from the diagnosis system (*sensors indicating the faulty system are highlighted in square shaped mark)	114
Figure 6-6 Sensor networked structure of solar panel	115
Figure 6-7 VBLS method and SNAP method for fault isolation when the diagnosis system is suffering from a loosening screw L2.	116
Figure 6-8 VBLS method and SNAP method for fault isolation when the diagnosis system is suffering from loosening screw of L4.	118
Figure 6-9 Sensors are networked on main body of satellite-like structure for virtual beam construction. *Four regions (i.e. P1, P2, P3 and P4) are frequently suffering from the loosening screw fault.	119
Figure 6-10 VBLS method and SNAP method for isolation of P1 fault.	122
Figure 6-11 VBLS method and SNAP method for isolation of P2 fault.	123
Figure 6-12 VBLS method and SNAP method for isolation of P3 fault.	124
Figure 6-13 VBLS method and SNAP method for isolation of P4 fault.	125
Figure 7-1 Examples of multiple faults on satellite-like structure	128
Figure 7-2 Correlation of time and frequency under sweeping-frequency excitation	129
Figure 7-3 The overall framework of virtual beam-like structure approach (VBLS) for multi-fault diagnosis	130
Figure 7-4 Response of sensors from the solar panel and virtual beams constructed by optimization method when diagnosis signals are measured from system suffering	

from loosening bolt at L1.....	133
Figure 7-5 Fault localization using TI-SNAP when diagnosis signals are measured from system suffering from loosening bolt at L1	134
Figure 7-6 Response of sensors on solar panel when it is suffering from loosening bolts at L1&L2	135
Figure 7-7 Virtual beams constructed using optimization method when the system is suffering from loosening bolts at L1&L2.....	136
Figure 7-8 Fault localization using TI-SNAP when diagnosis signals are measured from system suffering from loosening bolt at L1&L2	136
Figure 7-9 Response of sensors on solar panel when it is suffering from loosening bolts at L1&L3	137
Figure 7-10 Virtual beams constructed using optimization method when the system is suffering from loosening bolts at L1&L3	137
Figure 7-11 Fault localization using TI-SNAP when diagnosis signals are measured from system suffering from loosening bolt at L1&L3	138
Figure 7-12 Distribution of sensor networks when main body is studied for health condition	139
Figure 7-13 Alarm sensors among the sensor networks on solar panel when it is suffering from loosening bolt at connector P1	141
Figure 7-14 Virtual beams constructed using features when the system is suffering from loosening bolt at connector P1	141
Figure 7-15 Fault localization using TI-SNAP when the system is suffering from multiple loosening bolts at connectors P1.....	142
Figure 7-16 Response of sensors and virtual beams constructed using optimization method when the system is suffering from multiple loosening bolts at connectors P1&P2.....	143
Figure 7-17 Fault localization using TI-SNAP when the system is suffering from multiple loosening bolts at connectors P1&P2.....	144
Figure 7-18 Response of sensors and virtual beams constructed by optimization	

method when the system is suffering from multiple loosening bolts at connectors P1&P3.....	145
Figure 7-19 Fault localization using TI-SNAP when the system is suffering from multiple loosening bolts at connectors P1&P3	145
Figure 7-20 Response of sensors and virtual beams constructed using the sensors located on the main body and adjacent connectors	147
Figure 7-21 Fault localization using TI-SNAP when the system is suffering from L1&P3 faults.....	147
Figure 7-22 Response of sensors and virtual beams constructed using the sensors located on the solar panel	148
Figure 7-23 Fault localization using TI-SNAP when the system is suffering from L1&P3 faults.....	149
Figure 8-1 Procedures of the proposed fault diagnosis system	154
Figure 8-2 The fault diagnosis system.....	159
Figure 8-3 Examples of fault diagnosis using the proposed fault diagnosis system	161

List of Tables

Table 3-1 The range of classification accuracy rate and iterations using varying values of termination parameter P_{te}	54
Table 3-2 The range of classification accuracy rate and iterations using varying values of termination parameter P_{re}	56
Table 3-3 Average classification accuracy rate and iterations for the optimal when $P_{te}=40$	57
Table 3-4 Average classification accuracy rate and iterations for the optimal when $P_{te}=80$	57
Table 3-5 Average classification accuracy rate and iterations for the optimal when $P_{te}=120$	58
Table 3-6 Datasets for feature selection.....	60
Table 4-1 Typical time domain statistical features	68
Table 5-1 The results of fault indicators based on features	102
Table 6-1 Deviation ratio (IDev) and relative deviation ratio (Re_Dev) based on feature representatives from sensor chain	114
Table 6-2 Results of feature-represented sensors based on optimal virtual beams (L2 and L4 fault).....	117
Table 6-3 Results of feature-represented sensors based on optimal virtual beams (P1-P4 fault).....	120
Table 8-1 Time complexity of the proposed fault diagnosis system.....	163

Nomenclature

Abbreviations

VBLS	Virtual Beam-Like Structure
FS	Feature Selection
BAFS	Bacterial Algorithm based Feature Selection
BFO	Bacterial Foraging Optimization
BAs	Bacterial Algorithms
BFOLDC	BFO with Linear Chemotaxis Step Size
BFONDC	BFO with Nonlinear Chemotaxis Step Size
BC	Bacterial Chemotaxis
BCO	Bacterial Colony Optimization
PSO	Particle Swarm Optimization
EC	Evolutionary Computation
ACO	Ant Colony Optimization
GA	Genetic Algorithm
Dev	Deviation Ratio
IDev	Improved Deviation Ratio
Re_Dev	Relative Deviation Ratio
SNAP	Subtract on Negative Add on Positive
TI-SNAP	Trust Index Subtract on Negative Add on Positive
DSNAP	Decentralized Subtract on Negative Add on Positive
FTML	Fault Tolerant Maximum Likelihood
CE	Centroid Estimator
ML	Maximum Likelihood
BBHSs	Bolted-Base Hanging Structures
GA	Genetic Algorithm
ED	Euclidean Distance

K-S	Kolmogorov-Smirnov test
CDF	Cumulative Density Function
LCM	Life Cycle Model
RMS	Root Mean Square
FFT	Fast Fourier Transformation
FTs	Fourier Transforms
K-S	Kolmogorov-Smirnov
ANN	Artificial Neural Network
SVM	Support Vector Machine
STFT	Short-time Fourier Transform
WVD	Wigner-Ville distribution
WTs	Wavelet Transforms
DWT	Discrete Wavelet Transform
CWT	Continuous Wavelet Transform
WPT	Wavelet Package Transform
EMD	Empirical Mode Decomposition
HHT	Hilbert-Huang Transform
TID	Translation-Invariant Denoising
PCA	Principle Component Analysis

Symbols

A_s	Sensor
$X_{A_s}^{n_i}$	Time domain features measured from healthy system
$X_{A_s}^{diag}$	Time domain features measured from diagnosis system
D	Number of sensors for virtual beam
max_{row}	Maximum number of sensors to be selection in a row
Err_{num}	Number of sensors reflecting the abnormal system
$ub_{A_s}^n$	Upper boundary of the fault indicator value

lb	lower boundary of the relative fault indicator value
α_{A_s}	Selection of a sensor A_s by virtual beam
$Dis(A_s, A_{s+1})$	Distance between two sensors
β_s	Weight vector for a relative distance
R_c	Region of coverage
R_I	Region of influence

Chapter 3

P_{te}	Parameter used for premature termination
P_{re}	Parameter used to control the frequency of reproduction
P_{el}	Parameter used to control the frequency of elimination
$P_i(T)$	Position of the i^{th} bacterium during T^{th} iteration
P_{best}	Best position of an individual, recorded as local best
G_{best}	Best position of the population, recorded as global best
$\Delta(i)$	Direction angle of the i^{th} bacterium
$C(i)$	Chemotaxis step size
$R = [R_1, \dots, R_{NP}]$	A randomly generated vector consisting of 0 and 1
NP	Population size
$J_i(health)$	Health condition of the bacterium
$J(i, j, k, l)$	The fitness value
$\theta^i(j, k, l)$	The position of i^{th} bacterium, at the j^{th} chemotaxis, k^{th} reproduction, l^{th} dispersal
x_{max}, x_{min}	The maximum and the minimum of the variables
W	Weight matrix to record the performance
A	Matrix to record the occurrence of variables
Avg_Fit	The average fitness of the population

Chapter 4

$u_{ij} \in \{0,1\}$	Membership levels
----------------------	-------------------

W_{inert}	Inertia weight
c_1 and c_2	Acceleration constants to control the step distance of particles
M	Number of signals from the healthy system
T	Threshold
$\bar{X1}$	The mean of the sample
$P(x)$	The probability of the distribution
$F_1(x)$	The distribution of a sample
H_0	The null hypothesis

Chapter 5

$P_1, P_2, P_3,$ and P_4	The connectors
$[A_1, A_2, \dots A_D]$	The sensor chain
$Max_iteration$	The maximum iteration times
C	The number of columns in a sensor network
R	The number of rows in a sensor network
D	The dimensionality to decide the length of sensor chain

Chapter 1. Introduction

1.1. Background

In the field of mechanical engineering, fault diagnosis plays an essential role in monitoring the operational condition of pre-launching machines/structures and locating the faults in machines/structures to keep the operational sustainability and safety. The information used for detecting the health condition of mechanical systems includes those signals related to position, pressure, temperature, current, voltage and so on. Some of them can be directly obtained from records, while for most cases, data cannot be that easy to acquire, but can be measured in indirect ways by vibration sensors [1]. Those data, originally, are time series of measurement signals which represent the result of the observation on physical response phenomena and are sensitive to small changes (temporal variability) of the structure systems.

Generally, the techniques for fault diagnosis can be simply divided as model-based fault diagnosis and model-free fault diagnosis [2]. Model-based approaches would be more accurate when the structure is stationary and the physical information is easy to be described by models, but it would may confront the dilemma when the physical parameters as well as the fault modes are changing under different operation conditions. In time-varying system dynamics, it is not appropriate to apply the model-based approaches for fault diagnosis since the

system is rather complex with internal state variables which are inaccessible to be measured by common sensors [3]. Instead, model-free approaches focus on the study of the residuals rather than the construction of the observed and estimated models, which are much more flexible and applicable for monitoring the health condition of the complex dynamic systems. Therefore, the need for fault diagnosis in complex dynamic structure systems has led to the research impetus on time series data analysis.

The studies of time series data for fault diagnosis begin with extracting the informative features to represent the time series signals. Feature means an essential distinguishing characteristic, representing the physical properties of uniqueness and cardinality [4], or describing the attributes of the object that are most representative [5]. In feature-based analysis, features might be expressed as shape features [6, 7], texture features [8-10], power spectral density features [11], intensity features [12], structural features [13], etc. Taking fault detection in gearboxes as an example, 34 statistical features in function forms have been shown in [14] including the commonly used features like root mean square (RMS), kurtosis, skewness, standard deviation, variance, average absolute value, peak to peak, etc. Even more comprehensive feature description in such a research area can be seen in [15], in which 213 features have been identified as the potential for fault detection in planetary gearboxes. There are numerous types of feature definitions, while the challenge is the selection of features for fault detection since the features that are

effective in one system may not be valid in other systems.

Diagnosis techniques for machinery/structural health monitoring based on vibration signals are generally divided into three groups: time domain features, frequency domain features, and time-frequency features [16]. Being simple and interpretable, time domain features have been frequently used for health monitoring in mechanical engineering [17-21], which are calculated from raw signals without any transformation such as root mean square (RMS), kurtosis. The main advantage of this group of features is that time domain features are easy and quick in preprocessing without the tasks like filtering, windowing, framing, Fourier transformation, etc. [22]. The drawback of these group of features may lie in that it could be less robust in high noise environment. Frequency domain features like Fast Fourier Transformation (FFT), are transformed from the raw signals in the frequency domain [23, 24]. The advantage of frequency domain features is that it may be easier to isolate the faults at certain frequency components [16, 25]. However, Fourier related Transforms (FTs) have to rely on the fixed-time window. Though they are very useful tool for stationary cases, it is not suitable for non-stationary situations [16]. For that, time frequency analysis techniques like wavelets have been developed for dealing with the non-stationary signals [26]. Even so, the expensive cost to handle these group features is that they usually need much capital expense and computational complexity in signal processing [25], especially in dealing with high dimensional time series data. In such circumstances, time domain

features are considered to extract the characteristic features of time series using certain segments in my study.

Another challenge is that only a few approaches are applicable if available prior knowledge of the system is limited with regard to both the normal operational conditions and the abnormal conditions. Concerning the unavailability of prior knowledge of faults, it is often for signal-based methods to detect the health condition of the system by matching the fault or unknown features to the basis functions (or known signal pattern) [27, 28]. Among the data-driven based fault diagnosis, it is prevailing to adopt the unsupervised methods like PCA and ICA for fault detection since the amount of data from the health conditions could be used for training and establishing a knowledge basis [29, 30], while they have to confront the challenge of information missing, especially for the data in high dimensionality. Clustering methods are another reasonable choices for addressing the harsh conditions of no or less prior knowledge of faults [31, 32], but the reliability would be decreased when the samples are small in size and the knowledge basis is incomplete (i.e. limited prior knowledge of the normal conditions). Statistical methods and statistic-based indicators like Kolmogorov-Smirnov test (K-S test) [33-35], analysis of variance (ANOVA) [36, 37], rank-sum testing [38, 39], Bayesian hypothesis testing [40, 41], and likelihood-ratio test [42-44] have been widely used for fault or damage diagnosis since they are standardized and easy to implement though the available data might be limited. The thresholds associated with a

statistical test are obtained from p-values listed in statistical tables. They are used to measure the underlying probability distribution differences between two groups of data, which has been successfully applied to fault detection. However, a statistical test is effective as the reference data are highly representative and satisfy the assumptions related. The degree of correlation between two datasets does not prove the evidence of causation, and the results of a statistical test cannot reveal a definite event.

Therefore, there is a demand for the techniques that can detect the health condition of the system without overcomplicating the model and with less dependence on prior knowledge.

1.2. Objectives and Methodology

This study focuses on the development of an effective method for fault diagnosis of complex structures in which available data are extremely limited. Therefore, the objectives of this study mainly include:

- 1) An effective fault diagnosis technique should be investigated for both single and multiple fault detection and localization of the complex structures though there is no or less prior knowledge of the faults or/and limited historic data measured from the normal operational conditions involved;
- 2) The fault diagnosis method is practical and easy to implement in real structures with less parameter assumptions and no limitation on stationary data;

3) The fault diagnosis system consumes less computational cost, and can be done in a reasonable offline time or on line.

To accomplish the objectives above, a fault diagnosis method, named as virtual beam-like structure (VBLS) approach, is proposed. Figure 1-1 shows the overall flowchart of the methodology exploited in this study. Generally, the proposed methodology comprises a series of steps: data acquisition, signal preprocessing (feature characterization and fault indicators definition), fault detection, the optimal virtual beam construction, and fault localization based on virtual beams.

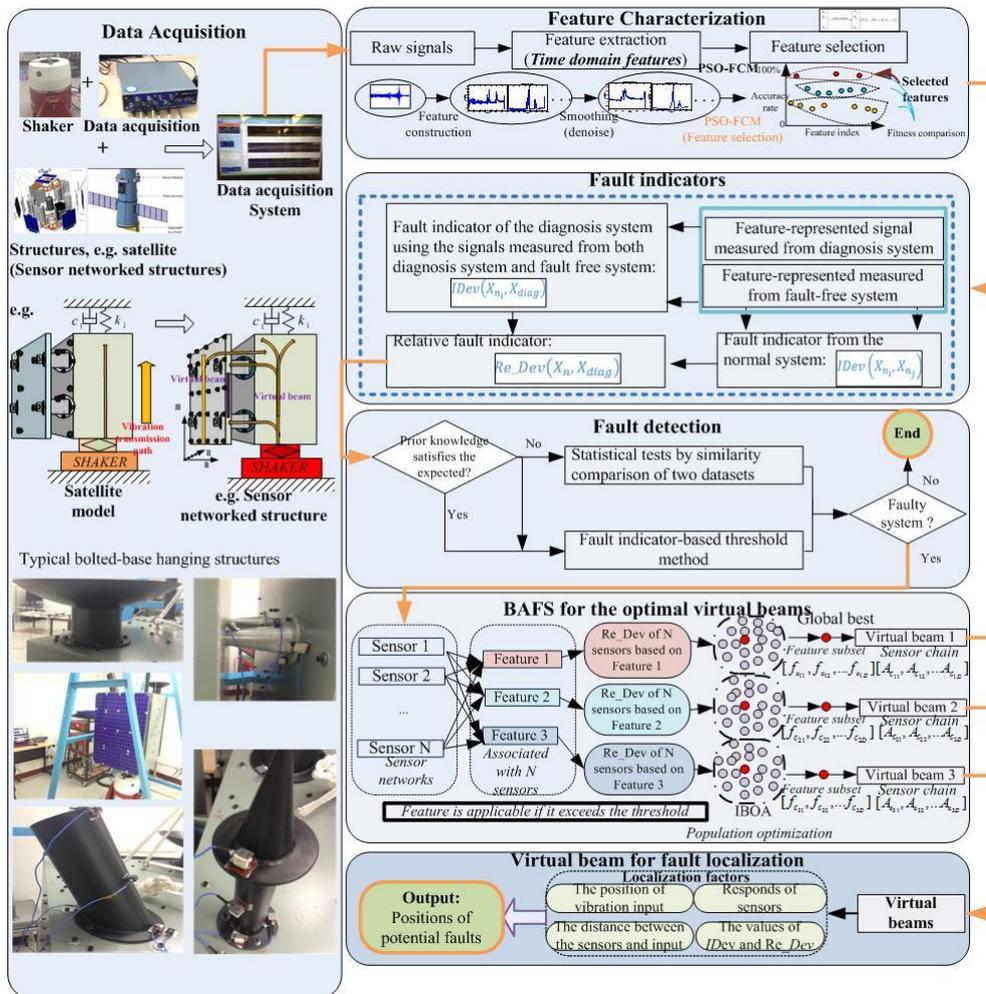


Figure 1-1 The overall framework of the proposed methodology

Considering the computational effort as well as the feature capability, in the signal

preprocessing stage, the appropriate time domain features are adopted to transform the high-dimensional raw time series signals into the feature-based statistical vectors with lower-dimensionality for fault diagnosis. To address the harsh conditions with unavailability of prior knowledge of faults and limited initial data from the normal conditions involved, statistical approaches towards fault detection, together with adaptive threshold techniques, are utilized in the method. A virtual beam based approach is developed for fault localization by taking into consideration of the damage information which could be captured by sensor networks with regard to the changes of vibration transmission. The vibration transmission path represented by the sensor chain is regarded as the virtual beam and is constructed automatically using an evolutionary optimization method.

Therefore, the main contributions of this study lie in that:

- A systematic virtual beam-like structure approach is proposed for the multiple-fault localization of complex structures without using prior knowledge of the faults, and using only limited prior knowledge of normal operational conditions.
- The proposed fault diagnosis method also employs a modified optimization methodology (i.e., an evolutionary optimization method) to automatically select the most discriminative sensor chains from the sensor networks for fault localization, which makes the proposed method more practical and easy to implement in real-world sensor-networked structures with less parameter

assumptions and no limitation on stationary data.

- To address the multiple fault localization problem, a new strategy named biased running is developed in the optimization methodology by constructing the multiple biased virtual beams so as to isolate the potential faults in a faulty component. Moreover, a statistical method together with adaptive threshold techniques are also presented for multi-fault detection to overcome the challenge of obtaining an appropriate threshold value with only limited experimental data.
- Sensor networks mounted on the structure are separated according to their regions of coverage and influence associated with the events, and the high-dimensional signals are characterized in forms of low-dimensional time domain features, which can help minimize the computational complexity and maximize the diagnosis accuracy.

Generally, with less or limited requirement of priori, a comprehensive fault detection and fault localization system based on an evolutionary optimization method is developed in this study for fault diagnosis of complex structures with single or multiple faults in various mechanical engineering systems using sensor network attached.

1.3. Outline of the Thesis

This thesis describes a sensor network based virtual beam-like structure method and its application in monitoring the health condition of complex structures with

limited or less prior knowledge. Figure 1-2 shows the overall structure of the thesis.

Specifically, the rest of the thesis is organized as follows.

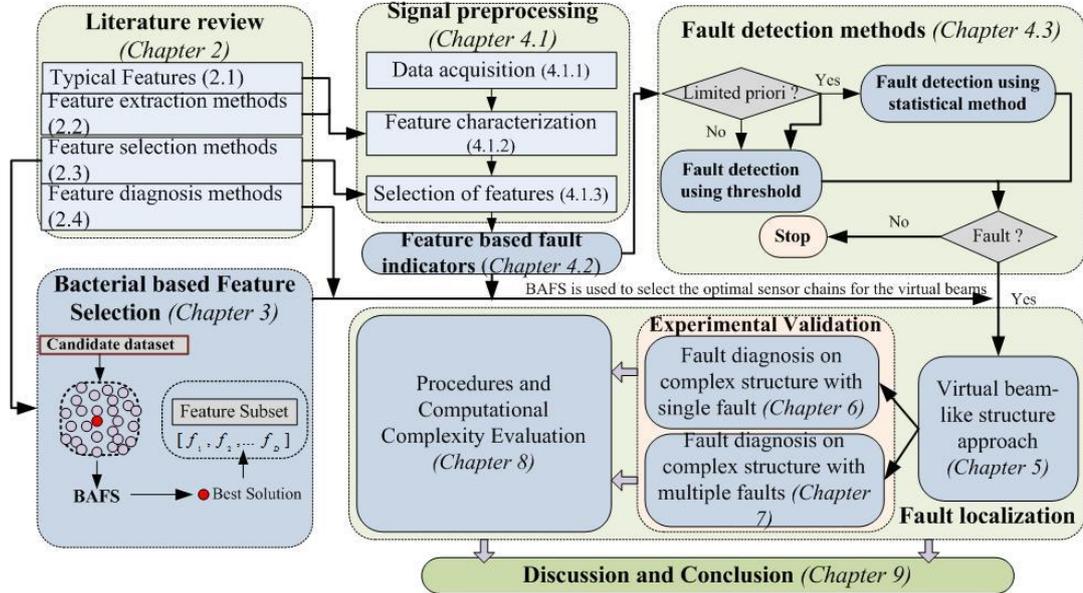


Figure 1-2 The overall structure of the thesis

Chapter 2 presents a literature review on the existing non-model based fault diagnosis methods according to their feature represented forms. This chapter consists of three subsections: typical features frequently used in engineering, feature extraction methods, and feature selection methods. This Chapter studies the advantages as well as the disadvantages of these methods and points out the research gaps of existing studies to provide the potential techniques for fault diagnosis.

Chapter 3 introduces a bacterial based feature selection method which will be served as the optimization method to select the appropriate sensors from the network to construct the “virtual beam”. The optimal ‘virtual beam’ consisting of sensor chain is essentially important for fault localization, which will be generated automatically using the optimization method proposed in this chapter. Meanwhile,

the effectiveness of the proposed optimization feature selection will be tested using several frequently used benchmark datasets.

In Chapter 4, the fault indicators and time domain features used for fault diagnosis are presented. An unsupervised fuzzy clustering method is developed to select the most informative features from the available feature representatives. The statistical tests and fault indicator-based threshold method are combined for fault detection to overcome the challenge of obtaining an appropriate threshold value with only limited experimental data.

Chapter 5 presents the new proposed vibration beam-like structure approach. Virtual beam plays the important role for fault localization. The construction of optimal “virtual beam” is realized by the optimization method proposed in Chapter 3, including the selection of sensor networks on complex structures, objective function and the constraints used by optimization methodology for virtual beams construction, and summarization of the general rules for fault localization based on the optimal virtual beams.

Chapter 6 demonstrates the effectiveness of the proposed virtual beam-like structure method by applying it to detecting and isolating the loosening screw (only one fault in the system) of a satellite-like structure. The typical complex substructures on the satellite-like model to be detected include solar panel, main structure body (or body unit), and band antenna.

In Chapter 7, the virtual beam-like structure approach is applied to the fault

diagnosis of the satellite-like structure with multiple loosening screws occurring on the various components without using prior knowledge of faults and using limited prior knowledge of normal conditions. The classical sensor network based multiple fault localization method is adopted for comparison.

Chapter 8 gives the detailed procedures of virtual beam-like structure approach, discusses the applicability of the proposed method for more complex structures, and studies the computational complexity of the proposed fault diagnosis system to provide a promising method for on-line application.

Chapter 9 summarizes the importance and significance of the proposed virtual beam-like structure approach. Meanwhile, the potential improvements in the future studies are discussed in this chapter.

Chapter 2. Literature Review

As noted previously, the need for fault diagnosis in a complex dynamic structure system has led to the research impetus on time series data analysis. Studies using time series data for fault diagnosis usually begin with the extraction of informative features while representing the time series signals. Any kind of time series can be represented by features irrespective of whether they work directly with raw data, or indirectly using raw derived models or extracted features; in short, almost all representations can be expressed in feature-based forms. A rising number of researchers have been investigating different features of the associated extraction algorithms for signal analysis.

This section reviews related topics involving feature based methods for fault diagnosis. In addition, the evolutionary optimization methods for feature selection are also discussed since they will be applied to selecting efficient sensors (represented by time domain features) while determining the optimal virtual beams (this problem is formulated as an optimization problem of feature selection). Finally, the existing fault diagnosis approaches with limited requirements of the prior knowledge are reviewed.

2.1. Typical Features for Pattern Recognition

The features described for pattern recognition can be classified as statistical, structural, and hybrid techniques [5]. They will be introduced separately.

2.2.1. Features in Statistical Techniques

Features in statistical techniques are commonly described in quantitative forms and the data are discriminated among different groups according to certain quantitative features such as root mean square, kurtosis, skewness, variance, crest factor and correlation coefficient. Statistical decision theories like T-statistics or p-value are usually used to determine the similarities or differences among two or multiple groups of data.

Statistical methods and statistic-based indicators based on the above theories have been widely used for fault detection [45-48]. The thresholds in statistical tests are obtained empirically in accordance with statistical tables (or p-values). The T-test and related non-parametric tests can be used for fault detection by comparing two datasets or two distributions (obtained from repeated measurements conducted on the same system). The T-test is a classical statistical method capable of comparing the differences in means and variance. Among the most frequently used T-tests, two sample (unpaired or independent) T-test and paired (or repeated measures) T-test are employed to compare the differences between two populations under the assumption that the samples are from the same distribution with equal variance. Such methods are more likely to be used as an effective tool while seeking to specify the threshold value for fault detection [47, 49-51].

If the samples are not from a Gaussian (or Normal) distribution, nonparametric tests are more applicable since they do not depend on the above assumption.

Commonly used non-parametric tests include the Wilcoxon signed rank test (or the Wilcoxon matched pairs test), Wilcoxon rank-sum test, Kolmogorov-Smirnov (K-S) test, Kruskal-Wallis test, and Friedman's test. The last two tests are not suitable for two-group data since Kruskal-Wallis test compares three or more unmatched groups and Friedman's test compares three or more matched or paired groups. Wilcoxon rank sum test detects the differences between median values and has been applied widely to fault detection [38, 52-56]. Kolmogorov-Smirnov test (K-S test) [57] measures the underlying probability distribution differences between two group data. It has been successfully applied to the fault detection for gears [35, 58].

Even so, a limitation of statistical methods is that the interrelationships embedded within the data might not be able to manifest in statistical features, which makes it difficult to discriminate among data drawn from different groups.

2.2.2. Features in Structural Techniques

In structural techniques, the objects are distinguished according to certain shape-based or structure-based features. Referred to as the primitives, structure features are used to represent the inherent relationships embedded within the data. Structural pattern recognition is applicable, in most cases, for discrimination among the groups based on inherent and identifiable natures of objects such as image data and time series data. Image data are recognized according to their visual rendering. In the example discussed in [59], a 1-D histogram contains certain structural and spatial information of the image, and the histogram is the image feature used for further

study. In image matching issues, points of interests (POIs) are extracted as features to enhance the image measure in question, so they are referred to as point-based features [60]. For high level geometrical subjects extraction, the contourlet transform has also been investigated to transform the given texture image into contourlet coefficients which are the data-based features to be evaluated [61]. Time series data, organized by time, are constructed using structural pattern recognition system [5].

But the implementation of the structural pattern recognition system is limited mainly by the extraction of the structural features. In addition, the structural features can be used to represent a particular object only if the structural information of the primitives (like interrelationships) have all been identified. As there is no general method for structural features, this study focuses on the statistical features and relevant feature extraction methods.

2.2.3. Features in Hybrid Techniques

Hybrid techniques combine the characteristics of both statistical and structural approaches so as to compensate for the drawbacks of statistical and structural methods, and to take advantage of the two groups. For example, both statistical features (e.g., kurtosis, skewness, etc.) and structural features (i.e., histogram features) have been extracted from the time domain vibration signals for diagnosis of bearing faults [62]. As the classification of the structural pattern recognition system is a challenging task [5], it has become popular to apply statistical

classification methods to distinguish objects according to their structural or shape-based features. In biological systems, structural properties relevant to gene regulatory networks have been defined and explored for observation and controlling using statistical or quantitative methods in [63]. In [64], time series data are represented by structural patterns, and problems embedded within the time series are classified by applying statistical classification algorithms (e.g., structural generative descriptions).

Though hybrid methods seek to preserve the merits of the two approaches, it is not a general method applicable to all pattern recognition systems in view of considerations pertaining to time-consumption, feature extraction, interrelationships, etc. There is no obvious guideline for choosing one test that is suitable for all applications. Considering the computational complexity and implementation, in this thesis, the available statistical features are adopted as representations of signals to facilitate the pattern recognition.

2.2. Feature Extraction Methods

Originally, the data measured by vibration sensors are time series data. The increasing need for fault diagnosis of complex structures has led to research on time series analysis. The noise as well as the wide spectrum of defective signals has contributed to the difficulty in fault diagnosis using time domain vibration signals directly [21]. Therefore, features are extracted from observed time series.

Definition: Feature Extraction. Given a feature set $X = \{x_i | i = 1, \dots, N\} \in R^N$,

find a mapping $Y = f(X): R^N \rightarrow R^M$, with $M < N$ such that the resulting feature vectors $Y = \{y_1, \dots, y_m\} \in R^M$ characters the most information that the original series possesses.

Feature extraction is used to transform high-dimensional raw signals into a different space with lower dimensionality, and a new subset is created after completing the process of feature extraction. Actually, the purpose of feature extraction algorithms is to describe the most representative attributes of objects. Many approaches have been developed by researchers to achieve this. Some have simply divided features into time domain features, frequency domain features, and time-frequency domain features [65-68]. In this section, methods based on different feature characteristics are studied and analyzed.

2.2.1. Time Domain Features based Techniques

Since they are simple and interpretable without need of laborious tasks like framing, windowing, filtering, one of the main advantages of time domain-features (e.g., RMS, variance, kurtosis, crest factor, correlation coefficient) is to avoid complexity of preprocessing, which has led to their wide application.

To identify faults, time domain vibration signal segments have been represented by time domain features (e.g., RMS, variance, skewness, kurtosis) for on-line condition monitoring of machines [21]. The time domain features (chosen in that study) of signals identified from the fault and fault-free systems associated were employed for training, while the features of signals collected from the diagnosis

system were classified using an ANN-based diagnosis approach. In [18], the disturbance impact factors were studied by employing the individual or/and combined time domain features for electromyographic (EMG) pattern recognition. The results indicate that an approach with combined EMG features has the advantage of improving the classification performance. To analyze non-stationary and nonlinear characteristics of vibration signals, time domain features like peak-to-peak were applied in [69], for fault diagnosis of fans using an improved k-means clustering method. Elsewhere [19], twelve time domain features were employed for identifying the gear defects SVM as a classifier.

The common characteristics of the above methods are that they are easily implemented, lower in terms of computational cost, and are applicable for on-line condition monitoring. It is normal to apply an effective classifier for classification tasks. However, one of the main limitations associated with these methods is the necessity for prior knowledge concerning the system, including normal and defective states with different fault types and extents.

2.2.2. Frequency Domain Features based Techniques

Frequency domain features (e.g., Fast Fourier Transform) are transformed from time domain data to isolate the fault at certain frequency components [16], which enable their wide application to fault diagnosis [24, 70, 71].

Fourier Transforms (FTs) are among the most traditional techniques used for feature extraction, which convert the sources into the time-frequency domain. Many

variations of the approach have already been developed. Fast Fourier Transform (FFT) is an FT method that speeds up the operational process. It is generally applied to extract the features embedded in a vibration signal using FFT and then applies classifiers such as KNN and SVM to assess the condition of machine conditions [23, 72, 73].

Despite their common application, FTs have to rely on a fixed-time window. As commented in [74] this type of technique has to "make a trade-off between frequency and temporal resolution". Though FTs are very useful tools for isolating faults at certain frequency components in stationary cases, it is not suitable for non-stationary situations. Subsequently, time-frequency domain features like wavelets have been developed for dealing with non-stationary signals.

2.2.3. Time-Frequency Domain Features based Techniques

Time-frequency domain methods were developed as the potential techniques for fault detection and diagnosis of machinery structures by analyzing both the time and frequency contents of signals. These methods include Short-time Fourier Transform (STFT), Wigner-Ville distribution (WVD), and Wavelet Transforms (WTs) [75].

Short-time Fourier Transform (STFT) is a time-frequency analysis used to overcome the disadvantage of FFT and provide both the time and frequency contents of given vibration signals. In [76], the time-frequency domain signals consisting of multi-frequency components were analyzed using STFT to identify the faults of a voltage source inverter at the early stage. However, a major drawback of STFT is

that the resolution obtained is provided for all frequency components. To address this disadvantage, STFT has been combined with wavelets [77, 78]. As a result, variable windowed STFT was developed to enable an examination of multi-resolution associated with the frequency components.

WVD is a quadratic-form of analysis offering resolution in both time and frequency domains. It has been widely applied to fault diagnosis through an analysis of given vibration signals. The main advantages associated with the technique are simplicity and low computational cost, whereas the main limitation associated with the classical WVD is the challenge of interpretation because of the interference terms [79]. A parameter called NP4 was developed to enhance the capability of WVD and improve the interpretation in gear fault detection in [80]. In [81], a weight factor was introduced to WVD so as to alleviate the influence of linear frequency modulation.

Wavelets are capable of treating non-stationary signals [82-85]. A key point is that temporal resolution enables a family of wavelets to outperform the FT method. Unlike STFT, wavelets allow different window sizes to analyze the varying frequency components embedded in vibration signals (e.g., a long window size for low frequencies and a short window size for high frequencies). Two classical forms of wavelet transforms for machinery fault diagnosis are continuous wavelet transform (CWT) [86] and discrete wavelet transform (DWT) [87, 88]. The main difference between the two wavelets lies in the handling of the subset scale for

translation values. The former seeks to represent the signal with all possible subset scales while the latter aims to translate using a specific subset scale (without redundant information). Apart from the above two methods, the wavelet family consists of a series of variant wavelets, such as Wavelet Package Transform (WPT) [89], and Morlet Wavelet Transform [90].

In addition to the above mentioned time-frequency domain features, it is popular to apply EMD decomposition to handle the non-stationary signals arising in fault diagnosis. To decrease the mode-mixing problem in EMD, an ensemble of EMD (EEMD) was utilized in [91] to realize the continuity of signals over varying scales. Compared to those from STFT, WVD, and DWT, experimental results have demonstrated the superiority of EEMD in revealing hidden knock signatures. As one of time series energy representation, Hilbert-Huang Transform (HHM) combines EMD and Hilbert spectral analysis (HSA) to analyze non-stationary and non-linear time series signals. However, in performing HHT, it becomes necessary to overcome the drawbacks associated with EMD. In [92], HHM was improved by adopting a specific strategy of decomposition of narrowband signals. In [93], translation-invariant denoising (TID) was used as the preprocessing technique to alleviate the interference of noise and enhance the quality of decomposition of HHT. A comparative discussion study has been presented in [94], which indicates that HHT is more appropriate for analyzing the signals of larger size than using wavelets.

However, the effectiveness of time-frequency domain features is offset by the

high computational cost, especially in dealing with high dimensional time series data, which makes it difficult to apply them to online health monitoring.

2.3. Feature Selection Methods

Except for feature extraction, another effective method to address the high dimensional feature characterization problem relies on feature selection (FS) which is to select a relatively small but potentially useful feature subsets from available features [95]. Since there is no specific feature selection technique that is applicable in all situations, numerous feature selection methods have been proposed for a variety of applications.

This section provides the background for feature selection methods and presents certain related works seeking for effective feature selection methods adopted in further studies.

2.3.1. Traditional Feature Selection Algorithms

From the perspective of the evaluation criteria, feature selection techniques can be divided simply into filter, wrapper, hybrids and embedded methods [96].

The filtering approaches select the subset of features, firstly, according to their own discriminant attributes evaluated by numerous metrics before using learning machines. The metrics include hypothesis tests on the coefficients of features by means of t-statistics and p-values [97], relief algorithms (RF) [98], information gain (IG) [99], consistency-based feature selection [100], minimum redundancy

maximum relevance (MRMR) [101], interaction feature selection [102], dependency margin [103], etc. Filter approaches are easy to implement and expected to be faster than other types of feature selection algorithms since the learning machines are used only in the final stage of performance assessment, i.e., after filtering undesirable features [104]. Even so, the filter methods do not need any feedback from the learning machine in feature selection. Further, more often than not, the features with the highest-ranked values are selected on the basis of statistical criteria such as p-values but without scoring the power of combined features [105]. As a result, they might not be able to obtain the best subset of features in the total feature space because the highest-ranked features may have redundancy among them [106]. In practice, it is more common that filters are embedded or cooperate with wrapper feature selection algorithms, called hybrid methods [106-112].

Wrappers seek to find the best subset from the feature space according to predetermined performance assessments conducted via classification algorithms such as the K-Nearest Neighbor algorithm (KNN) and Naive Bayes (NB). Greedy search based sequential forward selection (SFS) and sequential backward selection (SBS) are two attractive classical wrappers used in computation. However, their main drawback is that the features cannot be selected (or removed) again once they have been removed (or selected) [113]. Besides, to select the optimal feature sets, all possible subsets of features will need to be measured by the wrappers—which is impossible while classifying high dimensional microarray gene expression cancer

problems because of high computational cost. To address this problem, wrappers based on evolutionary algorithms are applied without the necessity of evaluating all possible subsets [114] [115]. Compared to filters, wrapper-based feature selection methods are generally more accurate since the quality of the selected feature subsets is assessed using a predefined learning technique.

Hybrid techniques are often developed with the combination of the filters and wrappers, or/and combination of the wrappers and wrappers to enhance the performance. The main trend related to the combination of filters and wrappers is to integrate the filters (such as the T-test and the χ^2 -test) with the evolutionary algorithms [106, 108, 110-112], in which the filters are served as the preprocessing step to decrease the redundant and unimportant features in huge dimensionality while wrappers are used to evaluate the relevance of the remaining feature subsets according to a given learning machine or classifier (with a view to reducing the computational problems associated with the wrappers and enhancing the effectiveness of feature selection). The problem is that the features seem to be less responsive to the labels but are useful in the feature sets for classification that might have been removed during the preprocessing stage, which cannot guarantee the optimal feature sets in this situation. Among the combination of wrappers and wrappers, hybrid evolutionary algorithms have become particularly popular [107, 109, 116-118] by taking advantages of certain brilliant mechanisms associated with each wrapper or looking for the complements among other wrappers to improve the

robustness and stability of the final solution. It is a good strategy to improve the performance of wrappers if the hybrid strategies could enhance the complement.

Embedded approaches search for features in the process of training and are usually specific to a given learning model using a limited number of features [119]. For instance, the regularizations or shrinkage methods like least absolute shrinkage and selection operator (lasso) [120], ridge regression [121] and elastic net [122] that have been incorporated into the development of a classification or regularization model [123]. The feature weightings are used to record the importance of features in classification. Thus, in comparison to wrappers, embedded methods usually need less intensive computation and become less prone to overfitting. However, typical embedded methods generally use quite strict model structure assumptions. Sometimes, the classification performances of embedded methods might turn out to be inferior in comparison to the filters and wrappers [124].

2.3.2. Evolutionary Feature Selection Algorithms

Among different feature selection techniques, evolutionary feature selection algorithms such as Genetic Algorithm (GA) [125, 126], Differential Evolution (DE) [127], Particle Swarm Optimization (PSO) [113], Ant Colony Optimization (ACO) [109] have become particularly popular search techniques over the past few years. Inspired by natural as well as artificial ideas related to evolution, Evolutionary Computation (EC), as a sub-field of artificial intelligence, is generally considered as an effective technique capable of addressing expensive optimization problems.

There is little doubt that feature selection problems with numerous redundant and repeated features belong to some of the most expensive problems. In most evolutionary algorithms, each unit represents a potential solution which can be updated continuously according to the machine utilized. The advantages associated with such evolutionary algorithms are that the methods are available for optimization by specializing certain agents suitable for the specific problem, and that the results are generally creative and qualified [128].

Inspired by bird flocking or fish schooling and known for its high convergence capabilities, PSO is a highly common evolutionary algorithm in the domain of feature selection [113, 129, 130]. Even so, a disadvantage associated with it is that it gets easily trapped into local minima and therefore cannot guarantee a globally optimal solution.

Apart from PSO, other swarm algorithms, such as ACO [131], Artificial Bee Colony (ABC) [132], Artificial immune systems [133], have also been used for feature selection. The search for the optimal feature subsets in ACO is implemented via a group of ants traveling through the graph to minimize the number of informative nodes. The main advantage of this method is that the best solution can be obtained by a global sharing strategy. The drawback is that the optimal feature subset may not always be arrived at, when the feature space being searched is large.

Further, genetic or evolutionary based algorithms such as Differential Evolution (DE) [108, 134, 135], GA [126, 136, 137], Genetic programming (GP) [114, 138]

which have been frequently shown to be effective in solving problems under complex conditions are presented for FS. Most of these are binary description algorithms, where each bit represents a feature, e.g., “1” represents a selected feature while “0” means a feature that has not been. However, the parameters used in genetic or evolutionary based algorithms need to be properly defined to attain acceptable performance.

Since swarm algorithms such as PSO and ACO are good at local research and evolutionary algorithms such as GA have superior global searching abilities, several researchers have come up with combinations of swarm algorithms and evolutionary algorithms: e.g., Hybrid GA-ACO [139] and Hybrid GA-ACO-PSO [140]. Even so, the effectiveness of feature selection is decreased when the size of search space increases.

Most of such feature selection methods have been found to suffer from computational complexity or local optima. In addressing high dimensionality feature selection problems, there is a clear need for an effective feature selection method with global search ability.

2.3.3. Bacterial based Feature Selection Algorithms

The bacterial foraging optimization (BFO) method proposed by Passino [141] and the bacteria chemotaxis (BC) approach developed by Müller [142] are perhaps the two earliest bacteria based algorithms used in addressing optimization problems. A common feature of the two algorithms is that they were initially inspired by the

chemotactic (foraging) behavior of *E. coli*. As a group of EC techniques, due to their global searching capabilities during control and optimization, BAs have recently gained attention for feature selection.

Often, bacteria-based FS methods cooperate with other methods. Cho [143] combined the BFO with the notion of mutual information for feature selection for the purpose of classification. In [144], an agent genetic algorithm based on a bacteria foraging strategy (BFOA-L) was integrated with a neural network while implementing fuzzy logic reasoning for feature selection. Also, BFO working with particle swarm optimization (PSO) named bacterial foraging-particle swarm optimization (BFPSO) was developed for the feature selection associated ECG signals [145]. Almost all such bacteria-based FS algorithms, generally, have tried to mend the defects of the original BAs in computational cost as well as improve convergence speed by combining with other techniques.

Though BAs have been developed for feature selection they still have several limitations. Firstly, the random orientation strategy utilized in BAs has indeed increased the amount of search conducted globally although it has been found to consume larger time while searching randomly for the optimum. The resulting excessive computation time has limited the application of feature selection in high dimensional datasets. Besides, the circulation machines embedded in original BAs could not guarantee acceptable convergence speeds. Therefore, those BAs were still almost universally immature. In this thesis, strategies for improving capability and

efficiency of bacteria-based algorithms for feature selection will be developed and further employed for sensor subsets selection to isolate the potential faults occurring in a complex structure.

2.4. Fault Diagnosis Techniques using Limited Prior Knowledge

Prior knowledge of the system involving the normal operational states and abnormal conditions is very important for fault detection and fault diagnosis. However, in real applications, the prior knowledge of the system is not always available. Therefore, in the section, the *fault diagnosis techniques using limited prior knowledge* are reviewed from two aspects: model-based fault diagnosis methods and model-free fault diagnosis methods.

2.4.1. Model-based Fault Diagnosis Methods

The models of the processes or systems, in the model-based methods, could be built according to observers in the normal operation states with no prior knowledge of faults, and fault detection procedures normally depend on the residuals. Generally, hardware redundancy and software redundancy (or analytical redundancy) concepts have been presented for fault diagnosis. The main principle of the hardware redundancy concept is to compare the duplicated output signals using the same input [146]. Though the hardware redundancy methods are reliable, they are costly and consume more space for storage, which impetuses the development of analytical

redundancy methods since 1980s. For the analytical redundancy methods, there are two main trends for fault diagnosis: parameter estimations and residual generations.

Parameter estimations have attracted increasing interests for fault detection, in which faults are detected by comparing the actual parameters with the reference parameters obtained from the healthy conditions. This group of methods is brilliant for fault diagnosis if the physical features could be mapped to the explicit model parameters. In [147], the system was modeled as the interconnection of subsystems which were separately modeled as the transfer functions. The linear function of parameters named as diagnostic parameters was applied to fault detection by comparing the average residual energy with the threshold value. Without any requirement of prior fault information, a fault parameter tracking law was designed in [148] for detection of a submarine with the potential bow and stern plane faults. To estimate the model parameters, a recursive subspace identification algorithm was proposed for online fault diagnosis of dam-gallery open channel system [149]. Other similar studies were also made using the subspace-based method to estimate the parameters for fault detection [150-153].

Except for the parameter estimations, residual generations were developed for both linear and nonlinear systems with unknown inputs. A well-known model-based fault diagnosis method is Kalman Filter (KF). Developed by R. E. Kalman [154], the KF is proposed for fault detection by keeping track of estimates. The occurrence of faults was detected by KF-based residuals using hypothesis tests [155-157]. Since

The original application of KF is to the linear systems, recent studies have been made on the modified KF approaches like extended KFs [38], adaptive unscented KFs [158, 159], and augmented state KFs [160] for the nonlinear states of the studied systems. Moreover, Bond Graph (BG) has been widely applied to the fault detection because of its capability of modeling the complex structure uniformly and analysis of the redundancy using causality concept to avoid the initial unknown condition [161]. In [162], a hybrid bond graph (HBG) model-based method was presented for multiple faults diagnosis using fault discrimination vector to distinguish the fault types after the abnormal state was detected according to the residual. To implement the easy-handling for fault detection, a quantitative approach based on the signed bond graph (SBG) was developed in [161] for multiple faults diagnosis. To ensure the online adaptation and self-optimization of the online system, the combination of the discrete-time Lyapunov stability theory and computational intelligence method was adopted in [163]. Though the bank of observer-based residuals is generally needed for fault isolation [146], a small number of studies were developed for fault isolation with unknown inputs observer using the fuzzy systems [164, 165], in which the domain knowledge can be obtained from the experts (obtaining the knowledge from the assembly process) rather than the historic failing data.

Generally, model-based fault diagnosis methods are capable of on-line fault diagnosis without prior knowledge of fault conditions, but the known model based

on normal states should be well built initially. For analytical redundancy methods, the effectiveness greatly depends on the accuracy of an initially known model. If the physical information of the structure is well-defined, the fault diagnosis can be reliable for the model-based fault diagnosis methods. Since the initial model is not that easy to acquire, especially in complex dynamic systems, there has been growing interest in model-free based fault diagnosis methods.

2.4.2. Model-free based Fault Diagnosis Techniques

Being simple and flexible for fault diagnosis without dependence on specific structures, model-free based methods have been widely applied to fault diagnosis. Generally, model-free based fault diagnosis methods can be divided as signal-based approaches and data-driven approaches (or knowledge-based approaches) [166]. The former group prefers to adopt the similarity comparison between the references and the estimates, while the latter focuses on the explorations of the underlying information based on available data.

As illustrated previously, signal-based approaches are more likely to detect the health condition of the system by comparing the reference signals measured from the normal conditions with the estimate signals collected from the diagnosis system. In existing methods, it is prevailing to explore the feature extraction methods in cooperation with classification and identification processes. Concerning the unavailability of prior knowledge of faults, it is often for signal-based methods to detect the health condition of the system by matching the fault or unknown features

to the basis functions (or known signal pattern) [27]. For multiple fault conditions, in [28], both the valve spring faults and valve clearance faults were studied on a diesel engine using only one accelerometer. The novelty of this method has been confirmed in detection and classification of the multiple faults occurring in the multiple elements without prior knowledge of the relation between the faults, but various historic data are needed for training in classification. The strategy of separating the signals into fault-related bandwidths was developed in [167] for multiple fault diagnosis, and the presence of single or combined faults was identified using a spectral analysis based on the MUSIC algorithm.

Data-driven approaches are to explore the underlying knowledge from a large amount of historic data, which could be grouped as supervised and unsupervised approaches. In supervised methods, classifiers are trained using the historic data containing both the normal condition and abnormal conditions. Different from the supervised methods, the data from the health conditions are applied to training and establishing a knowledge basis. Both algorithms have their advantages and disadvantages, but only the unsupervised methods could be applicable of the system with no prior knowledge of faults. A cold start fault detection framework was proposed in [168] using the historic data only measured in the normal system. The data-drive methods PCA, SVM, and PCA-SVM were studied for fault detection, indicating that PCA-SVM method had the advantages in comparison to two other methods. PCA and ICA are two most popular unsupervised fault detection methods

[29, 30]. Additionally, the clustering methods are another reasonable choice for addressing the harsh conditions of no or less prior knowledge of faults. In [31], clustering method combined with feature extraction and nearest neighbor based-outlier removal was proposed for health monitoring of bridges with limited knowledge of structural faults. The combination of the multiple unsupervised classifiers and fuzzy-C clustering was developed in [32] to measure the wear status of slurry pumps when available historic data were extremely limited. To address the one-class classification, in [169], support vector data description was considered as the promising idea for the extreme condition that only data from normal operation condition were available.

Generally, there is no distinct difference between the signal based methods and data-driven based methods for fault diagnosis when the prior knowledge is limited. Fault detection mainly relies on the similarity comparison between the unknown states with the known basis obtained from the initially healthy conditions. Even so, there is a challenge for above methods to obtain an appropriate threshold value for fault detection. Statistical methods and statistic-based indicators like Kolmogorov-Smirnov test (K-S test) [33-35], analysis of variance (ANOVA) [36, 37], rank-sum testing [38, 39], Bayesian hypothesis testing [40, 41], and likelihood-ratio test [42-44] have been widely used for fault or damage diagnosis since they are standardized and easy to implement though the available data might be limited. The thresholds associated with a statistical test are obtained from p-values listed in statistical tables.

The measures the underlying probability distribution differences between two group data, which has been successfully applied to fault detection of gears.

After the fault detection, the next step is to isolate the potential faults occurring in the structure. Among the model-free localization methods without a requirement of prior knowledge of faults, a series of binary estimators are proposed for fault localization because of simple, low-cost and fault-tolerant. In [170], a method named as subtract on negative add on positive (SNAP) was proposed for fault localization using the sensor networks. A likelihood matrix was created to record the (negative or positive) contribution of each sensor node based on their observations. The potential fault is isolated at the event with the maximum value in that matrix. To address the multiple fault localization, the method named as decentralized subtract on negative add on positive (DSNAP) [171] was developed, in which all the alarm sensor nodes were the fusion center and tracking the response of their neighbors. In [172], the frequently used binary estimators like fault tolerant maximum likelihood (FTML), Centroid Estimator (CE), Maximum Likelihood (ML), and SNAP were compared, indicating that the SNAP was superior to other three methods for fault localization in terms of accuracy and computational complexity. To decrease the negative influence of faulty sensors for localization accuracy, the strategy of trust index was adopted in SNAP, abbreviated as TISNAP [173], to assign the weights to the nodes from the sensor network based on their historic records in failing alarm. Generally, the localization based on sensor

networks is effective and easy to implement for monitoring the large space structures.

2.5. Conclusion

In this Chapter, existing feature-based methods have been summarized in order to seek the potential application in fault diagnosis of the dynamic signal in machinery engineering. Considering the simple, interpretable properties in signal preprocessing, time domain features are adopted as the signal representatives for further signal analysis. In addition, for optimization of sensor network distribution and sensors selection for fault diagnosis, numerous feature selection methods have been studied. As a result, bacteria-based feature selection algorithms were chosen as the candidates because of the global search ability and computational efficiency.

Since the prior knowledge is not always available in the real applications, the fault diagnosis methods considering the harsh conditions are reviewed. Model-based methods are excluded since their effectiveness greatly relies on the accuracy of the initial models which are rather complex to build. Compared to model-based fault diagnosis, model-free techniques are more flexible for fault diagnosis without dependence on specific structures. To overcome the challenge to obtain an appropriate threshold value for fault detection, statistical methods are considered for fault detection since they are standardized and easy to implement though the available data are limited. Additionally, with the continuous advances in sensor technology and sensor placement methods, the sensor network can be implemented for fault

diagnosis of large space structures with low-cost and high-efficiency. Therefore, in this study, a systematic sensor network-based model-free approach is proposed for fault diagnosis of complex structures without using prior knowledge of the faults, and using only limited prior knowledge of normal operational conditions.

Chapter 3. Bacterial Optimization Algorithm for Feature Selection

The optimal ‘virtual beam’ consisting of sensors is essentially important for fault diagnosis in our study. In this Chapter, an improved bacterial based optimization method is developed to select the most sensitive sensors (represented by features) for the optimal virtual beam.

3.1. Definition of Feature Selection

In engineering application, feature reduction methods are frequently developed to convert the high dimensional space into lower dimension space. To eliminate the redundant and irrelevant features, feature selection techniques are widely developed to reduce the feature number, decrease the store space and computational time, as well as improve the classification performance.

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} \xrightarrow{\text{feature selection}} \begin{bmatrix} X_{i_1} \\ X_{i_2} \\ \vdots \\ X_{i_n} \end{bmatrix}, \{1, 2, \dots, m\} \supset \{i_1, i_2, \dots, i_n\} \quad (3-1)$$

As shown in Eq.(3-1), the outcome of the feature selection is the subset of the features by selecting the most informative and relevant feature subset to represent all features, Main advantages of feature selection methods lie in the minimizing the data missing and avoiding the difficulty in creating the new features.

3.2. Bacterial Algorithms

In the following, the bacterial based algorithms will be introduced and discussed firstly before presenting the proposed improvement.

3.2.1. Typical Bacterial Algorithms

Inspired by the foraging or chemotaxis processes of *E. coli* bacteria, Bacterial Foraging Optimization (BFO) [141] and Bacterial Chemotaxis (BC) [142] are two earliest bacterial based algorithms for optimization problems. These two algorithms are credited with starting a new heuristic family in computational intelligence and provide a global searching capability for control and optimization. Since their advent, several Bacterial Algorithms (BAs) and extensions to them have been applied to the research areas such as fault detection [145, 174], signal processing [175-177], and pattern recognition [178, 179].

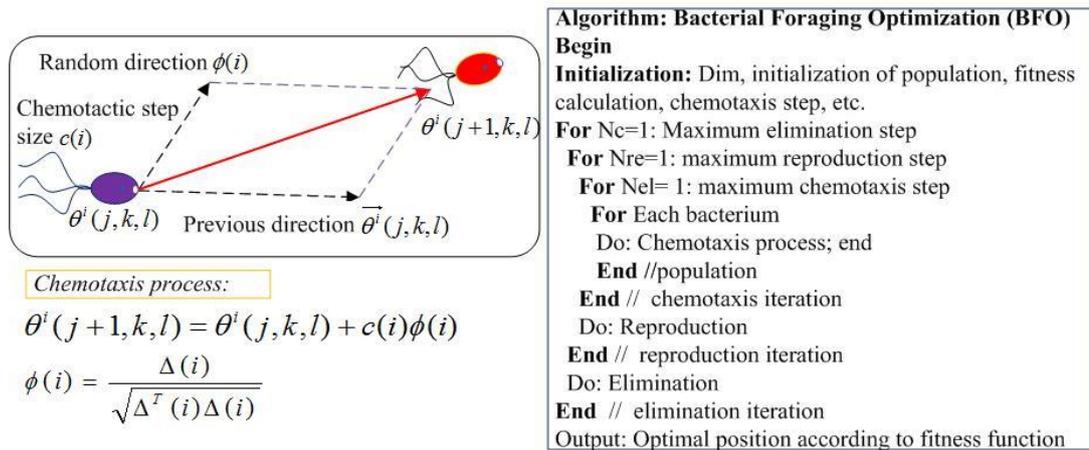


Figure 3-1 The overall framework of BFO

Figure 3-1 shows the overall framework of BFO. From the figure, one of main disadvantages associated with BFO or its extended Bacterial Algorithms (BAs) is

the expensive computational time needed for conducting the global search for the best solutions. The embedded interior circulation renders it impossible to solve the high dimensional problems arising while analyzing on-line systems. Many strategies have been applied to overcome this drawback. Among these, Bacterial Colony Optimization (BCO) [180] is the only method (to the best of our knowledge) capable of improving the search capability, fundamentally on the basis of the bacterial life-cycle, by adopting the conditional behavior modes. This method has been demonstrated to yield faster convergence and be more effective for search. However, BCO was originally proposed for solving continuous optimization problems; it has not been applied to discrete optimization problems (all variables are integers). By contrast, the optimization method proposed in this thesis is suitable for solving the discrete combinatorial problem of feature selection arising while selecting the most sensitive sensors (represented by time domain features) from the sensor networks. The motivation is to select the best subsets to achieve high performance. However, it results in a significant increase in the difficulties associated with the optimization method used during variable selection, an improved BCO will be developed in this thesis.

3.2.2. Bacterial Colony Optimization

Bacterial Colony Optimization (BCO) [180] has been derived from the two original bacterial behavior-based algorithms: Bacterial Foraging Optimization (BFO) [141] and Bacterial Chemotaxis (BC) [142]. Compared to bacterial

algorithms like BFO or BC, BCO designs the LCM and employs rule-based conditions to avoid excessive search for the best solutions. The learning mechanisms used by bacteria adopt multiple communication topology structures—such as dynamic neighbor oriented studies (or random oriented studies) and group oriented study—so as to attain the significant reductions in the computational cost associated with efforts to randomly search and increase the convergence speed.

- **Chemotaxis** As shown in Figure 3-2, chemotaxis process consists of two alternate strategies: running (or swarming) and tumbling. More specifically, the current position during the running process—see in Eq.(3-2)—is decided by the previous position and the learning experiences are oriented by the global or the individual exchanges. In comparison to running, tumbling includes an exceptional randomness to avoid the local optimal and pursue the diversity of the population—see in Eq.(3-3).

Running process:

$$P_i(T) = P_i(T-1) + R_i * (G_{best} - P_i(T-1)) + (1 - R_i) * (P_{best_i} - P_i(T-1)) \quad (3-2)$$

Tumbling process:

$$P_i(T) = P_i(T-1) + R_i * (G_{best} - P_i(T-1)) + (1 - R_i) * (P_{best_i} - P_i(T-1)) + C(i) * \Delta(i) / \sqrt{\Delta^T(i)\Delta(i)} \quad (3-3)$$

where $P_i(T)$ is the position of the i^{th} bacterium during T^{th} iteration, which is determined by the previous position $P_i(T-1)$, the best position of the population G_{best} (recorded as the global best), and the exchanged best position of an individual recorded as local best P_{best} . The method used for arriving at the exchanged best position P_{best} is given using the pseudo-code shown in Figure 3-2. $\Delta(i)$ is the

direction angle of the i^{th} bacterium randomly generated in $[-1, 1]$, and $C(i)$ is the chemotaxis step size. The larger values of $C(i)$ contribute to more randomness of the population for global search, while smaller values make it necessary to conduct more local search. $R = [R_1, \dots, R_{NP}]$ is a randomly generated vector consisting of 0 and 1. NP is the population size. More details can be referred to [180].

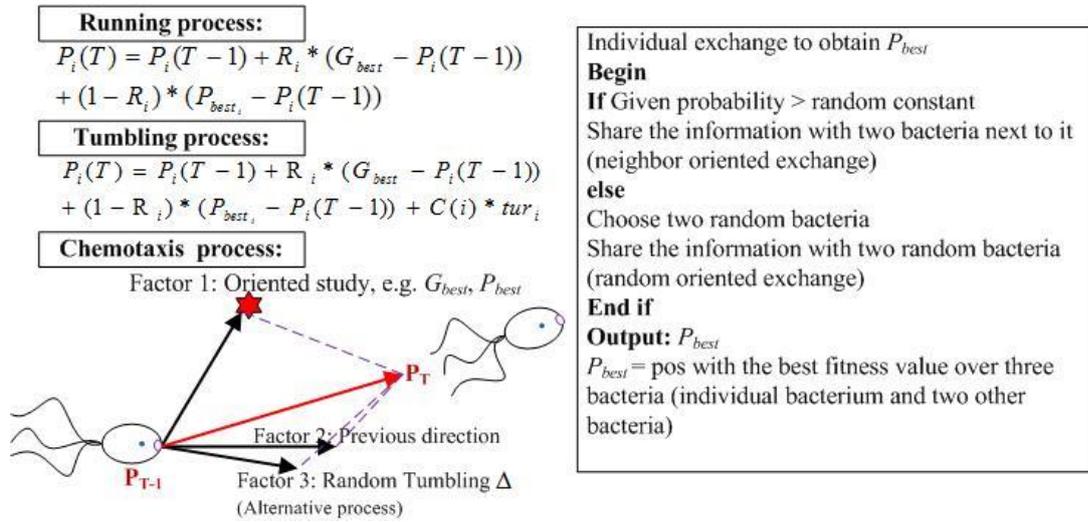


Figure 3-2 The main principle of BCO

Reproduction. After a predefined chemotactic step size, a reproduction mode will be employed. The reproduction process relies on the health condition of bacteria over the past chemotactic process. The first half ranking bacteria with the better performance will be kept in population and the second half bacteria with poor search capability will be replaced by the first half bacteria. The health index used to evaluate the searching capability of the j^{th} bacterium could be calculated as:

$$J_i(health) = \sum_{j=1}^{Nc} J(i, j, k, l) \quad (3-4)$$

and the population would be updated as:

$$\theta^{i+S_r}(j, k, l) = \theta^i(j, k, l) \quad (3-5)$$

where $J(i, j, k, l)$ is the fitness value of the i^{th} bacterium at the j^{th} chemotaxis, the k^{th} reproduction, the l^{th} dispersal, and $S_r = S/2$. If the optimization problem is to

minimize the objective fitness function, then the smaller values of $J_i(\text{health})$ means the better health condition of the bacterium.

- ♦ ***Elimination and dispersal.*** Following the predefined chemotactic steps and reproduction time, elimination-dispersal is taken to move the bacteria to the dynamic position.

$$\theta^i(j, k, l) = x_{min} + (x_{max} - x_{min}) \times rand \quad (3-6)$$

where x_{max} and x_{min} are the maximum and the minimum of the variables, and $rand$ is a randomly generated constant ranging from 0 to 1.

The reproduction in bacteria based algorithms is used to multiple the high-quality bacteria and remove the poor ones from the population, while the merit of the elimination operation is that it helps realize diversity through random dispersal or migration. More details can be referred to [180].

3.3. The Proposed Bacterial Algorithm based Feature Selection (BAFS)

Having described the proposed BCO approach for solving the continuous optimization problems, we turn now to the improvement—bacterial algorithm based feature selection (BAFS)—made with a view to addressing the combinatorial problem arising during the feature selection phase while trying to obtain the feature subsets from datasets.

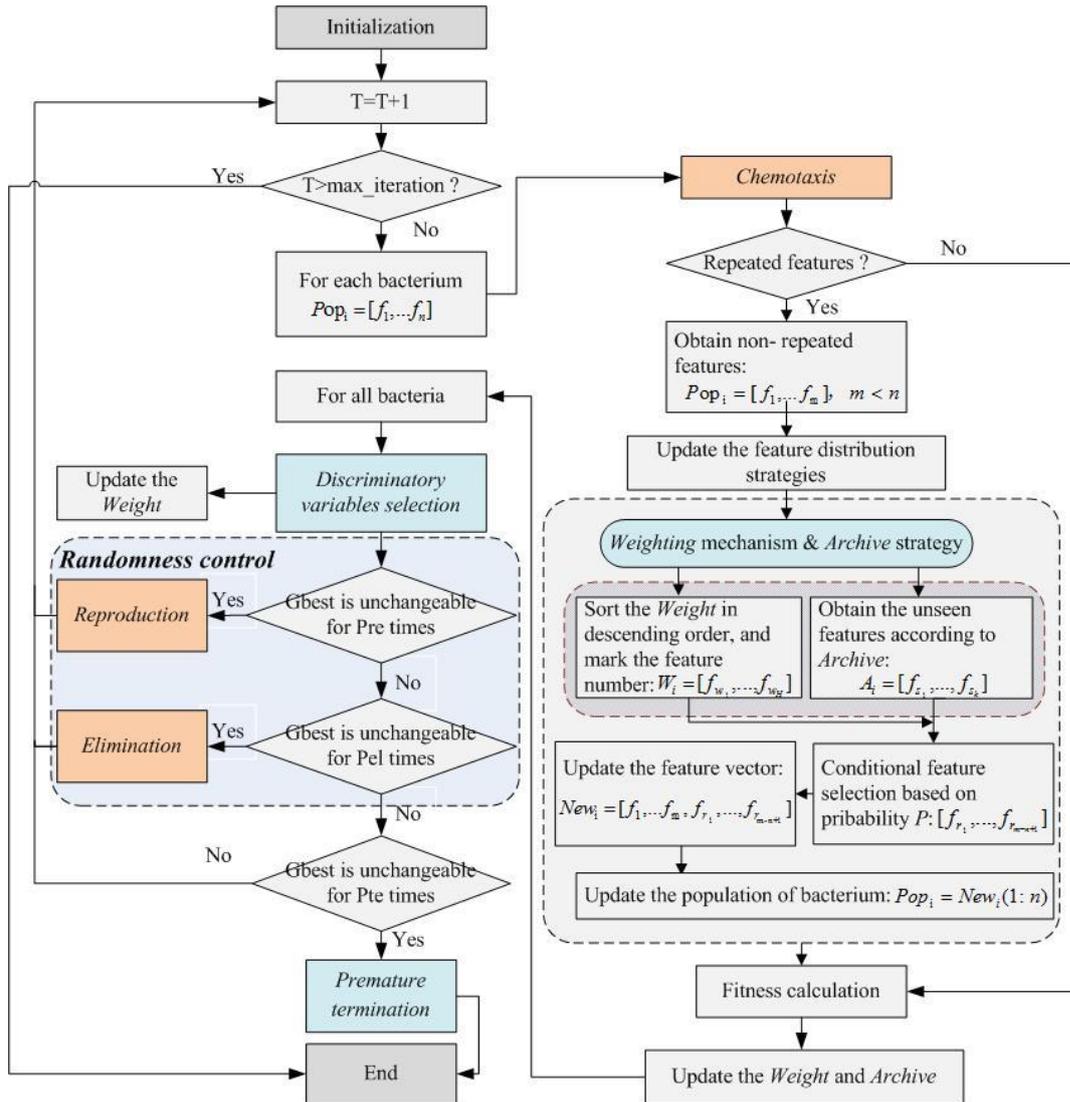


Figure 3-3 The overall framework of the Bacterial Algorithm based Feature Selection

The overall framework of the proposed feature selection method is given in Figure 3-3. The chemotaxis, reproduction and elimination process are the same as those in the original BCO, which can be referred to [180]. It is important to note here that, since the number of sensors for the optimal is predefined, the expected number of features in BAFS is known. Further, other researchers [106, 129, 181] have found that, among the thousands of features under consideration, only a few (numbering in tens) are usually of interest during the optimal solution of the feature

selection problems. A further reason is that the constrained feature selection method is useful in comparing the performances of different algorithms

Several strategies are proposed to enhance the effectiveness of BAFS including weighting mechanism & ‘Archive’ strategy, discriminatory variables selection, and premature termination. The main ones will be described in the following sections.

3.3.1. Weighting Mechanism and ‘Archive’ Strategy

To evaluate the candidate variables (all variables listed as the candidates for selection), a weighting mechanism and archiving strategy are proposed to record the variable performances and store the occurrences of candidate variables, respectively. In BAFS, the number of features to be selected are given in each run, and the repeated features in the same vector are removed according to two matrices ‘*Weight (W)*’ and ‘*Archive (A)*’. The performance of candidate variables is evaluated by a predefined fitness function. To simplify the problem, it is assumed that a larger value of the fitness function indicates better performance of the variable subset.

Assume that the number of candidate variables is H , the dimension of variable vector for evaluating the fitness function is D , and the bacterial population is N . The performance scores used to record the performance of features associated with each particle are stored as a $H \times N$ matrix W . The initial weight values in W are *zeros*. If the attendance of a variable yields to the higher fitness function value, the variable is deemed to be more effective than the replaced one provided that the new adding variable is assigned with higher weight in comparison to that of the replaced variable.

Finally, the weights representing the performance of variables in each individual are updated.

If $Fit_{(X_i,m)} > Fit_{(X_j,m)}$, then

$$W(i, m) = W(j, m) + \frac{|Fit_{(X_i,m)} - Fit_{(X_j,m)}|}{Fit_{(X_i,m)}} \quad (3-7)$$

Otherwise

$$W(i, m) = W(j, m) - \frac{|Fit_{(X_i,m)} - Fit_{(X_j,m)}|}{Fit_{(X_i,m)}} \quad (3-8)$$

The replacement of variable X_j with X_i contributes to the increase (or decrease) of fitness. The related weights are updated using Eq.(3-7) (or Eq.(3-8)) for the individual (i.e. m^{th} bacterium). If the selection of variable X_i is to replace the variable X_j in the m^{th} bacterium, and the current fitness values $Fit_{(X_i,m)}$ is larger than previous fitness value $Fit_{(X_j,m)}$, then the weight of X_i in m^{th} bacterium— $W(i, m)$ —will be larger than the weight of X_j in a same bacterium, i.e. $W(j, m)$. Otherwise, the weight of X_i would be smaller than the weight of X_j .

In the BAFS, the occurrence of variable across the optimization is stored in a

$H \times N$ matrix A , and the matrix is initialized as: $A = \begin{Bmatrix} 0, \dots, 0 \\ \vdots, \dots, \vdots \\ 0, \dots, 0 \end{Bmatrix}$. If the i^{th} feature

appears in m^{th} bacterium, the related element is updated as: $A(i, m) = 1$. The main contribution of matrix ‘ A ’ is to record the appearance of variables and ensure that the unseen variables from candidate variables have the chance appearing in the optimization process. Therefore, recording the occurrence of variables in a vector

could avoid the situation that certain variables are frequently selected while others never appear during subset assessment.

3.3.2. Discriminatory Variables Selection

To enhance the quality of the selected variable subsets, the variable is deemed to be of value, if it has appeared frequently at subsets exhibiting above average performance. The parameter F is used to indicate the distribution of the features. The weight (W) is then updated according to Eq.(3-9) and Eq. (3-10).

$$F = \left(\frac{G_i}{G_i + B_i} \right) / \max_j \left(\frac{G_j}{G_j + B_j} \right) \quad (3-9)$$

$$W(i, m) = W(i, m) + F \quad (3-10)$$

where G_i is the number of times that the i^{th} feature has been used in the subsets whose classification accuracies have been found to be higher than the average fitness value. Similarly, B_i is the number of times that the i^{th} feature has been employed in the subsets whose classification accuracies are lower than the average.

3.3.3. Randomness Control Mechanisms

To avoid the redundant searching for the optimal and qualify the population for feature selection, the randomness control mechanisms are investigated in the proposed bacterial feature selection method. Considering the computational cost, the proposed bacterial method adopts the life-cycle mechanisms in BCO using the predefined processing rules to operate the reproduction and elimination process. These predefined rules in the proposed method are re-designed for improving the

optimization capability and avoiding the randomness searching in typical bacterial based algorithms.

Three parameters are used for randomness control mechanisms. More specially, the parameter P_{te} is used to control the over-searching for the optimal, and parameters P_{re} and P_{el} are exploited for controlling the frequency of reproduction and elimination-dispersal strategies, respectively (see Pseudo-code 3-1).

Pseudo-code 3-1: Randomness Control

```

If global best is unchanged
    Record=Record+1;
else
    Record=0;
End if
If Record= $P_{te}$ 
    Success=1; Stop the searching process // termination to avoid the redundant searching
End if
If Record= $P_{re}$ 
    Do: Reproduction process
End
If Record>> $P_{el}$ 
    Do: Elimination-Dispersal process
End

```

The larger value of parameter P_{te} indicates the longer iterations waiting for the best solution. However, this parameter cannot be too large to bring the computational burden, while the value cannot be too small to obtain the local optimal rather than the best solution. Parameters P_{re} and P_{el} are developed to decide the process of reproduction and elimination. Reproduction is used to replace the poorer bacteria from the population and improve the quality of the partial population rather than the global best. To enrich the population, elimination-dispersal process is adopted to improve the diversity of the individuals and strength the capability for the global optimization. Smaller value of P_{el} brings more randomness to the population, while larger value indicates the less diversity. Therefore, the

randomness of the individuals relies on two parameters P_{re} and P_{el} , while the termination parameter P_{te} is applied to control the redundant searching.

3.3.4. Modified Reproduction and Elimination Strategies

The reproduction strategy in BCO is originally investigated for the continuous problems according to the historic performance of individuals. However, it seems to be less effective in solving the discrete combinatorial problems, e.g. feature selection. Therefore, the reproduction strategy in typical bacterial based algorithms is modified in the proposed method for feature selection.

1). *Reproduction* To improve the quality of the population, the individuals achieving the lower classification accuracy rate are updated through the reproduction process. In BAFS, the bacteria achieving the better performance than the average (i.e., mean fitness of the population Avg_Fit referring to Eq. (3-12)) are used to replace the bacteria achieving the poorer performance (worse than the average).

$$Avg_Fit = \frac{1}{Popsize} \sum_{i=1}^{Popsize} Fitness_i \quad (3-11)$$

$$Fitness_i = \frac{FS_i}{TS_i + FS_i} \quad (3-12)$$

The fitness of the optimization method is the classification error rate, where TS and FS represent the number of testing samples that have been classified into true and false groups, respectively.

It is assumed that the objective of the optimization is to minimize the fitness function. The bacteria with the fitness value larger than Avg_Fit are replaced by bacteria with the fitness smaller than Avg_Fit . The fitness values are sorted in ascending order, and the bacteria locating at the front are more effective than the ones locating at the back. If the fitness of $(Popsize-i+1)^{th}$ bacterium are smaller than Avg_Fit , then reproduction process is operated as follows:

$$\theta(Popsize - i + 1) = \theta(i) \quad (3-13)$$

where $\theta(i)$ represents the position of i^{th} bacterium. $Popsize$ is the total number of bacteria.

2). Elimination-Dispersal It is used to guarantee the diversity and global optimization of the proposed method. As the variables in feature selection is not continuous and the problem is a combinatorial problem to select the best subsets of features from the dataset, the elimination and dispersal process for the randomness should be modified to make the method suitable for the float optimization. The matrix A is initialized with all elements zeros. If the i^{th} feature used in j^{th} bacterium, then element $A(i, j)$ equals to 1. According to Eqs. (3-7) and (3-8), the features are ranked according to their contributions to the feature subsets. The features with lower weighting indexes are replaced with unseen features (never used for evaluation, e.g., $A(i, j)=0$) according to matrix A . Assumed that the i^{th} bacterium is $\theta(i) = [f_{i1}, \dots, f_{iD}]$. The features are sorted in descending order according to weighting vector (see Eqs. (3-7) and (3-8)), then the last feature f_{iD} has the lowest weighting index in W . Therefore, the feature f_{iD} in the i^{th} bacterium will be

replaced by a new feature that has never been appeared in that bacterium, and the bacterium will be updated as follows:

$$\theta(i) = [f_{i1}, \dots, f_{iD}, f_N]. \quad (3-14)$$

where f_N is the feature that is not appeared in the i^{th} bacterium with the occurrence record $A(N, i)=0$.

3.3.5. Bacterial Algorithm based Feature Selection Method

The pseudo-code of the proposed method, abbreviated as BAFS, is given in Algorithm 3-1.

Algorithm 3-1: Pseudo-Code of BAFS for feature selection on datasets

01 **Input:** dataset for training and testing: Tr and Te ; number of features to be selected: D
02 **Initialization:** $T1, T2, T3, C, Max_iteration$;
03 **Fitness function:** develop the fitness function //classification error rate achieved by KNN
04 Calculate the fitness of all bacteria, and define $Current_iteration=0$
05 **Optimization process:**
06 **If** $Current_iteration < Max_iteration$
07 $Current_iteration = Current_iteration + 1$;
08 Obtain the P_{best} and G_{best} // P_{best} is calculated according to the pseudo-code in Figure 3-1, and G_{best} is the best position of the group
09 **For** each bacterium
10 Running: Adapt the position of bacteria using Eq.(3-2)
11 Obtain the fitness f and compare with the original fitness Fit
12 **If** $Fit < f$ (assume that the objective is to minimize)
13 Tumbling: Adapt the position of bacteria using Eq.(3-3).
14 **end if**
15 Adapt the W using Eqs. (3-7) and (3-8)
16 **End for** //end for each bacterium
17 **For** all bacteria
18 Adapt the $weight$ using Eqs. (3-9) and (3-10) and global best G_{best}
19 **If** global best is unchanged, i.e. $G_{best}^{iter} = G_{best}^{iter-1}$
20 Record=Record+1;
21 else
22 Record=0;
23 **end if**
24 **If** $Current_iteration > Max_iteration/2$ and Record= P_{te}
25 Success=1; // termination to avoid the redundant searching
26 **end if**
27 **If** $Current_iteration > Max_iteration/2$ and Record= P_{re}
28 Do: Reproduction (refer to Eqs. (3-11) to (3-13))
29 **end if**
30 **If** $Current_iteration > Max_iteration/2$ and Record>> P_{el}
31 Do: Elimination-Dispersal (refer to Eq. (3-14))
32 **end if**
33 End for //end for all bacteria
34 **end If**
35 **Output:** selected feature subsets

3.4. Parameters Estimation and Decision

In this section, three parameters (P_{te} , P_{re} , and P_{el} .) used for randomness control are studied and discussed.

3.4.1. Parameter Settings and Benchmarks

To choose the appropriate parameters for randomness control, the dataset ‘9_tumors’ is used for evaluation. All attribute values in the database have been entered as numeric values. The dataset includes 60 instances with 5726 features belonging to 9 classes, which is available online <http://www.gems-system.org/>. Since the motivation of the experiments described below is to study the choosing of parameters, the number of features expected to be selected is defined as 30 (from 5726 features). The instances are divided randomly into two sets: 75% for training and 25% for testing.

The fitness function is the classification error rate achieved by a classifier. In this study, K-Nearest Neighbor (KNN) with $K=5$ is considered as the classifier to evaluate the performance of the feature selection method. The population size of the BAFS is 50, the maximum iteration for the optimal is 300. The parameter P_t used to avoid the redundant searing is studied ranging between 20 and 120, and two other parameters for reproduction P_{re} and elimination P_{el} are varied between 5 and P_{te} .

3.4.2. Experimental Results on Parameters P_{te} , P_{re} , and P_{el}

As is illustrated previously, the smaller values of parameters P_{re} (and/or P_{el}) indicate the more frequent reproduction (and /or elimination) to update the population by replacing the poorer individuals (and/or renewing the population by adding the randomness). Results are given in the following tables using different values of three controlling parameters. All results are obtained over 30 runs. ‘*Accu*’ in Tables represents the average classification accuracy, and ‘*Iter*’ indicates the real iteration times used by BAFS.

Table 3-1 shows minimum, maximum, average and standard variance of iterations by BAFS to achieve the associated classification accuracy (%) with various values of control parameters. Table 3-2 to Table 3-5 give the detailed results in Table 3-1. Figure 3-4 shows the results given in Table 3-2 to Table 3-5. The values of average classification accuracy rate higher than 95% are highlighted and the highest value of accuracy rate among all points is texted.

Table 3-1 The range of classification accuracy rate and iterations using varying values of termination parameter P_{te}

Parameters		Maximum	Minimum	Average	Std. Var
$P_{te} = 20$	Accu	97.7800%	71.1100%	86.6598%	4.7054%
	Iter	42.6000	30.6000	36.1612	3.0192
$P_{te} = 40$	Accu	97.7800%	72.2200%	88.2306%	1.3857%
	Iter	76.7667	61.9333	68.3844	3.3058
$P_{te} = 80$	Accu	97.7800%	75.5600%	88.8343%	2.4382%
	Iter	170.8000	95.2000	125.5906	17.4291
$P_{te} = 120$	Accu	98.8900%	72.2200%	89.7919%	5.6077%
	Iter	242.8000	138.0000	175.3129	18.3019

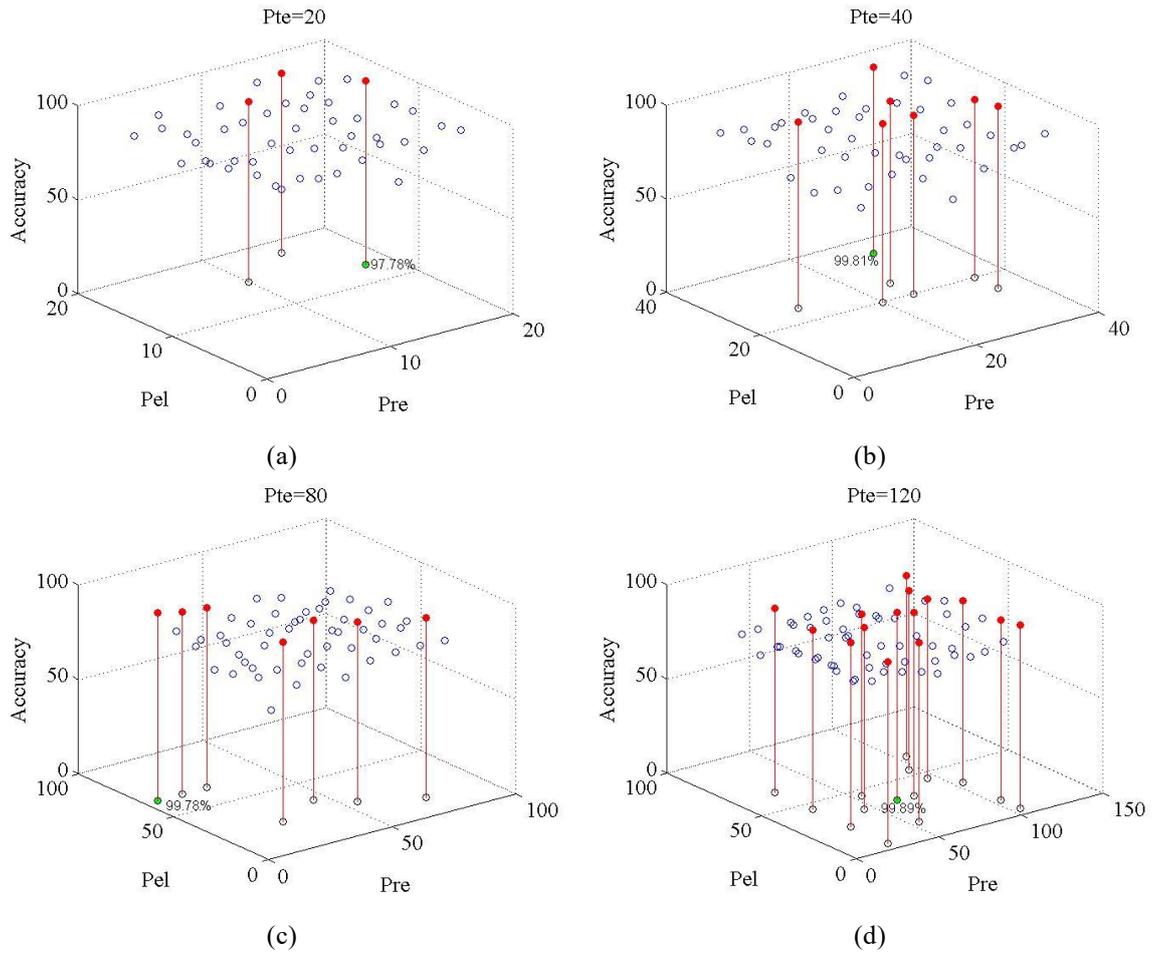


Figure 3-4 Average classification accuracy achieved with varying P_{te} , P_{re} , and P_{el}

According to Figure 3-4, the larger value of P_{te} contributes to the higher probabilities (with more points are highlighted) for obtaining the good feature subsets to achieve the higher classification accuracy rate. Even so, we cannot ignore the computational complexity associated with large P_{te} . As shown in Table 3-1, longer computation time (i.e. represented by iteration times) is needed to search for the best solutions though higher average classification accuracy rate is obtained with the larger value of termination parameter P_{te} . Among four cases in consideration, the highest classification accuracy rate and average classification accuracy rates are achieved when the termination parameter P_{te} is 120, but the computational cost for

the best is also the highest. Additionally, larger iterations cannot guarantee the best solution. Though the average and the maximum of classification accuracy rate are achieved when parameter $P_{te}=120$, the computational cost consumed by searching process is rather expensive and over twice of the case with the termination parameter $P_{te}=40$. Actually, the minor priority in classification accuracy rate is realized at the expense of expensive computational cost.

The effectiveness of optimization method relies not only on termination parameter P_{te} , the parameters P_{re} and P_{el} are also playing the essential roles for updating the population. Since the strategies like reproduction, elimination are eligible for improving the quality of the population, these two parameters (e.g. P_{re} and P_{el}) are employed to control the frequency of reproduction and elimination process. The classification accuracy rate achieved exceeding 95% are highlighted in bold.

Table 3-2 The range of classification accuracy rate and iterations using varying values of termination parameter P_{re}

		$P_{el}=3$	$P_{el}=5$	$P_{el}=8$	$P_{el}=10$	$P_{el}=13$	$P_{el}=15$	$P_{el}=18$
$P_{re}=3$	Acc	91.11	92.22	88.89	87.22	80.00	94.44	83.33
	Iter	34.30	34.30	41.8	34.5	35.1	34.3	37.1
$P_{re}=5$	Acc	91.67	81.11	88.89	85.00	77.78	87.78	91.11
	Iter	42.6	40.0	39.3	38.5	37.5	34.3	39.4
$P_{re}=8$	Acc	88.89	81.67	90.00	88.78	93.33	85.00	71.11
	Iter	34.2	33.7	31.2	37.9	37.1	39.4	33.7
$P_{re}=10$	Acc	92.56	94.44	87.22	93.33	94.44	96.33	87.22
	Iter	33.2	31.1	35.6	31.7	33.4	36.8	38.8
$P_{re}=13$	Acc	75.56	91.11	88.89	92.22	92.22	90.00	94.
	Iter	37.1	31.2	32.0	35.7	36.5	33.4	38.8
$P_{re}=15$	Acc	88.89	88.89	84.44	90.00	91.67	91.11	95.78
	Iter	37.1	34.0	33.6	38.1	37.1	36.3	41.8
$P_{re}=18$	Acc	94.44	92.22	93.33	92.22	97.78	94.44	86.67
	Iter	37.8	35.9	30.6	35.7	38.8	41.4	38.2

Table 3-3 Average classification accuracy rate and iterations for the optimal when $P_{te}=40$

		$P_{el}=5$	$P_{el}=10$	$P_{el}=15$	$P_{el}=20$	$P_{el}=25$	$P_{el}=30$	$P_{el}=35$
$P_{re}=5$	Acc	81.11	84.44	77.78	80.00	92.22	94.44	86.67
	Iter	69.67	63.83	72.47	65.40	67.23	65.23	61.93
$P_{re}=10$	Acc	94.44	82.22	92.22	90.00	85.56	91.11	77.78
	Iter	70.97	67.07	71.87	67.03	65.37	71.43	72.47
$P_{re}=15$	Acc	87.78	94.44	90.00	92.22	91.11	94.44	83.33
	Iter	72.60	66.63	70.20	66.67	67.10	65.53	70.03
$P_{re}=20$	Acc	72.22	88.89	82.22	95.56	93.33	91.11	81.11
	Iter	63.93	62.53	74.73	63.93	64.67	66.80	68.97
$P_{re}=25$	Acc	84.44	90.00	84.44	95.56	97.78	87.78	88.89
	Iter	68.37	76.77	66.50	70.37	65.60	69.10	67.87
$P_{re}=30$	Acc	91.11	92.22	92.22	83.33	88.89	86.67	78.89
	Iter	68.47	68.87	65.63	67.80	67.20	69.53	74.37
$P_{re}=35$	Acc	94.44	82.22	97.78	95.56	87.78	94.44	91.11
	Iter	67.47	68.43	69.40	66.86	75.63	70.23	70.07

Table 3-4 Average classification accuracy rate and iterations for the optimal when $P_{te}=80$

		$P_{el}=5$	$P_{el}=15$	$P_{el}=25$	$P_{el}=35$	$P_{el}=45$	$P_{el}=55$	$P_{el}=65$
$P_{re}=5$	Acc	75.56	93.33	85.56	83.33	91.11	94.44	97.78
	Iter	114.2	113.2	112.2	134.6	115.2	115.6	143.6
$P_{re}=15$	Acc	85.56	88.89	80.00	87.78	93.33	86.67	96.67
	Iter	130.6	127.4	118.2	107.4	112.2	132.6	121.8
$P_{re}=25$	Acc	91.11	88.89	95.56	88.89	75.56	81.11	95.56
	Iter	115.2	167.4	160.8	106.2	101.4	127.6	132.6
$P_{re}=35$	Acc	82.22	94.44	85.56	94.44	87.78	87.78	86.67
	Iter	113.6	154.8	124.0	162.0	148.2	116.8	127.8
$P_{re}=45$	Acc	87.78	90.00	94.44	95.56	90.00	90.00	93.33
	Iter	115.0	120.0	107.6	123.0	104.8	106.0	117.6
$P_{re}=55$	Acc	88.89	91.11	95.56	85.56	93.33	83.33	90.00
	Iter	132.4	123.4	124.4	105.8	121.6	124.8	114.4
$P_{re}=65$	Acc	88.89	93.33	91.11	83.33	84.44	88.89	78.89
	Iter	134.0	170.8	119.2	152.2	104.6	105.4	129.8
$P_{re}=75$	Acc	87.78	95.56	88.89	94.44	85.56	86.67	86.67
	Iter	144.2	120.4	148.4	117.2	167.2	151.4	145.0

Table 3-2 shows that the proposed BAFS could select the considerable subsets to achieve the average classification accuracy rate reaching to 97.78% within comparative small iterations (no more than 40 iteration time) when the two parameters are: $P_{re}=18$ and $P_{el}=13$. However, smaller parameters P_{el} (e.g. $P_{el} < 10$) associated with frequently utilizing the elimination for diversity seem to be less effective for improving the searching capability. As parameter P_{el} increases to

around 15 and parameter P_{re} increases to around 10, the average classification accuracy rate is achieved mostly around or over the average level.

Table 3-5 Average classification accuracy rate and iterations for the optimal when $P_{te}=120$

		$P_{el}=5$	$P_{el}=15$	$P_{el}=25$	$P_{el}=35$	$P_{el}=45$	$P_{el}=55$	$P_{el}=65$
$P_{re}=5$	<i>Acc</i>	92.22	92.22	94.44	92.22	91.11	82.22	88.89
	<i>Iter</i>	221.0	177.4	176.6	162.4	184.6	157.2	139.6
$P_{re}=15$	<i>Acc</i>	88.89	84.44	87.78	86.67	86.67	84.44	88.89
	<i>Iter</i>	145.4	164.8	194.8	204.6	184.0	183.4	159.4
$P_{re}=25$	<i>Acc</i>	96.67	88.89	97.78	81.11	95.56	93.33	97.78
	<i>Iter</i>	202.6	169.8	156.6	212.2	161.0	176.2	172.4
$P_{re}=35$	<i>Acc</i>	94.44	88.89	88.89	94.44	88.89	90.00	87.78
	<i>Iter</i>	191.4	149.0	145.8	154.4	189.4	160.4	166.0
$P_{re}=45$	<i>Acc</i>	86.67	82.22	77.78	96.67	86.67	91.11	88.89
	<i>Iter</i>	186.2	207.8	174.8	177.8	182.8	166.4	185.2
$P_{re}=55$	<i>Acc</i>	83.33	95.56	88.89	84.44	96.67	84.44	90.00
	<i>Iter</i>	197.8	156.0	186.4	180.4	166.4	138.0	212.4
$P_{re}=65$	<i>Acc</i>	91.11	83.33	92.22	98.89	92.22	90.00	91.11
	<i>Iter</i>	200.8	202.4	161.0	183.6	191.6	176.2	150.4
$P_{re}=75$	<i>Acc</i>	87.78	87.78	84.44	97.78	90.00	86.67	86.67
	<i>Iter</i>	193.4	198.6	151.8	170.6	184.2	172.2	162.2
$P_{re}=95$	<i>Acc</i>	91.11	81.11	90.00	92.22	95.56	95.56	92.22
	<i>Iter</i>	161.0	172.0	168.2	159.0	167.2	148.8	189.2
$P_{re}=105$	<i>Acc</i>	97.78	95.56	92.22	96.67	92.22	87.78	96.67
	<i>Iter</i>	188.8	152.8	172.2	179.0	164.0	193.0	182.2

Similarly, from Table 3-4 and Table 3-5, we can find that the smaller parameters (e.g. $P_{re} < 10$, $P_{el} < 10$) associated with the earlier process of reproduction and/or frequent randomness from elimination for diversity might contribute to poorer performance for feature selection. But the value of parameter P_{el} cannot be too large. When termination parameter P_{te} is ranging between 80 and 120, the accuracy rate is decreasing when the parameter P_{el} is larger than 65, so the cases with parameter P_{el} larger than 65 are not considered in the study.

Therefore, the iterations for the optimal by BAFS do not need to be too large. Even though the predefined maximum iterations for the optimal is large, the

termination parameter P_{te} could be used to avoid the redundant searching for the best solution. It is appropriate to assign the termination parameter P_{te} around 40, employ the reproduction control parameter P_{re} larger than 20 and smaller than P_{te} , and apply the elimination control parameter P_{re} larger than 15 but smaller than 65 and P_{te} .

3.5. Benchmark Experiments

To test the performance of the proposed feature selection method, typical bacterial based algorithms are employed and compared for feature selection on several frequently used datasets.

3.5.1. Comparison Techniques and Datasets

In this section, the effectiveness of the proposed BAFS will be tested on datasets in comparison with four other bacterial algorithms: BFO [141], BFO with linear chemotaxis step size (BFOLDC) [182], BFO with nonlinear chemotaxis step size (BFONDC) [182], and BCO [183].

Features used in the currently available mechanical analysis datasets are small in scale and easy to implement. In our study, the proposed optimization method is to select the sensor chains from the networks with more than 100 sensors (or even more than 1000 in real application). Therefore, the datasets—frequently used for feature selection with high dimensional variables (more than 2000 features) — are provided together with mechanical analysis dataset to test the effectiveness of the

proposed bacterial based method. The first dataset is available online from:

<http://cilab.ujn.edu.cn/datasets.htm>. The last nine datasets are from:

<http://www.gems-system.org/>. Table 3-6 gives the detailed information of datasets.

Table 3-6 Datasets for feature selection

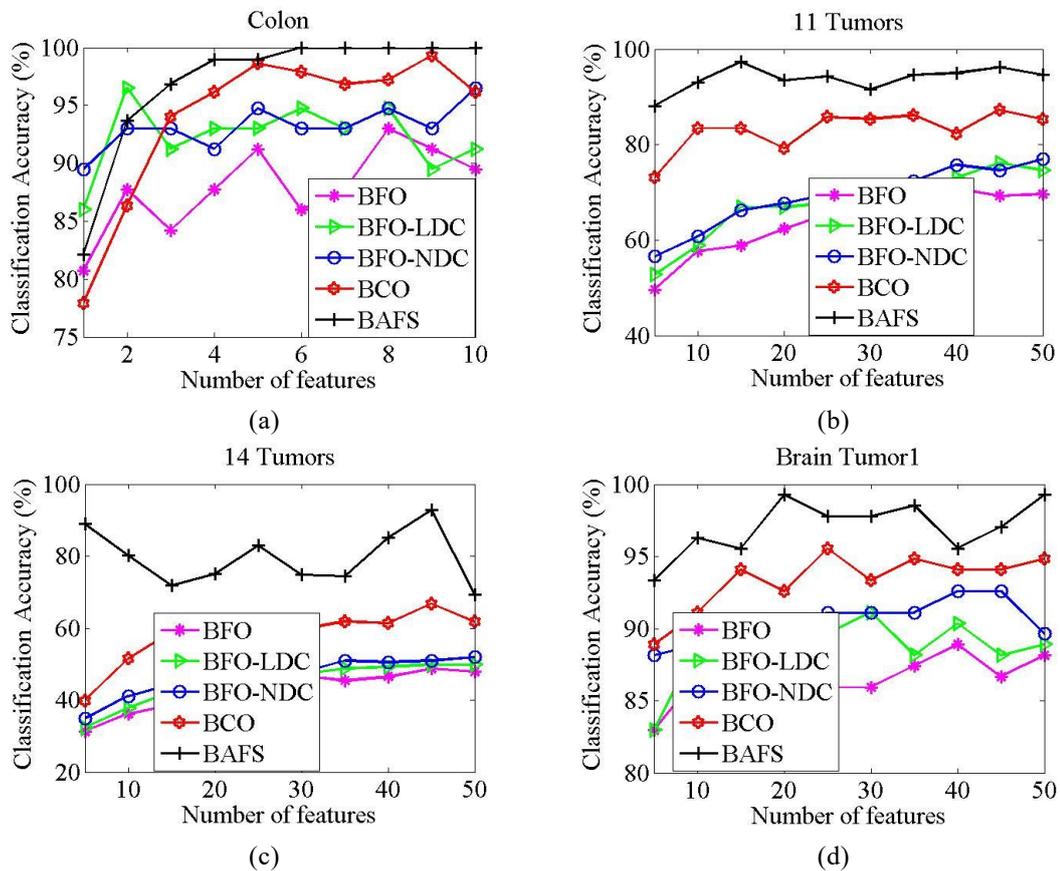
Dataset	Features	class	instances
Colon	2000	2	62
11_Tumors	12,533	11	174
14_Tumors	15,009	26	308
Brain_Tumor1	5920	5	90
Brain_Tumor2	10367	4	50
SRBCT	2309	4	83
Leukemia1	5328	3	72
Leukemia2	11225	3	72
Prostate_Tumor	10509	2	102
DLBCL	5470	2	77

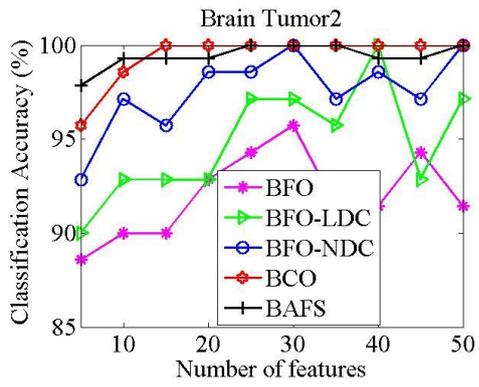
The population size in all feature selection methods are defined as $NP = 50$.

The iteration time for BCO and BAFS is 200. For remaining three bacterial based algorithms, the parameters are defined as: chemotaxis iteration $N_c=50$, reproduction iteration $N_{re}=5$, and elimination iteration $N_{el}=2$. The parameters for chemotaxis step strategies in BFOLDC and BFONDC are assigned as: $C_{min} = 1$ and $C_{max} = 5$. According to the study on the three randomness control parameters in BABOAFS, they are defined as: $P_t=40$, $P_{re}=25$, and $P_{el}=20$. For each dataset, the instance are divided randomly two sets: 70% for training and 30% for testing. The fitness for all the optimization method is the classification error rate evaluated by KNN with $K=5$. The number of features expected to selection cannot exceed 10 for Colon and 50 for remaining high dimensional datasets.

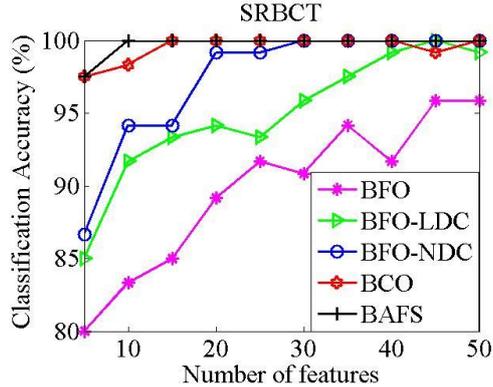
3.5.2. Experimental Results and Analysis

Figure 3-5 shows the average classification accuracies across 30 repeated runs for varying subset size on datasets. The experimental results show that the proposed BAFS outperforms all other bacterial based methods to select the feature subsets for achieving higher classification accuracy rate, and BCO performs better than BFO, BFO-LDC and BFO-NDC. Meanwhile, BFO-NDC shows the minor superior to BFO-LDC in some cases like SRBCT, Leukemia 2, and DLBCL. Though, the improvement of chemotaxis step strategies on BFO have improved the effectiveness of the original BFO, the exchange strategies in BCO seems to be more eligible in comparison to chemotaxis strategies based algorithms (i.e. BFO-LDC and BFO-NDC).

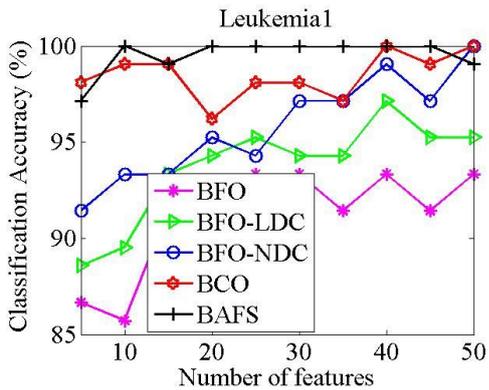




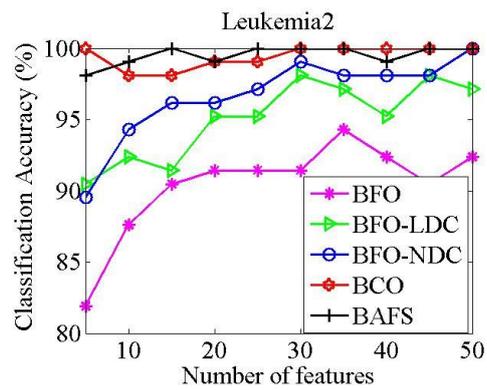
(e)



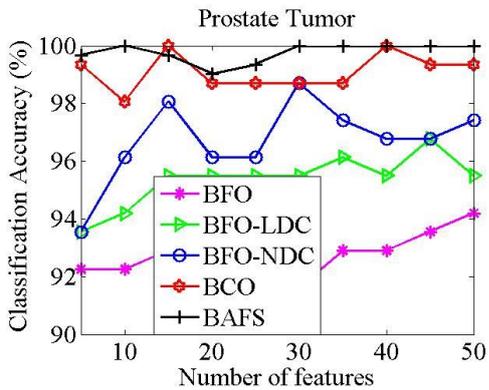
(f)



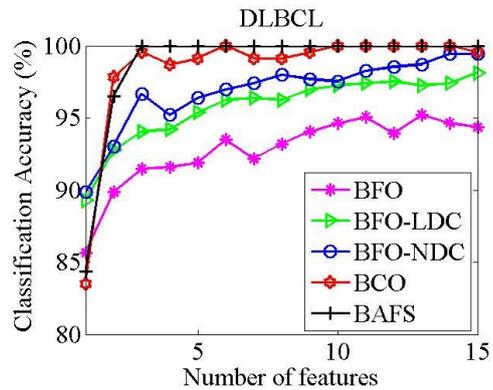
(g)



(h)



(i)



(j)

Figure 3-5 Average classification accuracy across 30 repeated runs for varying subset size on datasets

Among all bacterial based algorithms, BFO, BFO-LDC and BFO-NDC have the largest iteration time for the best solutions, i.e. 500 iterations. The iteration time used in BCO for searching is 200. While BAFS has the smallest iteration, i.e. no more than the maximum iterations (i.e. around 100 and no more than 200). Thus, the computational cost consumed by BAFS is even smaller than BCO, while BFO

and BFO based algorithms spend larger computational time for the best solution. Figure 3-6 shows that BAFS and BCO spend less computational time, and BFO, BFO-LDC, BFO-NDC consume more than twice of computational time for searching the best solutions.

Though randomness mechanism is employed in BCO, the randomness embedded in BCO is realized only by tumbling process. In the proposed BAFS, the randomness is not only manifested in tumbling strategy, but also applied by elimination to replace the poorer features with unseen features from recording matrix. The importance of features are evaluated by the probability of their contribution to the subsets achieving higher classification performance. As a result, the capability of BAFS is obviously superior to BCO in most cases with lower computational time.

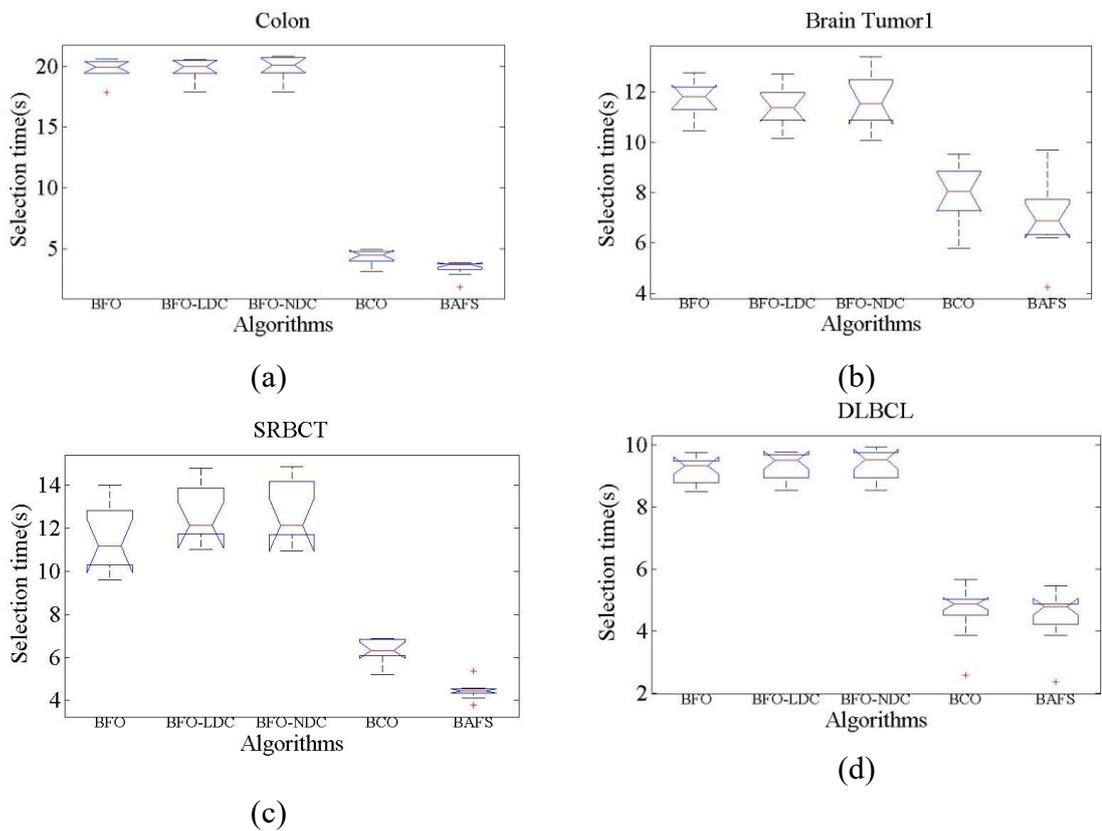


Figure 3-6 Computational time spent by feature selection methods across the 30 runs

3.6. Discussion and Conclusion

The proposed bacterial algorithm based on randomness control mechanisms was investigated for feature selection in classification. A premature parameter is defined to avoid the redundant searching for the best results and proved to be effective in reducing the computational time for selecting the feature subsets from datasets. Meanwhile, two additional parameters are applied to decision the process of reproduction and elimination. Experimental studies on choosing the parameters indicate that the frequency of elimination for diversity cannot be too high (waiting for over 15 times), and it is appropriate to apply the reproduction mechanism when the best position of the population is unchangeable over 15-20 times. The learning mechanism in BCO is more effective in comparison to chemotaxis step strategies for feature selection in classification.

This higher capability of the proposed method for feature selection indicates that the proposed method BAFS makes it more suitable for the selecting the most appropriate sensor chains while constructing virtual beams for fault localization. Therefore, in the further study, the proposed BAFS will be applied to the sensors selection to construct the virtual beams for fault diagnosis of complex structures.

Chapter 4. Feature Characterization and Fault Detection

In this Chapter, signal preprocessing and fault detection methods are presented. Signal preprocessing consists of feature characterization and feature selection. To obtain the most informative features from the available feature representatives, an unsupervised clustering method is developed for feature selection. The combination of statistical tests and fault indicator-based adaptive threshold is presented for fault detection. In what follows, they will be presented in details.

4.1. Signal Preprocessing

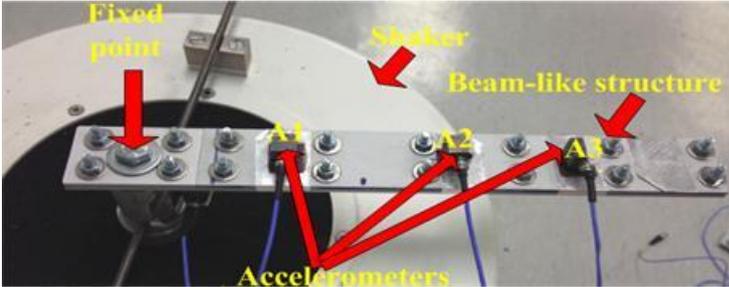
Since the data are firstly transformed into the form of features based on segmentation of the time series. Signal preprocessing in this stage mainly consists of feature characterization and feature selection.

4.1.1. Data Acquisition

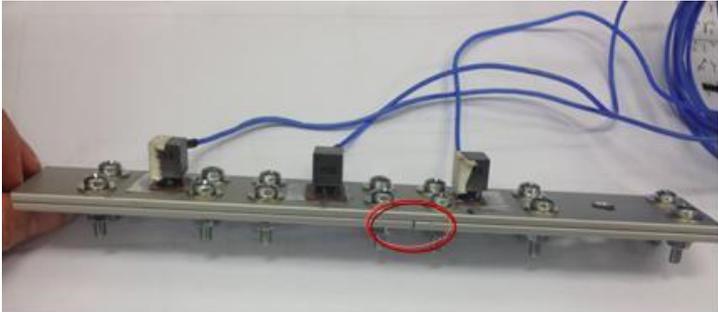
In this thesis, a beam-like structure, served as the benchmark structure, is studied for fault diagnosis. Thus, the data measured from this system is provided for feature selection.

The beam-like structure shown in Figure 4-1 consists of the two layers bolted steel panels of square size (length 40cm* wide 5cm*high 1cm). Breathing-like cracks or loosening-bolt can be easily created in one layer at any known positions

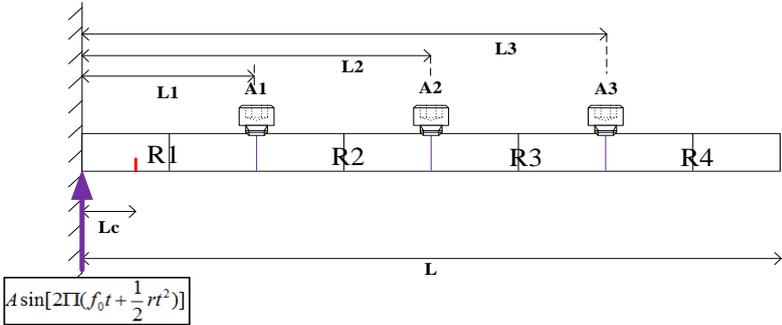
for benchmark studying (see Figure 4-1 (b)). As shown in Figure 4-1 (c), faults (e.g. cracks) at different positions can be denoted by $R1$, $R2$, and $R3$, and referred to as *root crack*, *middle crack*, and *top crack*, respectively. Three accelerometers are placed regularly on the beam at shown positions denoted by $A1$, $A2$, and $A3$.



(a). Beam-like structure system



(b). Crack on beam-like structure



(c) Illustration for the beam-like structure

Figure 4-1 The designed beam-like structure

The system is firstly measured without any cracks or loosening bolts as a reference signal and then is excited with known faults by using the same input. The first measured time series data acts as the reference signal which can usually be taken in the health monitoring record of a structural system. The second

measurement time series data under fault situations is used for fault feature characterization and diagnosis. The input excitation used in the experimental testing is not necessarily any specific signal. As a benchmark study, different inputs have been tested. Actually, the study focuses on the changes of vibration transmission, and thus the excitation magnitude of the input has no effect on the feature characterization and fault diagnosis. As an example, in this thesis, the input is a sweeping-frequency excitation from 20 Hz to 200 Hz with the exciting intensity 2g, the sample frequency 8192Hz, and velocity 4 oct/min.

4.1.2. Feature Characterization

In the process of feature construction, the raw time series data of long time dimension will be converted into features of a lower time dimension. Considering that the frequency domain features may not be suitable for non-stationary signals, and time-frequency domain features have more computational complexity, in our study, time domain features are adopted in this thesis. Those 16 features are: maximum value(f1), Minimum value (f2), Average absolute value (f3), Peak to peak (f4), Variance(f5), Standard deviation (f6), Skewness (f7), Kurtosis (f8), Root mean square frequency (f9), Crest factor (f10), Clearance factor (f11), Impulse factor (f12), Peak to peak (f13), Peak-magnitude-to-RMS ratio (f14), Energy Ratio (f15), energy operator (f16). The detail formation can refer to [184]. Time domain features are presented in Table 4-1.

Table 4-1 Typical time domain statistical features

(1) Maximum value: $Max\{x(i)\}$	(2) Minimum value: $Min\{x(i)\}$
(3) Average absolute value: $\frac{1}{n} \sum_{i=1}^n x(i) $	(4) Peak to peak: $Max\{x(i)\} - Min\{x(i)\}$
(5) Variance: $\frac{1}{n} \sum_{i=1}^n (x(i) - \bar{x})^2$	(6) Standard deviation: $\sqrt{\frac{1}{n} \sum_{i=1}^n (x(i) - \bar{x})^2}$
(7) Skewness: $\frac{1}{n} \sum_{i=1}^n (x(i) - \bar{x})^3 / (\frac{1}{n-1} \sum_{i=1}^n (x(i) - \bar{x})^2)^{3/2}$	(8) Kurtosis $\frac{1}{n} \sum_{i=1}^n (x(i) - \bar{x})^4 / (\frac{1}{n} \sum_{i=1}^n (x(i) - \bar{x})^2)^2$
(9) Root mean square frequency: $\sqrt{\sum_{i=1}^K f(k)^2 \times X(k) / \sum_{i=1}^K X(k)}$	(10) Crest Factor (or Crest RMS) $Max\{x(i)\} / \sqrt{\frac{1}{n} \sum_{i=1}^n x(i)^2}$
(11) Clearance factor: $Max\{x(i)\} / \frac{1}{n} \sum_{i=1}^n x(i)^2$	(12) Impulse factor: $Max(x(i)) / \frac{1}{n} \sum_{i=1}^n x(i) $
(13) Peak to peak: $\sqrt{\frac{1}{n} \sum_{i=1}^n x(i)^2} / \frac{1}{n} \sum_{i=1}^n x(i) $	(14) Peak-magnitude to RMS ratio: $\ X\ _{\infty} / \sqrt{\frac{1}{n} \sum_{i=1}^n X_i ^2}$
(15) Energy Ratio $\sqrt{\frac{1}{n} \sum_{i=1}^n (d(i) - \bar{d})^2} / \sqrt{\frac{1}{n} \sum_{i=1}^n (x(i) - \bar{x})^2}$	(16) Energy Operator $\frac{1}{n} \sum_{i=1}^n (\Delta x(i) - \Delta \bar{x})^4 / (\frac{1}{n} \sum_{i=1}^n (\Delta x(i) - \Delta \bar{x})^2)^2$
<p>where X is the frequency spectrum of $x(i)$, while $X(k)$ is the k^{th} measure of the frequency spectrum of $x(i)$, $f(k)$ is the frequency value of the k^{th} spectrum line, $k=1, \dots, K$. $\Delta x(i)$ is the obtained piecewise. n is the total number of data points in time record. \bar{x} represents the mean of $x(i)$. $d(i)$ indicates the “diff” function which calculates differences between adjacent elements of $x(i)$, and $d_m(i)$ is the m^{th} time record. \bar{d} is the mean of $d(i)$. More exactly, $diff(x)$ returns a vector, one element shorter than X of differences between adjacent elements: $[x(2) - x(1), x(3) - x(2), \dots, x(n) - x(n-1)]$</p>	

4.1.3. Selection of Features for Signal Analysis

Sixteen predefined time domain features for each sensor are the candidates for

feature representatives. Therefore, 48 features in total for three sensors are provided for feature selection. Each feature is regarded as an independent individual for fault clustering and feature subsets are used for fault classification. The time interval for segmentation is 1 second. 40 groups of signals with each ten groups of signal under the condition of normal, root crack, middle crack, and top crack are provided for clustering:

- ◆ *Normal condition*: ten group of signals measured from fault free system;
- ◆ *Root crack* (R1 region): ten group of signals measured when crack is near closed to fixed boundary ($L_c/L=0.0625$);
- ◆ *Middle crack* (R2 region): ten group of signals measured when the crack is in the middle of the beam ($L_c/L=0.35$);
- ◆ *Top crack* (R3 region): ten group of signals measured when the crack locates at the top of the beam ($L_c/L=0.65$).

The clustering methods are capable of identifying the underlying features of the data, which provide useful algorithms for data mining, computer vision, pattern recognition, document clustering, etc. Among various clustering techniques, unsupervised fuzzy clustering methods have been widely developed since the results obtained include the association degree between the objective and clusters rather than the complete belonging of the objective to one group in hard clustering methods like K-means clustering. The fuzzy clustering techniques favor contiguous clusters in time and enable to detect changes in the hidden structure of multivariate time-

series, which is very useful in fault diagnosis with time series data analysis. Even so, most fuzzy clustering methods are great dependent on the initialization and less robust to the global optimum. Particle Swarm Optimization (PSO) [185] is a widely used stochastic global optimization method. In our study, an improved version of fuzzy c-means clustering, i.e. combined Particle Swarm Optimization (PSO) and fuzzy c-means clustering method (FCM), is proposed and applied to feature selection. In the combined fuzzy clustering method, the random characteristic in the PSO will be integrated with the FCM to avoid to be trapped into the local minima. In what follows, the main principle of the relevant methods will be illustrated.

Fuzzy C-means clustering (FCM). The clustering method is to classify the data from the observations into groups depending on the underlying structure. Assume that the single observed data vector consists of n variables: $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]$, $x_i \in R^n$.

Therefore, when there are N observations, it could be represented as:

$$X = \{x_i \mid i = 1, 2, \dots, N\} = \begin{bmatrix} x_{11}, x_{12}, \dots, x_{1n} \\ x_{21}, x_{22}, \dots, x_{2n} \\ \vdots \quad \vdots \quad \quad \quad \vdots \\ x_{N1}, x_{N2}, \dots, x_{Nn} \end{bmatrix} \quad (4-1)$$

For supervised clustering techniques, clustering number is predefined and denoted as c . For fuzzy clustering methods, the partition matrix consists of the probabilities of belonging to predefined number of group, which could be represented as:

$$U = \{u_{ij} \mid i = 1, \dots, N, j = 1, \dots, c\} = \begin{bmatrix} u_{11}, u_{12}, \dots, u_{1c} \\ u_{21}, u_{22}, \dots, u_{2c} \\ \vdots \quad \vdots \quad \quad \vdots \\ u_{N1}, u_{N2}, \dots, u_{Nc} \end{bmatrix} \quad (4-2)$$

$$\sum_{j=1}^c u_{ij} = 1, \quad i = 1, 2, \dots, N$$

where the membership levels are ranging from 0 to 1, i.e. $u_{ij} \in [0,1]$. For hard clustering methods, the membership level is either 0 or 1, i.e. $u_{ij} \in \{0,1\}$. The fuzzy C-means clustering method is a well-known fuzzy clustering technique, and it can be formulated as [186]:

$$\min \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m D_{ijA}^2 \quad (4-3)$$

$$v_i = \frac{\sum_{j=1}^N u_{ij}^m x_j}{\sum_{j=1}^N u_{ij}^m}, \quad i = 1, 2, \dots, c \quad (4-4)$$

$$D_{ijA}^2 = \|x_j - v_i\|_A^2 = (x_j - v_i)^T A (x_j - v_i) \quad (4-5)$$

$$\sum_{j=1}^c u_{ij} = 1, \quad u_{ij} \in [0,1], \quad i = 1, 2, \dots, N, \quad u_{ij} \in \{0,1\} \quad (4-6)$$

In the FCM, the matrix A is $n \times n$ diagonal matrix and could be formulated as the common form, i.e. $A = I$ or consists of variances of X , i.e.

$$A = \begin{bmatrix} (1/\sigma)^2 & 0 & \dots & 0 \\ 0 & (1/\sigma)^2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (1/\sigma)^2 \end{bmatrix}. \quad \text{The membership level is either 0 or 1, i.e. } u_{ij} \in \{0,1\}.$$

The membership degree u_{ij} is ranging from 0 to 1, i.e. $u_{ij} \in [0,1]$. It is rather common that $m=1$ or 2.

Particle Swarm Optimization (PSO). Inspired by bird flock or fish school, the Particle Swarm Optimization (PSO) [185] is known for its simplicity and

considerable performance on optimization problems. In the PSO, the variables in an optimization problem are represented by particle position. For each particle, it can be represented as: $x_i = [x_{i_1}, x_{i_2}, \dots, x_{i_D}]$ in the D -dimensional space. In N -observations fuzzy clustering optimization, each particle is represented as $N \times c$ matrix with the assumption that the number of cluster is n . Here, the column of matrix is $x_{i_k} = [x_{ik_1}, x_{ik_2}, \dots, x_{ikN}]^T$. The position of each particle in next moment is determined by its velocity, the local best (the best fitness value of that particle in history) as well as the global best (the best fitness value of all the particles).

In each time step t , the particles are updated according to the following equations [185]:

$$x_i(t+1) = x_i(t) + v_i(t) \quad (4-7)$$

$$v_i(t+1) = W_{inert}v_i(t) + c_1(pb_{est_i} - x_i(t)) + c_2(gb_{est} - x_i(t)) \quad (4-8)$$

where W_{inert} is an inertia weight, which employs the linear strategy in the optimization process, i.e. $W_{inert} = W_{end} + (W_{start} - W_{end}) \times (iter_{max} - iter) / iter_{max}$. Parameters c_1 and c_2 are acceleration constants to control the step distance of particles.

Combination of PSO and FCM (PSO-FCM). To overcome the defects of the FCM, the improved method PSO has been combined with the FCM (called PSO-FCM) for feature selection. The main idea is to take advantage of the random and population-based initialization and learning mechanisms of PSO to overcome the disadvantages of FCM (i.e., less robust for the global optimum). Moreover, the combination

method could also keep the priority of the fuzzy clustering methods such that the objects are not ‘forced’ to assign to the predefined clusters, which is very important to our further application where the prior knowledge of the system is unknown.

The pseudo-code of PSO-FCM has been given in Algorithm 4-1.

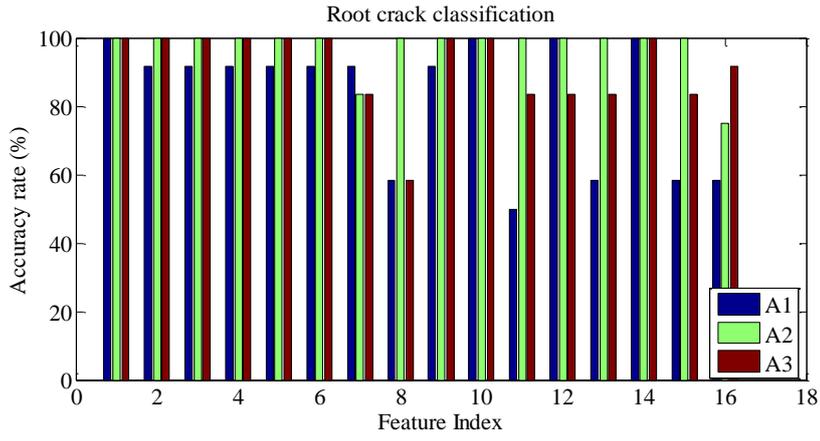
Algorithm 4-1: Pseudo-Code of PSO-FCM

```

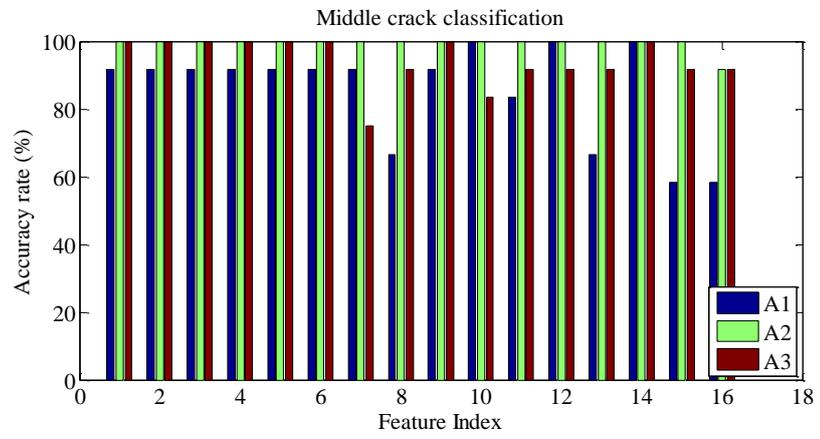
01  Input: extracted features transformed by vibration signals measured when the system is
      in varying health conditions.
02  Initialization: Population size  $Pop$ , the parameters  $c_1$  and  $c_2$ ,  $Max\_iteration$ , the
      inertia weight of velocity  $w_0$ 
03  Population initialization:
04  Particles in group 1: Randomly generate a matrix like Eq (4-1) and satisfy the
      constraints defined in Eq. (4-6) as the initial position of the particle.
05  Particles in group 2: The position of particle is initialized using the  $U$  partition matrix
      obtained by FCM.
06  Evaluation: Compute the fitness value using Eqs (4-3)-(4-5), and find the local best (the
      best fitness value of that particle in history) as well as global best (the best fitness value
      of all the particles) of the particles by fitness comparison.
07  Updating: Adapt the position matrix of the particles
08  Optimization process
09  While  $Max\_iteration$  is not met
10  Update the position using Eq. (4-7) and satisfy the constraints defined in Eq. (4-6);
E   If the current iteration =  $Max\_iteration/2$ 
12   Do: Half poorer particles would be optimized by FCM;
13   end
14  Evaluation: Compute the fitness value using Eqs (4-3) to (4-5), and obtain the local best
      ( $pbest$ ) as well as global best ( $gbest$ ) of the particles by fitness comparison
15  end while
16  Output: the possibility of features to be selected and corresponding accuracy rate
      achieved by each feature.

```

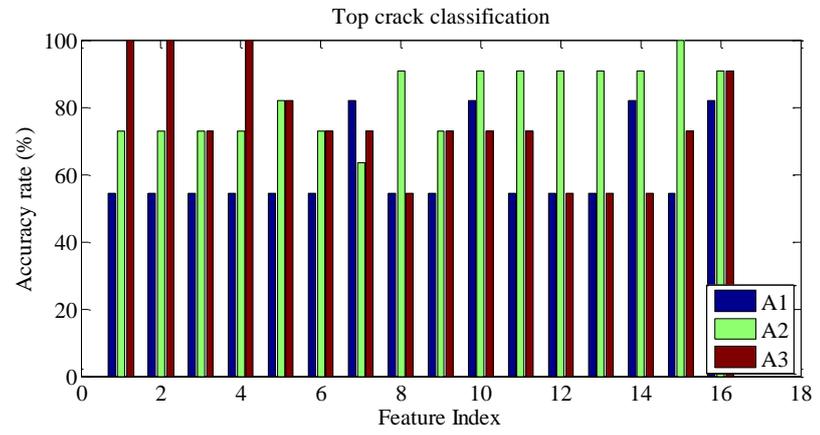
To select the most informative features, the PSO-FCM is applied to clustering the four different healthy conditions of the beam-like structure. The results are shown in Figure 4-2.



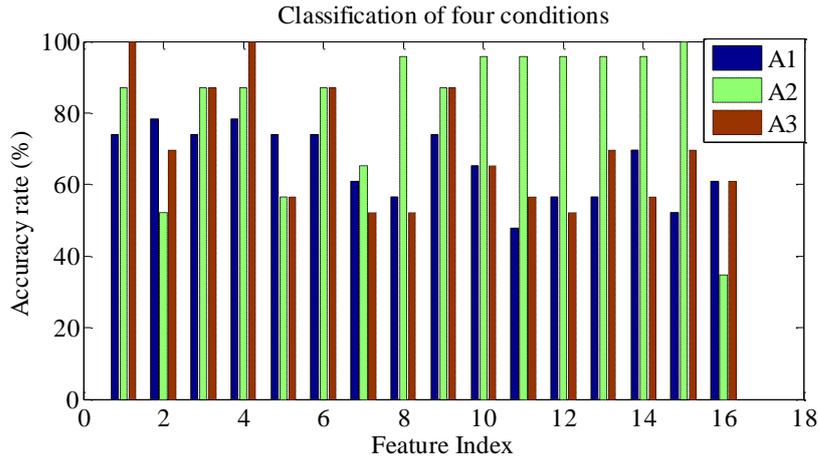
(a)



(b)

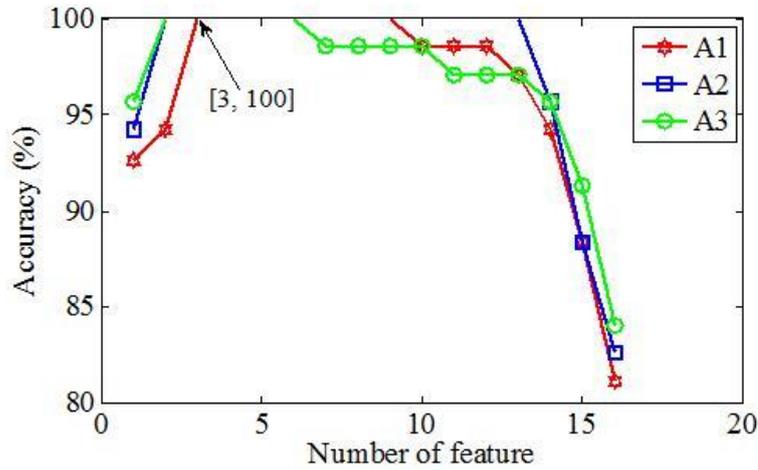


(c)



(d)

Figure 4-2 Accuracy rate achieved by clustering method using the data provided by three sensors with individual time feature as signal representative



Feature No.	3, 8, 10	4, 8, 15	3, 4, 12	4, 10, 15	Total
Times	9	8	6	7	30

Figure 4-3 Classification accuracy achieved by three sensors (i.e. A1, A2 and A3)

From Figure 4-2 (a), the energy ratio (f15) is the most sensitive feature for fault classification of four health condition in terms of sensor A2, while maximum value (f1) and peak to peak (f4) are most effective features in terms of sensor A3, and peak to peak (f4) is also the most informative features for sensor A1. Except above-mentioned features, crest factor (f10) and peak-magnitude-to-RMS ratio (f14) also show better the performance for fault classification according to three other

classification cases in Figure 4-2 (b) to (c). However, the maximum value (f1), peak to peak (f4) and the peak-magnitude-to-RMS ratio (f14) are similar in feature characterization.

According to Figure 4-2 (d), the peak to peak (f4) is the most informative feature for sensor A1, the energy ratio (f15) is the most sensitive feature for fault classification of four conditions in terms of sensor A2, while the maximum value (f1) and peak to peak (f4) are most effective features for sensor A3. Considering that the feature characterization of the maximum value (f1) and peak to peak (f4) are similar, one of them will be selected to avoid the feature redundancy. Combined with the results from sensor A1, the peak to peak (f4) is selected.

Even so, Figure 4-3 shows the classification accuracy achieved by the KNN (K=4) [127] using the optimization method, indicating that the number of features should be 3 to balance the best performance of the fault classification as well as the computational complexity. The feature crest factor (f10) seems to be more sensitive than the remaining features according to Figure 4-2 (d).

Therefore, three features are selected for fault diagnosis: peak to peak (f4), crest factor (f10), and energy ratio (f15) are finally employed for fault diagnosis.

4.2. Feature Based Fault Indicators

Euclidean Distance (ED) is a commonly used approach for similarity measurement, but is susceptible to small distortions. Thus, the notion of deviation ratio is proposed as the feature indicator for fault diagnosis, which is to evaluate the

overall relative difference of two signals, defined as follows:

$$Dev(X_{n_i}, X_{diag}) = \frac{\sum_k (X_{n_i}^k - X_{diag}^k)^2}{\sum_k (X_{n_i}^k - \bar{X}_n^k)^2} \quad (4-9)$$

where $X_{n_i} = [X_{n_i}^1, \dots, X_{n_i}^N]$ ($n_i = 1, \dots, M$) represents the time domain features of n_i^{th} time series data measured from the previous fault-free or normal system, and the element $X_{n_i}^k$ in X_{n_i} indicates the feature of k^{th} time interval data. N is the feature dimensionality. $X_{diag} = [X_{diag}^1, \dots, X_{diag}^N]$ denote the time domain features of diagnosis signal. $\bar{X}_n = [\bar{X}_n^1, \dots, \bar{X}_n^N]$ are the mean values of the time domain features in the fault-free system, namely $\bar{X}_n^k = \frac{1}{M} \sum_{n_i=1}^M X_{n_i}^k$. M is the number of signals measured from the fault-free system. If there is only one dataset measured from healthy system, the deviation ratio is calculated as:

$$Dev(X_{n_1}, X_{diag}) = \frac{\sum_k (X_{n_1}^k - X_{diag}^k)^2}{\sum_k (X_{n_1}^k - \bar{X}_n^k)^2} \quad (4-10)$$

In principle, the deviation rate should be close to zero if there is no fault and noise.

However, input magnitudes at different measurements may not be identical (or sometimes very inconsistent), which would largely contribute to the variance of measurement data and decrease the accuracy of fault detection. To address the problem of this distance measure, the time series data are normalized with respect to the measured input magnitude. An improved deviation ratio ($IDev$) is thus proposed as the feature indicator for fault diagnosis, which is to evaluate the overall relative difference of two signals, defined as follows:

$$IDev(X_{n_i}, X_{diag}) = \frac{\sum_k (S_{n_i}^k - S_{diag}^k)^2}{\sum_k (S_{n_i}^k - \bar{S}_n^k)^2}$$

$$S_{n_i}^k = \frac{X_{n_i}^k}{B_{n_i}^k}, S_{diag}^k = \frac{X_{diag}^k}{B_{diag}^k}, \bar{S}_n^k = \frac{\bar{X}_n^k}{\bar{B}_n^k} \quad (4-11)$$

where $B_{n_i} = [B_{n_i}^1, B_{n_i}^2, \dots, B_{n_i}^N]$ is the real input magnitude of n_i^{th} time series data measured from fault-free system, and $\bar{S}_n = [\bar{S}_n^1, \bar{S}_n^2, \dots, \bar{S}_n^N]$ is the mean values of the real input magnitudes. To evaluate the change of the deviation ratio, the following relative index is adopted:

$$Re_Dev(X_n, X_{diag}) = \frac{1}{M} \sum_{n_i=1}^M \frac{IDev(X_{n_i}, X_{diag})}{\sum_{n_j=1}^M IDev(X_{n_i}, X_{n_j}) / M} \quad (4-12)$$

where X_{diag} represents the signal to be detected, and X_{n_i} is the n_i^{th} reference signal measured from the normal/healthy condition. M is the number of signals from normal/healthy system.

4.3. Fault Detection

It is a rather challenging task for fault detection when the prior knowledge of the normal operation system is limited and the prior knowledge of faulty states are unavailable. In our study, the adaptive threshold and statistical methods are combined for fault detection.

4.3.1. Fault Indicator based Adaptive Threshold

With limited prior knowledge of normal conditions, the threshold value is difficult to define. Above that, the applicable of threshold for fault detection might be even impacted by the number of samples from the normal state for knowledge basis. Since there is no standard guideline to the appropriate sample size for

threshold, in this study, an adaptive threshold is adopted and combined with statistical tests for fault detection when there are more than six groups of signals (except the outliers) measured from the normal conditions involved.

It is assumed that there is no noise signals and outliers among the available data from the normal operational system. Otherwise, to decrease the disturbance, the k -nearest neighbors (KNN) is applied to remove the outliers and noisy signals through an iterative process [31]. Similarity, a threshold is presented based on that concept for fault detection as shown in Eq. (4-13).

$$\min_{n_i} \{ IDEv (X_{A_s}^{n_i}, X_{A_s}^{diag}) \} > T \quad (4-13)$$

$$u_{A_s}^n = \frac{1}{M(M-1)} \sum_{i,j=1}^M IDEv (X_{A_s}^{n_i}, X_{A_s}^{n_j}), \quad T = u_{A_s}^n + k\sigma(A_s^n)$$

where k is a scale factor. The $u_{A_s}^n$ represents the mean of the $IDEv$, and $\sigma(A_s^n)$ is the standard deviation of such fault indicator. Thus, $u_{A_s}^n + k\sigma(A_s^n)$ is the basis threshold for fault detection. The k value should be tuned to ensure the accuracy of classification. Too large of a value might lead to the fail of fault detection. For the application in this thesis, the tune value is ranging from 0.5 to 2, i.e., $k = 0.5 + 1.5e^{-M/100}$. When monitoring the system, observations are probably collected in a consistent way. The data collected from the repeated measurement are kept, and the threshold values should be adaptive over the times.

4.3.2. Statistical Methods

When encountering with the system with limited knowledge of priori, especially the data measured from the normal operational system are limited and small in scale

(e.g., two or three groups), the threshold is so difficult to obtain. To address this harsh conditions, statistical tests like T-test and related non-parametric tests are capable of fault detection by similarity comparison between the two datasets or two distributions (which are obtained from the repeated measurements on the same system).

In a typical experiment, it is more practical to detect the system using two datasets with one dataset collected from the healthy system (for control) and another dataset collected from the diagnosis system (for measurement). Those disadvantage factors have greatly increased the challenge of fault diagnosis. Statistical methods and statistic-based indicators have been widely used for fault detection [45-48]. The threshold associated with a statistical test is obtained from p-value listed in a statistical table, and they are standardized and easy to implement. It is assumed that the signals measured from the healthy system are representative with lower noise disturbance.

Although there is no obvious guidelines for choosing tests suitable for our application, we can say that it would not be advisable to use paired algorithms where there are any extreme outliers. In this thesis, three statistical methods (i.e. independent T-test, Kolmogorov-Smirnov test, and Wilcoxon rank-sum) frequently used for fault detection are adopted to measure the similarity of the two datasets with one reference dataset measured from the healthy system and another one diagnosis dataset measured from the diagnosis system.

(1) *Unpaired or independent t test* Two-sample t test is used to test whether the means of two samples are different, i.e. the null hypothesis: the means of two samples are equal.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1X_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$S_{X_1X_2} = \sqrt{\frac{(n_1 - 1)S_{X_1}^2 + (n_2 - 1)S_{X_2}^2}{n_1 + n_2 - 2}}$$

where \bar{X}_1 and \bar{X}_2 are the means of two samples, S_{X_1} and S_{X_2} are the standard variance of two samples, and n_1 and n_2 are the corresponding capacity. It can be used only when samples are from Gaussian distribution or approximate shape of Gaussian distribution.

(2) *Kolmogorov-Smirnov test* As a nonparametric hypothesis test (which do not depend on the assumption that values were sampled from Gaussian distributions), Kolmogorov-Smirnov test (K-S test) [57] is a popular statistical method to measure the probability that two datasets are from the same distribution (using two-sample K-S test) or a chosen univariate dataset is drawn from a given model (using one- sample K-S test). Two datasets do not need to be in same size. Though Chi-square goodness of fit test is one of most popular method to test the goodness-of-fit tests, this method is applied with large samples (> 30). K-S test has no constrict on sample size and is applicable for testing with small sample size. To measure the time series data, theoretical cumulative density function (CDF) will be obtained to compare data with the

reference distribution [35]:

$$P(x) = P(X_i \leq x) = i/N$$

where X_i is the value at i^{th} point when the data series X is re-arranged in ascending order. N is the total number of the data series. The distance D is defined as the maximum absolute distance between two CDFs. Mathematically, the KS test statistic is represented by:

$$D = \max_{x_1 \leq x \leq x_n} (|F_1(x) - F_2(x)|)$$

The similarity probability of two series data using K-S test could be defined using mathematical formulation as follows [35]:

$$P(D) = 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 \mu^2}, \quad \mu = D \left(\sqrt{\frac{N_1 N_2}{N_1 + N_2}} + 0.12 + \frac{0.11}{\sqrt{N_1 N_2 / (N_1 + N_2)}} \right)$$

If the two vibration signals are similar in distribution, then the probability $P(D)$ is approaching to 1. Otherwise, the similar probability $P(D)$ is approaching to 0 if two the distribution of two signals are different.

(3) *Wilcoxon rank-sum test* [54]

$$W = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h_{i,j} + \frac{N_1(N_1+1)}{2}, \quad h_{i,j} = \begin{cases} 1, & x_i < x_j, \\ 0.5, & x_i = x_j, \\ 0, & \text{otherwise} \end{cases}$$

where N_1 and N_2 are the sizes of two samples. The main advantage of Wilcoxon test is that it is valid for the data from any distribution and less sensitive to outliers in comparison to two sample T-test. However, Wilcoxon test cannot reflect the location of differences.

In statistical tests, $F_1(x)$ is assumed to be a target distribution and $F_2(x)$ is a reference distribution. The null hypothesis (H_0) is that two samples are from the

same distribution, otherwise, the null hypothesis is not accepted. This can be formulated as follows.

$$H_0: F_1(x) = F_2(x), \quad H_1: F_1(x) \neq F_2(x).$$

Since p-value is usually used to measure the probability that the distribution of reference signal is different from the diagnosis signal, it is widely accepted that the null hypothesis (H_0) is to be rejected when $p \leq 5\%$. Otherwise, the null hypothesis fails to be rejected.

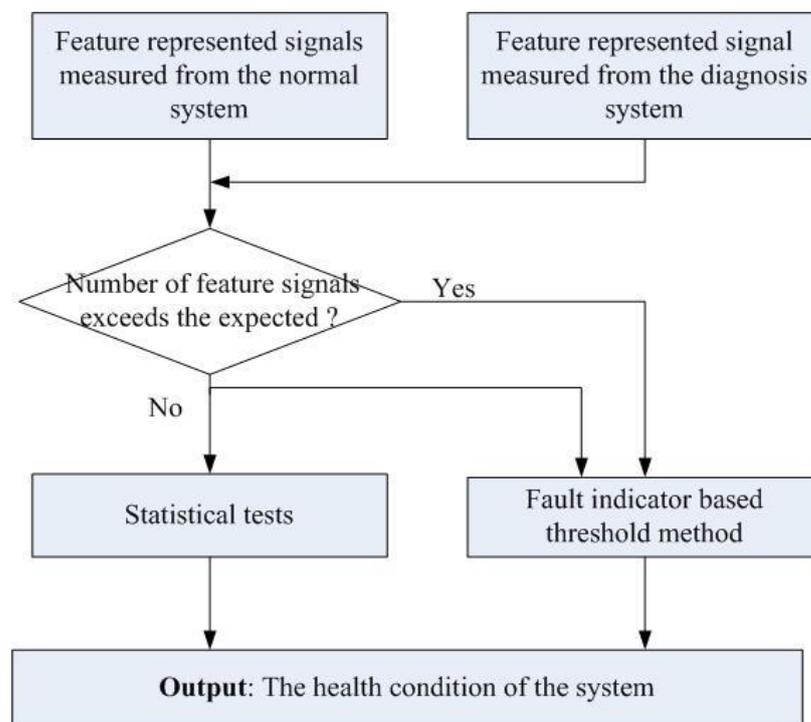


Figure 4-4 The flowwork of the combined fault detection method

4.3.3. Combined Method for Fault Detection

Fault indicator based threshold method is combined with statistical tests for fault decision making. Figure 4-4 shows the overall framework of the combined method. If the number of signals collected from the normal operational system does not exceed the expected (more than six groups in this study), both the fault indicator

based threshold method and statistical tests are provided for fault detection. The abnormal system is suspected if a fault is detected by any of these two methods.

4.4. Conclusion

Signal preprocessing including feature computation and selection is provided before the fault detection and fault localization. Two methods with different requirements of priori are studied for fault detection. Specifically, the threshold method is suggested for fault detection if there are considerable signals measured from the healthy conditions containing the representative features of the system, while statistical tests are adopted when the prior knowledge of the system are limited. Since the main objective of this thesis is to present a comprehensive fault diagnosis method with the limited prior knowledge, those two methods are combined for fault detection.

Chapter 5. Virtual Beam-like Structure Approach

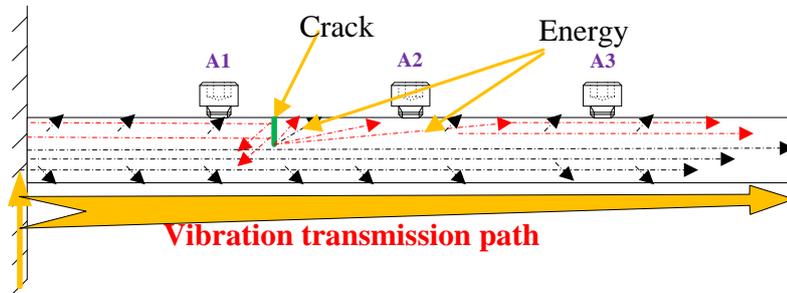
In this Chapter, a virtual beam-like structure (VBLS) approach is proposed for both single and multiple faults localization. The concept of virtual beam, the optimization methodology to construct the optimal virtual beam, and virtual beams for fault localization are all included.

5.1. Concept of Virtual Beam-like Structure

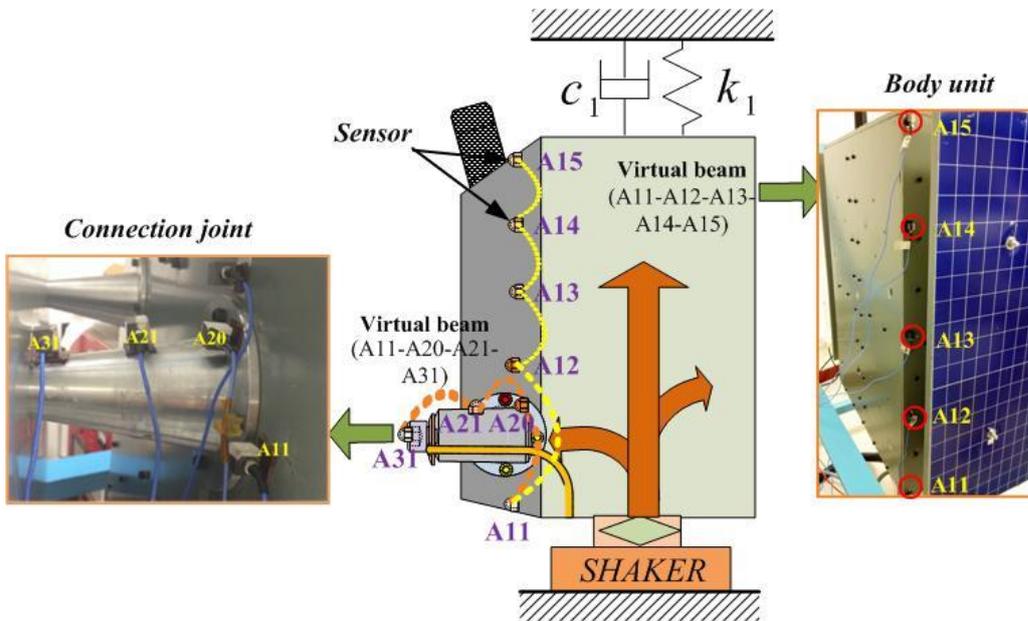
The concept ‘virtual beam’ stems from the inspiration that beam-like structures are basic structural components in various complex structural systems [187]. When the structure is excited with an appropriate input, the vibration transmission paths from the vibration source to other aspects of structure could be regarded as the “virtual beams” which could be applied to fault detection and localization since the energy intensity as well as the vibration transmission paths at some components are influenced by the occurrence of faults.

In [188], it has been formulated that the signal contains the information content of energy received by corresponding sensor. The significant differences in signals are caused by the energy changes which could be captured by sensor nodes, and the information like damage localization and orientation can be obtained according to sensor distribution in the node. Still, this study focused on the development of a set of energy correlated damage indices based on transmission and reflection, it ignored the limitation of the priori. Based on this general knowledge, virtual beam-like

structure approach is proposed for fault diagnosis, and the fault diagnosis of complex structure can be transformed into a problem of optimal selection and fault diagnosis of virtual beams.



(a). Vibration transmission path along beam-like structure



(b). Examples of virtual beams on a satellite-like structure

Figure 5-1 Two typical examples of vibration transmission path along the structures

To be clearer, a beam-like structure and satellite-like structure are applied as examples for formulation. A complex structure is regarded as a combination of numerous virtual beam-like structures considering the vibration transmission path from vibration sources to each sensor. From Figure 5-1 (a), energy transmission path is from the left bottom (i.e. vibration source) to the right of the beam, while entire

beam-like structure can be regarded as a vibration transmission path from the root to the top end represented by the sensor chain: A1-A2-A3. To apply this method to complex structures, virtual beams along vibration transmission paths in the complex structure can always be found by optimally selecting a chain of sensors from sensor networks distributed in the complex structures. Figure 5-1 (b) shows two typical virtual beams consisting of two sensor chains assigned on satellite-like structure from the bottom to other component along vibration transmission path (since the structure is excited from the bottom). The connection joint shown in Figure 5-1 (b) locates at the middle of those two sensor chains. That means the occurrence of the fault like loosening bolt on this connection joint, the signals measured by sensor chains can reflect the vibration transmission changes, especially the sensor located at the closer and after the damage, e.g. A20, A12. To be convenient, sensor on the virtual beam closer to the vibration source is named as root sensor (e.g., A11), while the one far away from the vibration source is regarded as top sensor (e.g., A15 and A31).

Thus, it is rather important to select the appropriate sensors to construct optimal virtual beam for damage localization and orientation.

5.2. Sensor Networks Distribution and Selection

In large space structures (LSS), it is practical to apply the sensor network to fault diagnosis, but it was not effective to use all of them in terms of computational complexity and efficiency [189]. Thus, the methods for sensor placement have been

developed to obtain the fewer sensors for fault diagnosis [189-192]. With the continuous advances in sensor technology and sensor placement methods, the sensor networks can be implemented for fault diagnosis of LSS with low-cost and high-efficiency. In this study, the sensors are mounted on the complex structure as a network for fault diagnosis.

Since the complex structure composes of various components, accordingly, sensors are grouped into numerous sub-networks. The separation of the sensor networks is studied based on components and search space to come up with optimal solutions and improve the fault localization accuracy. The concepts of *region of influence* (ROI) and *region of coverage* (ROC) in [170] are adopted for sensor network separation. ROI is the area around an event, while the sensors within that region are capable of fault detection in high probability (over 50%). ROC is the area around a sensor node where the occurrence of fault could be detected in high probability (over 50%). The distance between two sensors cannot too large considering the coverage of sensor network. While this distance also cannot too small in term of computational complexity and efficiency (The increase of the sensors would also bring the possibility of the negative influences of false information from the environment noise and other unexpected factors). For that consideration, the distance between two sensors is suggested to be as follows: $R_c/2 \leq \text{dis}(A_1, A_2) \leq 2R_c$, where R_c is the radius of the “region of coverage”.

Take the small satellite-like structure as example. A satellite-like structure is

hanged from the top with a spring and a shaker is attached at the bottom. Potential faults can be fatigue cracks, bolt-loosening or fractures etc. This study focuses on bolt-loosening fault since it is easy to create in an experimental setup. Bolt-loosening can happen in various hanging structures with a bolted-base connected with the main satellite-like structure, referred to as bolted-base hanging structures. The complex structure is probably divided as various components. Figure 5-2 shows some typical bolted-base hanging structures in a satellite-like model such as band antenna, solar panel, connectors.

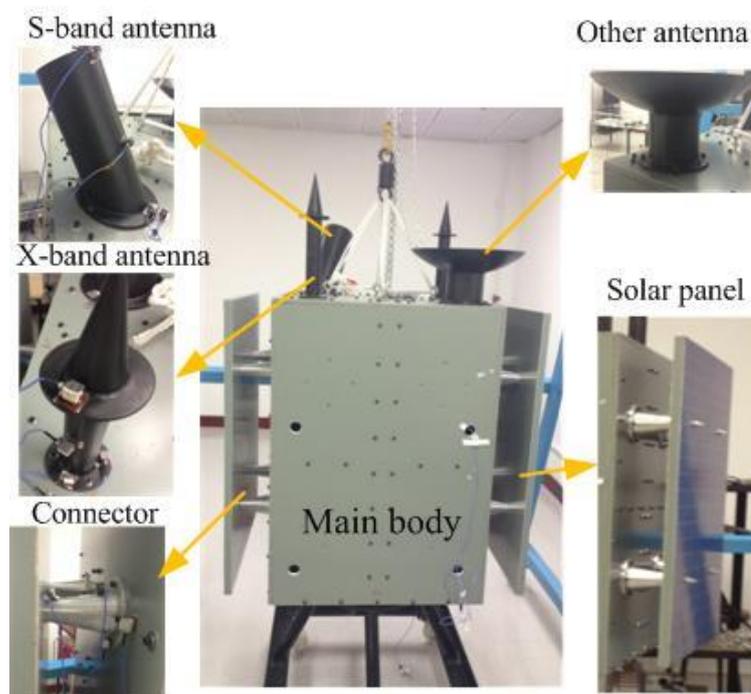


Figure 5-2 Typical bolted-base hanging structures in a satellite-like model

As shown in Figure 5-3, accelerometers are evenly distributed on the substructures, named as sensor-networked structures. For the convenience of description, the sensor located on the i^{th} row and j^{th} column of the t^{th} component is named $S_t A_{ji}$. Four connectors are named as $P_1, P_2, P_3,$ and P_4 , and the sensor on the

bottom of the connector P_i is named as AB_i , and the sensor located on the top of the connect P_i is named AT_i . The sensor networks are grouped on basis of the sub-structures. Sensors on each solar panel are grouped as a network, and the sensors on the body are grouped as another network. Since sensors on the overlapping span of two or more sub-structures are capable of fault detection in neighbor regions, they could belong to any of the sub-structures in the process of virtual beam construction. For example, sensors AB_4 and AT_4 could be used for virtual beam construction in both solar panel and main body.

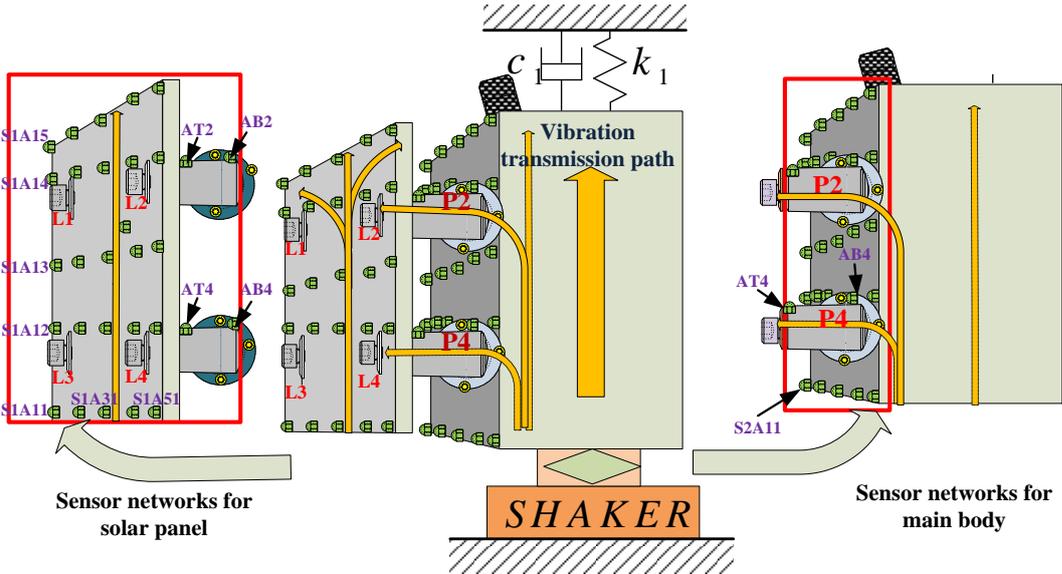


Figure 5-3 Sensor networks on the satellite-like structure

Once a component is detected for the abnormal condition, the sensor network related to that component is adopted and applied to virtual beam construction for fault localization. The candidate sensors used for virtual beams construction for a single solar panel include the sensors on that solar panel and the sensors on the four adjacent connectors, while the sensors used at the left plane of the main body in

Figure 5-3 are assigned to both that direction of the main body and on the associated adjacent connectors. Based on sensors response of sensor networks, the potential faults could be isolated with respect to certain components, which could greatly decrease the computational complexity for analysis of all sensors and improve the efficiency for fault diagnosis. To further narrow the potential faults in a given component with fewer sensors, virtual beams are constructed using the optimization methodology formulated in the next section.

5.3. Optimization Methodology for Virtual Beams

To apply the virtual beam-like structure method for fault diagnosis, the proposed BAFS method is adopted for the construction of the optimal virtual beams by selecting sensor based chains from appropriate sensor networks. Specifically, the optimization methodology involving the fitness and constraints is presented in Section 5.3.1, and application of BAFS for sensor selection is given in Section 5.3.2.

5.3.1. Objective Function and Methodology

A larger value of the relative deviation ratio indicates a larger difference in the signals measured from the diagnosis system and the fault-free system. The different continuous sensor chains in the sensor network are potential candidates for the optimal virtual beam. The length of this path cannot be too long while selecting the most informative sensors that are sensitive to dynamic changes of feature within a small region. The objective of the optimal virtual beam construction considers two

parts: the maximum of the relative deviation ratio and the minimum of sensor routing.

It is assumed that the sensor chain $[A_1, A_2, \dots, A_D]$ is selected in the optimization process. The dimensionality D indicates the number of sensors for virtual beam.

This resulting objective to be maximized is:

Fitness =

$$\sum_{s=1}^N \frac{\alpha_{A_s}}{M} \sum_{n_i=1}^M Re_Dev(X_{A_s}^{n_i}, X_{A_s}^{diag}) / \sum_{s=1}^{N-1} \alpha_{A_s} \alpha_{A_{s+1}} \beta_s Dis(A_s, A_{s+1}) \quad (5-1)$$

$$\text{with } \alpha_{A_s} = \begin{cases} 1, & \text{if sensor } A_s \text{ is used} \\ 0, & \text{otherwise} \end{cases}$$

Subject to:

$$\sum_{s=1}^N \alpha_{A_s} \leq D \quad (5-2)$$

$$\sum_{j=1}^R \alpha_{A_{r_j}} \leq max_{row} \quad (5-3)$$

$$|\alpha_{A_{r_k c_v}} \times v - \alpha_{A_{r_k c_j}} \times j| \leq 1, k = 1, \dots, R, v = 1, \dots, C, j = 1, \dots, C \quad (5-4)$$

$$\sum_{j=1}^R \sum_{k=1}^C r_j c_k = N \quad (5-5)$$

$$\sum_{s=1}^M \gamma_{A_s} \geq Err_{num}, \quad \gamma_{A_s} = \begin{cases} 1, & \text{if } \min_{n_i} \{Re_Dev(X_{A_s}^{n_i}, X_{A_s}^{diag})\} > T \\ 0, & \text{if } \min_{n_i} \{Re_Dev(X_{A_s}^{n_i}, X_{A_s}^{diag})\} \leq T \end{cases} \quad (5-6)$$

where N is the total number of sensors in a complex structure for fault detection—see Eq. (5-5) and M is the number of signals measured from the fault-free system.

$X_{A_s}^{n_i}$ represents the time domain features of n_i^{th} signal measured by sensor A_s when the system is in a healthy state. Similarly, $X_{A_s}^{diag}$ denotes the time domain

feature of diagnosis signal measured by the same sensor. The number of sensors to construct a virtual beam is D —see Constraint (5-2)—which cannot be too large or

too small. Selecting a large number of sensors could contribute to larger regions for potential occurrence, while a smaller number could decrease the accuracy of fault localization. In our experience, it is appropriate to construct the virtual beam using three to five sensors.

The virtual beam defined in the thesis represents one of vibration transmission path on the way from vibration sources to each sensor, and the most essential ones for the fault localization are aligned with the direction of vibration source. For convenience, we assume that the vibration source is in the direction of the corresponding column. The sensor network is denoted by row R and column C , and $A_{r_v c_k}$ represents the sensor located at v^{th} row and k^{th} column of the sensor network. For each row, the number of sensors cannot exceed max_{row} , as formulated using Constraint (5-3). Since the virtual beam is constructed with the sensors from varying rows, at least one sensor should be selected from each row ($max_{row} < D$). The sensors belonging to a sensor chain should be located as the neighbors if they are in the same row from the network, otherwise the selected sensor chain becomes invalid (see Constraint (5-4)).

Constraint (5-6) illustrates that a fault might occur in the system if at least Err_{num} sensors from the network indicate the abnormal. Larger number of alarm sensor means the system has a higher probability of being abnormal. However, if there is only one alarm sensor in a network, it is difficult to decide whether the system is healthy since the signal measured by the only alarm sensor might be

influenced by noise or other factors. The optimization process is terminated if Constraint (5-6) is not satisfied. This means that the system is detected in the normal state. The requirement of alarm sensor arises from the need to ensure that the system is in an abnormal condition. However, if the knowledge of priori is limited for threshold value (i.e. T), statistical methods are used for fault detection according to similarity measurement of two datasets.

The distance between two adjacent sensors from sensor chain are represented by $Dis(A_s, A_{s+1})$. The minimum of the length of a virtual beam (i.e. the shortest route that visit every selected sensor exactly once) is adopted to select the most sensitive sensors and refine the region for fault isolation. β_m is the weight vector for the relative distance (i.e., Dis) of the selected sensors, A_s and A_{s+1} . The denominator illustration of the fitness function in optimization problem, i.e., $\{\sum_{m=1}^M \alpha_{A_m} \alpha_{A_{m+1}} \beta_m Dis(A_m, A_{m+1})\}$ in Eq. (5-1), is solved as a ‘shortest route problem’ using Genetic Algorithm (GA) in Matlab toolbox [193]. After implementing the optimization process, the optimal shortest route can be obtained.

5.3.2. BAFS for Optimal Selection of Sensors

The improved bacterial algorithm based feature selection (BAFS) is used for optimal selection of sensors to construct the effective virtual beams for fault diagnosis. The pseudo-Code for virtual beam construction is given in Pseudo-Code 5-1.

Pseudo-Code 5-1 BAFS for virtual beam construction

```
01 Input: The values of  $Re\_Dev$  and  $IDev$  from sensor networks using a specific feature;  
     $max_{row}$ ; the least requirement of abnormal sensor number  $Err_{num}$ ; maximum number of  
    sensors to be used to construct the virtual beam:  $D$   
02 Initialization:  $Pop, T1, T2, T3, C, Max\_iteration$   
03 If Constraint (5-6) is satisfied or abnormal state is detected by statistical tests  
04 Do  
05 Fitness function: develop the fitness function according to Eq. (5-1);  
06 Calculate the fitness of all bacteria, and define  $Current\_iteration=0$   
07 Optimization process:  
08 If  $Current\_iteration < Max\_iteration$   
09  $Current\_iteration = Current\_iteration + 1;$   
10 Obtain the  $P_{best}$  and  $G_{best}$  //  $P_{best}$  is calculated according to the pseudo-code in Fig.  
    1, and  $G_{best}$  is the best position of the group  
11 For each bacterium  
12 Running: Adapt the position of bacteria using Eq. (3-2)  
13 Obtain the fitness  $f$  and compare with the original fitness  $Fit$   
14 If  $Fit < f$  (assume that the objective is to minimize)  
15 Tumbling: Adapt the position of bacteria using Eq. (3-3)  
16 end if  
17 Adapt the  $W$  and  $A$  using Eqs. (3-7) and (3-8)  
18 End for  
19 For all bacteria  
20 Adapt the  $W$  using Eqs. (3-9) and (3-10)  
21 If  $Current\_iteration > Max\_iteration/2$  and the  $G_{best}$  is unchangeable for  $T1$  times  
22 Reproduction: refer to BFO reproduction  
23 end  
24 If  $Current\_iteration > Max\_iteration/2$  and the  $G_{best}$  is unchangeable for  $T2$  times  
25 Elimination: refer to BFO elimination  
26 end  
27 If  $Current\_iteration > Max\_iteration/2$  and the  $G_{best}$  is unchangeable for  $T3$  times  
28  $Premature\ termination$  // end the optimization process  
29 end  
30 End for  
31 End If // the optimization process  
32 Output: The best position for the optimal represents the selected sensors (i.e. sensor chain  
    or virtual beam)
```

5.3.3. “Biased Running” based Optimization Methodology

The main challenge for multiple-fault localization is that the occurrence of one

fault (or main event) probably brings about the abnormal reflection on numerous sensors, and sensors suffering from this main event (i.e., the kernel fault with the most serious influence on the system) might only reflect this fault and ignore the other potential faults. Though some sensors may be pointing to the potential faults, the extent of response cannot exceed those from sensors distributed close to the main events, which contributes to diminished detection capability of side faults. Such factors will exacerbate the difficulties associated with multi-fault isolation since other potential faults cannot be correctly isolated.

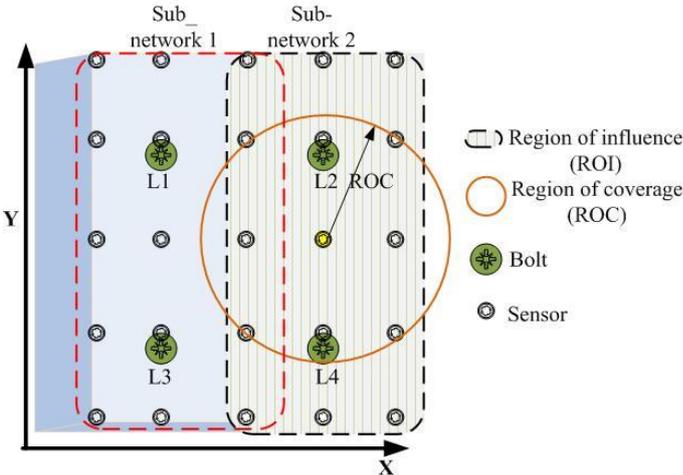


Figure 5-4 Subnetwork separation based on region of influence

(ROC is the area around a sensor node, in which the occurrence of fault could be detected in high probability (over 50%); ROI is the area around an event, while the sensors within that region are capable of fault detection in high probability (over 50%))

As shown in Figure 5-4, sensors on a component can be divided into several subsets according to the ROI (i.e., the area around an event, while the sensors within that region are capable of fault detection in high probability, over 50%) with some overlapping regions between two adjacent areas for fault detection. The loosening bolts of L1 or/and L3 located on the left of the solar panel structure are more likely

detected by the sensors in subnetwork 1, the sensors on the right region of subnetwork 2 are less effective in health monitoring these two bolts, and vice versa. Since the virtual beam consists of four to five sensors, the occurrence of the loosening bolt in L1 might be masked by loosening bolt of L4.

In this thesis, a strategy, named ‘biased running’, is presented and embedded within the optimization process for obtaining multiple virtual beams to isolate the potential faults so as to decrease the possibility that the presence of some faults are masked by others. The “biased running” indicates that multiple independent runs of the optimization process by BAFS consists of more than one virtual beams consisting of different sensor chains. In the first run, the first virtual beam consisting of sensors is selected from the networks according to the values of feature indicator (i.e., *Re_Dev*) associated with the sensors from the networks. The solution in the first run consists of a sensor chain regarded as the kernel virtual beam used to isolate the fault with the most significant abnormal indication. If the sensor network related to the faulty component contains more than one subnetworks, in the second or/and later runs, more than one virtual beam are constructed. To eliminate the disturbances arising from the main event (i.e., main fault), the sensors selected in the former runs with the largest value of fault indicator are ignored during the latter runs. If there are N subnetworks in a component, N virtual beams consisting of N different sensor chains (at least one sensor is different) are obtained after the “biased running” process. The strategy of two biased running is described in Pseudo-code 5-2.

Pseudo-code 5-2: Biased running

```
01 Input:  $Re\_Dev$  of sensors from networks using a specific feature
02 Calculation the number of subnetworks  $N$  according to Eq (5-7)
03 For  $i=1:N$ 
04   If  $i=1$ 
05     Dataset: fault indicator values  $Re\_Dev$  of sensors from the sensor network
       responsible for with the alarm component;
06   else
07     Dataset: fault indicator values  $Re\_Dev$  of sensors from the sensor network
       responsible for with the alarm component, ignoring the most sensitive sensors
       (with the largest  $Re\_Dev$ ) already selected in the previous virtual beams;
08   End if
09   Sensor selection process by BAFS, referred to Pseudo-Code 5-1
10   Solution: selected sensor chain (i.e. the virtual beam obtained in  $ith$  run)
11 End for
12 Output: Multiple virtual beams consisting of various sensor chains
```

Generally, the iteration times of optimization process for the multiple virtual beams rely on the number of subnetworks in a component. It is assumed that there are $C \times R$ sensors mounted on a component. The number of iteration times N is defined as:

$$N = \text{ceil} \left(\frac{C \times R}{D \times D} \right) \quad (5-7)$$

where D is the length of sensor chain for virtual beam. If $\left(\frac{C \times R}{D \times D}\right)$ is not an integer, then N equals to the closest neighbor integer of the element $\frac{C \times R}{D \times D}$. For example, if $\frac{C \times R}{D \times D} = 1.5$, then $N = 2$. However, if $C \times R$ is smaller than $D \times D$, the iteration time is $N = 1$. It means that there is only one virtual beam will be optimized for the fault localization.

In a large space structure, multiple virtual beams are obtained, it is possible that the same fault is isolated by more than one virtual beams. In addition, a single virtual beam consisting of a chain of sensors might be imposed by multiple faults. Thus,

the fault localization based on the optimal virtual beams will be studied according to a beam-like structure in the next section.

5.4. Validation of the “Virtual” Beam Method in a “Real” Beam-like Structure

In this section, a “real” beam-like structure with cracks at different positions is studied to preliminarily validate the effectiveness of the proposed “virtual” beam method.

5.4.1. The Beam-Like Structure

The beam-like structure shown in Figure 4-1 consists of two layer bolted steel panels of square size (40cm*5cm*0.4cm). Breathing-like cracks (around 0.3cm) or loosening-bolt can be easily created in one layer at any known positions for benchmark studying (see Figure 4-1 (b)). The faults (e.g. cracks) at different positions can be denoted by $R1$, $R2$, and $R3$, and referred to as *root crack*, *middle crack* and *top crack*, respectively. Three accelerators are placed regularly on the beam at shown positions denoted by $A1$, $A2$ and $A3$. The beam-like structure is free at the top end and attached to a shaker stringer at the root end, and then the “virtual” beam is represented by the sensor chain $A1$ - $A2$ - $A3$ as shown.

The system is firstly measured without any cracks or loosening bolts as reference signals and then is excited with known faults by using the same input. The measurement data under fault situations are used for fault feature characterization

and diagnosis. The input excitation used in our experimental testing is not necessarily any specific signal. For benchmark study, different inputs have been tested. Actually, our studies focus on the changes of vibration transmission, and thus the excitation magnitude of the input has no effect on the feature characterization and fault diagnosis. As an example, in this thesis, the input is a sweeping-frequency excitation from 20 Hz to 200 Hz with the exciting intensity 2g, the sample frequency 8192Hz, and velocity 4 oct/min.

5.4.2. Benchmark Testing on the Beam-Like Structure

As given in Section 4.1, three features are selected for fault diagnosis: crest factor, energy ratio, and peak to peak are finally employed for fault diagnosis. Based on those three time domain features, the mean value and the standard variance of fault indicators in four different fault cases are shown in Table 5-1 and Figure 5-5.

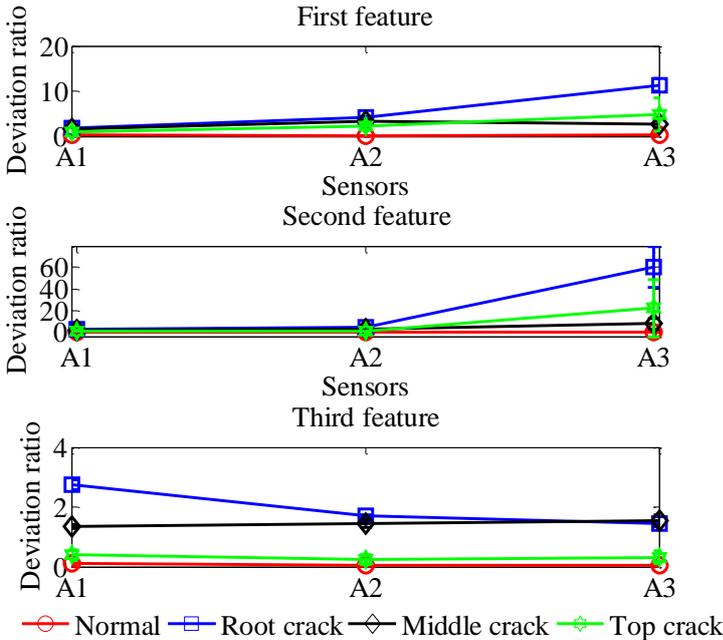


Figure 5-5 The mean and variance IDev of three sensors in different conditions

From Figure 5-5, the deviation ratios from normal states of signal features are

always the smallest, and as the deviation ratio exceeds the normal acceptable level (e.g., threshold method: value of relative deviation ratio exceeds the upper boundary of the normal), a fault might exist in the system. The mean values and variances of the deviation ratio (*IDev*) given in Table 5-1 are shown in Figure 5-5. The main results as revealed by the sensor chain for fault diagnosis are summarized as:

- 1) The occurrence of a root crack leads to the largest deviation ratios (*Re_Dev*) in all features from all three sensors in comparison to the deviation ratios (*Re_Dev*) obtained when the system is in the three other health conditions;
 - 2) When a crack located at the middle between sensor A1 and A2, the features from sensor A2 and A3 give the larger deviation ratios, especially for the second feature (energy ratio) and the third feature (peak to peak);
 - 3) The top crack located between A2 and A3 leads to larger feature deviation ratios in data from sensor A3, while the feature deviation ratios from sensor A1 and A2 have no significant difference from the normal cases.
- b). The above insights provide a very useful basis for fault diagnosis including the fault occurrence and positions. As for the beam-like structure, the vibration energy is transmitted from the root (or left) to the top end (or right), along which the three sensors A1, A2 and A3 are located.

Table 5-1 The results of fault indicators based on features
 (* F1: Crest factor; F2: Energy Ratio, F3: Peak to peak)

Sensors	Normal (N)		Root (R)		Middle (M)		Top (T)	
	<i>IDev (N)</i>	<i>IDev (R)</i>	<i>Re_Dev</i>	<i>IDev (M)</i>	<i>Re_Dev</i>	<i>IDev (T)</i>	<i>Re_Dev</i>	
A1	F1	0.34±0.26	1.74±0.19	5.12	1.57±0.09	4.63	1.04±0.53	3.07
	F2	0.24±0.19	2.72±0.21	11.23	2.21±0.36	9.15	1.52±0.41	6.29
	F3	0.09±0.08	2.74±0.17	27.57	1.34±0.06	13.49	0.39±0.17	3.92
A2	F1	0.11±0.06	4.14±0.36	38.33	3.26±0.36	30.46	2.19±0.67	20.44
	F2	0.04±0.03	5.04±0.22	126.88	3.09±0.23	77.83	1.01±0.25	25.23
	F3	0.05±0.05	1.69±0.05	31.07	1.46±0.13	26.69	0.24±0.16	4.49
A3	F1	0.39±0.22	11.36±1.41	29.15	2.76±0.63	7.08	4.83±3.59	12.41
	F2	0.14±0.13	60.37±19.16	436.54	8.01±5.14	57.90	22.36±26.2	161.67
	F3	0.05±0.04	1.44±0.04	29.27	1.54±0.19	31.29	0.32±0.21	6.54

The finding that the occurrence of a crack along the way of vibration transmission path is manifested in the responses from sensors located on that path has been explained via energy based analysis in literatures [188, 194]. Generally, if the crack is located closer to the root of the vibration transmission path, the feature deviation ratios would be larger in more sensors along the vibration transmission path. If the fault position is far away from the root of the vibration transmission path, fewer sensors can detect the fault features.

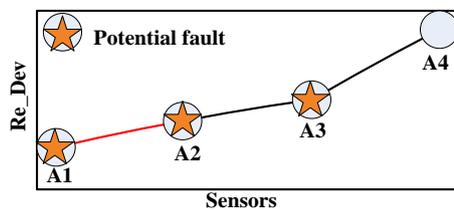
5.5. Fault Localization based on Virtual Beams

In this section, the method for fault localization is presented using virtual beams. Since one or more virtual beams might be obtained after the optimization process, fault localization based on single virtual beam and multiple virtual beams are formulated, separately.

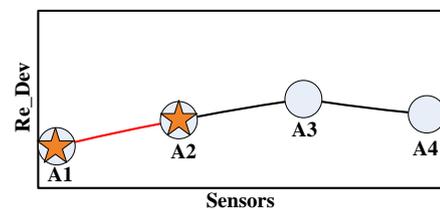
5.5.1. Fault Localization based on Single Virtual Beam

The virtual beam can be automatically constructed by the proposed bacteria based optimization method and labeled by a corresponding chain of sensors placed in the structures along the vibration transmission path. The next stage is to decide the fault location according to beam-like structure. A summary of this beam-like structure approach for fault isolation is given as follows.

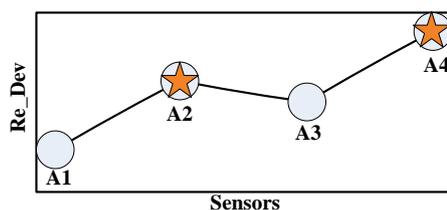
- 1) Occurrence of fault close to vibration source found more alarm sensors from the network with a large deviation ratio in almost all features of sensors on beam-like structure;
- 2) Occurrence of fault away from the vibration source may bring smaller alarm sensors with smaller deviation ratio in comparison to a fault close to the vibration source if the fault extent is the same;
- 3) A fault results in a larger relative deviation ratio in the sensors that are close to the fault or at the afterward of the transmission path;
- 4) The sensors locate closer to the vibration source are less capable of being detected if the fault is located far away from the vibration source.



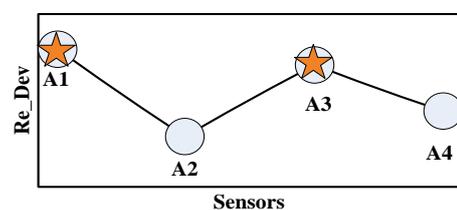
(a)



(b)



Sensors



Sensors

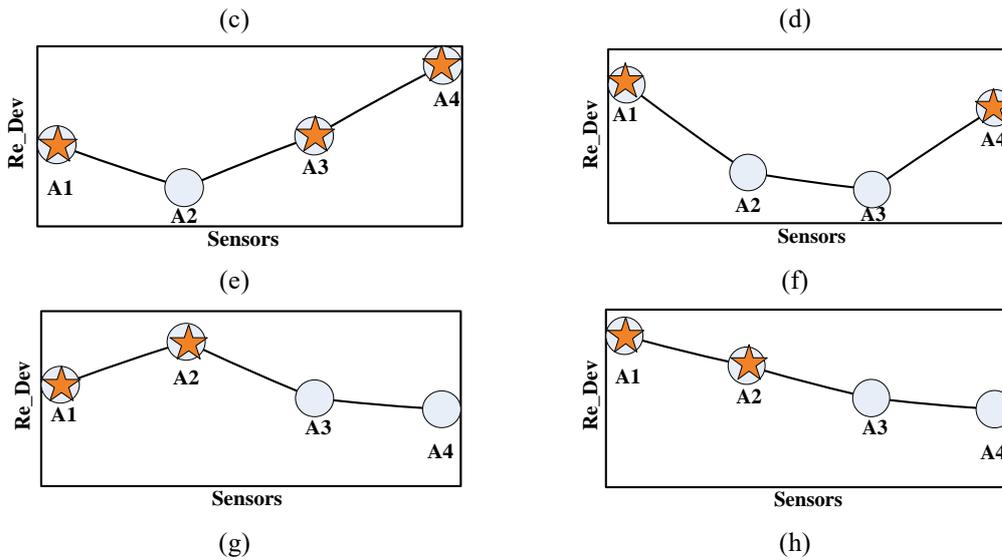


Figure 5-6 Typical examples of fault isolation according to the Re_Dev of the ‘virtual beam’ (* Vibration transmission path is A1-A2-A3-A4 through A1 to A4, and every two neighbor sensors on that path located as neighbors on the structure)

Accordingly, the summary from beam-like structure is presented above as a rule of generality for the virtual beam-like structure approach for fault localization. To be clear, some typical examples for fault isolation using ‘virtual beam’ have been shown in Figure 5-6. Four sensors (i.e., A1 to A4) from the sensor network are selected to construct a ‘virtual beam’. Here, for convenience, all four sensors are assumed to indicate the abnormal condition of the system though it is not a requirement for the optimal virtual beam (referred to Section 4.3). The vibration transmission path is from sensor A1 to sensor A4 and two neighboring sensors on that path are also the neighbors on the structure.

The first sensor A1 is closer to the vibration source, the occurrence of fault around sensor A1 would bring larger Re_Dev to the sensors afterward of transmission path, e.g., in Figure 5-6 (a), Figure 5-6 (b) and Figure 5-6 (g). Even so, the signals might be influenced by the noise or other unexpected factors which might bring rather

large or small Re_Dev of some sensor like Figure 5-6 (f), the multiple positions are to study for a potential fault. An event is suspected for the potential fault if it is localized inside the *region of convergence* (ROC) [170] of a sensor node highlighted in Figure 5-6.

In our study, three features are provided as the representatives for signals provided by sensors, so there might be more than one virtual beam for fault diagnosis. Thus, the rules for fault localization based on multiple virtual beams will be presented in the next section.

5.5.2. Fault Localization based on Multiple Virtual Beams

After the optimization process, more than one virtual beams might be obtained using time domain features. Though various virtual beams consisting of different sensor chains are available for fault localization, they might indicate a same potential fault position, which means that the system is probably suffering from only one fault rather than multiple case. However, if virtual beams reflect different fault position, we cannot conclude the system is definitely suffering from the multi-fault since the non-kernel virtual beam might be obtained because of the main event.

To be clear, virtual beams obtained from a component with two subnetworks are taken as an example to illustrate. If there are two subnetworks in a component, two virtual beams would be constructed on the basis of each feature. A summary of fault localization based on virtual beams from different features is provided and shown in Figure 5-7. For each feature, if two virtual beams indicate at least two different

potential positions of fault, then more than one regions are suspected for fault examination. Otherwise, there is only one potential fault being isolated if two virtual beams indicate the same potential fault. Based on multiple features, a fault is considered in the system if it has been detected by at least two virtual beams based on features. While a component is suspected for multiple faults if more than one faults are detected simultaneously by multiple features.

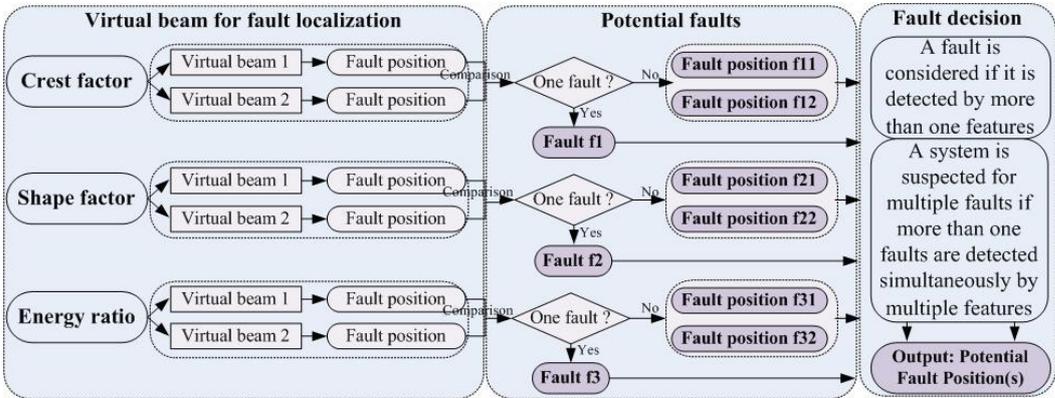


Figure 5-7 Virtual beams for fault localization based on time domain features

Even so, the sensor chain obtained after the first run is considered as the kernel virtual beam which would be given the first priority of fault localization in comparison to the virtual beam created at the second run. When there is a conflict of decision between kernel virtual beam in one feature and second virtual beam based on another feature, the second virtual beam would conform to the potential fault position localized by the kernel virtual beam.

5.6. Conclusion

A novel virtual beam-like structure method capable of optimizing the fault diagnosis process for complex structures using sensor network with limited or less

prior knowledge of faults has been developed and described in this section. The optimization algorithm presented in Chapter 3 is applied to construction of the optimal virtual beams using vibration sensors to enable automatic fault diagnosis. The beam-like structure is served as the benchmark system to provide some general rules to narrow the regions for the potential faults based on a given virtual beam. The validation of this optimization method based virtual beam-like structure approach for fault diagnosis will be conducted in the next section.

Chapter 6. Fault Diagnosis of Complex Structures with Single Fault

In this section, the VBLS approach is demonstrated by applying to detecting and isolating of a satellite-like structure with only one loosening screw. Figure 6-1 shows the overall framework of the proposed method for fault diagnosis of complex structure with only single fault. To distinguish with multiple fault diagnosis, the strategy of biased running is not adopted since there is only one fault, and the threshold based method is applied to fault detection with the assumption that the prior knowledge of the normal conditions are available.

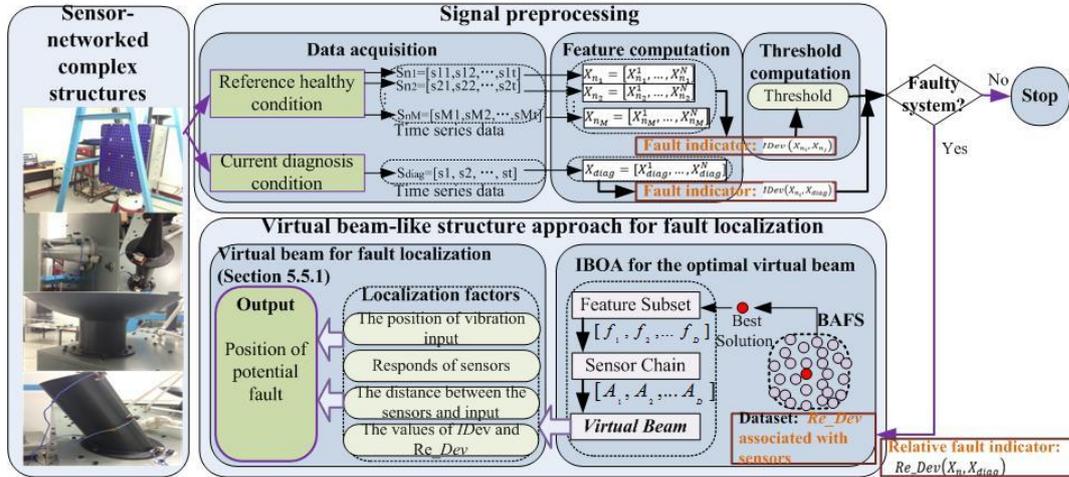


Figure 6-1 The overall framework of virtual beam-like structure approach (VBLS) for single-fault diagnosis

6.1. Experimental Platform

As shown in Figure 6-2, a satellite-like structure is hanged from the top with a spring and a shaker is attached at the bottom. Potential faults can be fatigue cracks,

bolt-loosening or fractures etc. This study focuses on bolt-loosening fault since it is easy to create in an experimental setup. Bolt-loosening can happen in various hanging structures with a bolted-base connected with the main satellite-like structure, referred as bolted-base hanging structures (BBHSs) like solar panel.

Feature dynamics without faults are available and the data with bolt-loosening faults can be obtained subject to a similar excitation by a shaker. Similar to beam-like structure, accelerometers (PCB Model 356M131 with sensitivity 10mv/g) are placed as the network (e.g., 25 accelerometers evenly distributed on solar panel structure in Figure 5-3) for monitoring the fault condition of the BBHS. There is no restriction on the stationary of the process, but all samples are collected at the same time instant (spatial correlation). The input is adopted here as a sweeping-frequency excitation from 20 Hz to 200 Hz with the exciting intensity 0.4g, the sample frequency 8192Hz, and velocity 4 oct/min.

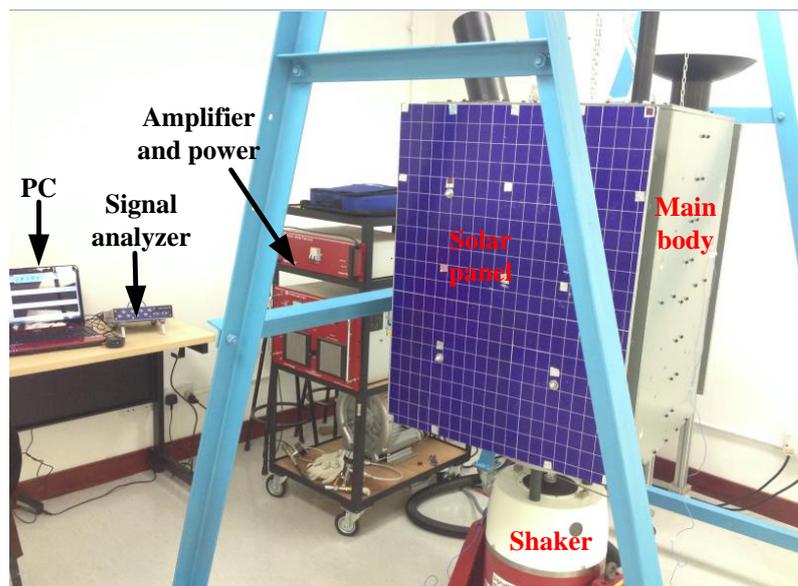


Figure 6-2 Satellite-like structure and testing platform

6.2. Experimental Testing and Results

Varying bolt-loosening faults on the satellite-like model are considered. The typical complex substructures on the satellite-like model to be detected include solar panel, main structure body (or body unit), and band antenna.

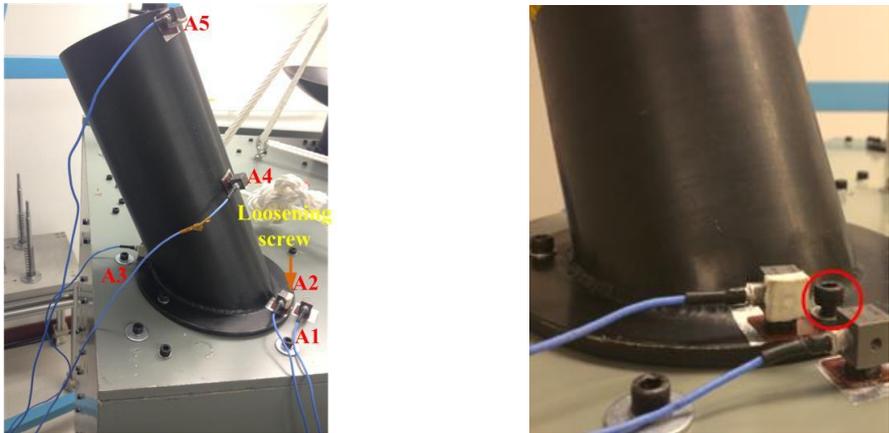
Though sensor location or placement approaches for optimal detection have been studied in the literature for modeling the system, few results have been obtained from the non-model methods. In existing studies, sensors are mostly placed at positions where faults occur frequently [184, 195], or installed along vertical, horizontal and axial directions to pick up the vibration signals [196, 197], or installed at positions determined by considering both the aspects [198]. The most popular method of isolating potential faults is to distinguish the fault according to faulty segments [184, 199], which is not applicable in our study since the unknown or limited priori. In our study, accelerometers were networked on structures to monitor the health condition of the system (e.g. the sensor networks in Figure 5-3).

Currently, many algorithms have been proposed for sensor network localization. Among the model-free localization methods without requirements of prior knowledge of faults, a series of binary estimators are proposed for fault localization because of simple, low-cost and fault-tolerant. The frequently used binary estimators like fault tolerant maximum likelihood (FTML), Centroid Estimator (CE), Maximum Likelihood (ML), and subtract on negative add on positive (SNAP) were compared in [172], indicating that the SNAP was superior to other three methods

for fault localization in terms of accuracy and computational complexity. The SNAP [170] was proposed for fault localization using a likelihood matrix to record the (negative or positive) contribution of each sensor node based on their observations. The potential fault is isolated at the event with the maximum value in that matrix. In this section, the proposed VBLS approach is compared with the SNAP for single fault localization. The studied component like solar panel is divided into a grid g with 100×100 cells and grid resolution is $g = 1$. The radius of ROC is the length of 35 cells, i.e., $R_c = 35$ (e.g. the distance between the sensor S1A15 and screw L1 in Figure 5-3), and $R_c = R_f$.

The sensor chain in our study is regarded as ‘virtual beam’ consisting of four sensors ($D=4$) used for initialization of the dimensionality (D) in BAFS. The minimum number of sensors with the abnormal signal are defined as 2 (i.e. $Err_{num} = 2$) in two complex structures (i.e. solar panel and body unit) and 1 in the simple structure (i.e. band antenna). Since the maximum sensor for each row max_{row} should be smaller than D , we set $max_{row} = 3$. The remaining parameters in the optimization method are initialized the same way as in the feature selection process. The threshold for each sensor for fault detection is determined by T (Eq. (4-13)) on the basis of prior information in normal conditions. P-value for statistical tests is 0.05 (i.e. 95% confidential range). The dataset for the fault indicator is obtained through six independent measurements taken on the normal system and

two measurements from the diagnosis system. The weight vectors for the relative distance of sensors in this study are the same (i.e., equal to 1 for all elements).



(a) Sensor network for fault diagnosis (b) Loosening screw state for detection

Figure 6-3 Sensor networked structure of band antenna

6.2.1. Fault Diagnosis on Band Antenna Structure

The band antenna structure located at the top of satellite model was fixed on the bus unit with four screws along each direction. As shown in Figure 6-3, sensors were assigned on the band antenna along the vibration transmission path. Sensor A1 was located at the adjacent region of the band antenna and bus body could be used for construction of virtual beams in both two structures. Since the band antenna studied was a relatively simple structure with smaller sensors from the sub-network (A1, A2, A3, A4 and A5), optimization method was not necessary for determining optimal virtual beam and the virtual beam was constructed using all sensors if a fault was detected by any sensor. The vibration source was excited from bottom of the satellite, and transmission path was from sensor A1 to sensor A5. The diagnosis signals were measured from the system with one loosening screw (close to A1 and A2) shown in Figure 6-3 (b).

Table 6-1 shows the deviation ratio ($IDev$) and the relative deviation ratio (Re_Dev) for each sensor based on three feature representatives. The sensors that reflect the abnormal state of the system are highlighted in bold. The boxplots in Figure 6-4 indicate similar conclusions using the feature representative ‘peak to peak’ (with almost all alarm sensors). The changes of Re_Dev for the diagnosis signals (associated with sensors) are presented in Figure 6-5. Constraint (5-7) is not considered (since the number of sensors consisting of the sub-network is 5 without exceeding 10 required in BAFS). Since at least one of sensors from the virtual beam reflects the abnormal system, all features are appropriate during fault diagnosis.

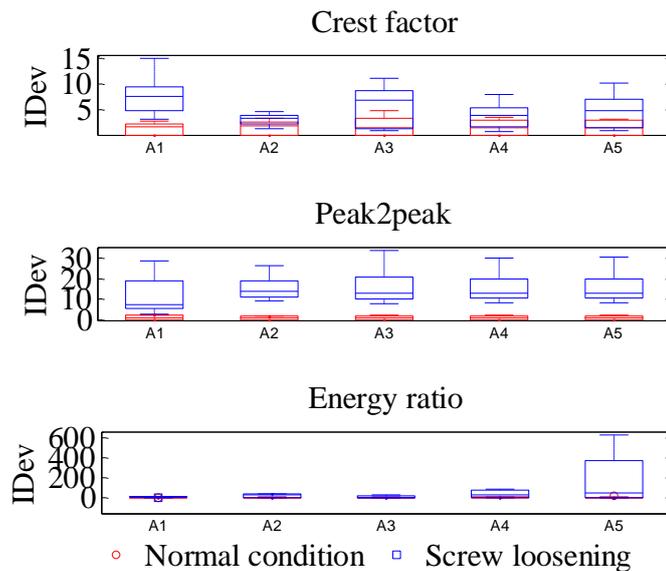


Figure 6-4 Comparison between the normal system and fault system according to $IDev$ of sensors

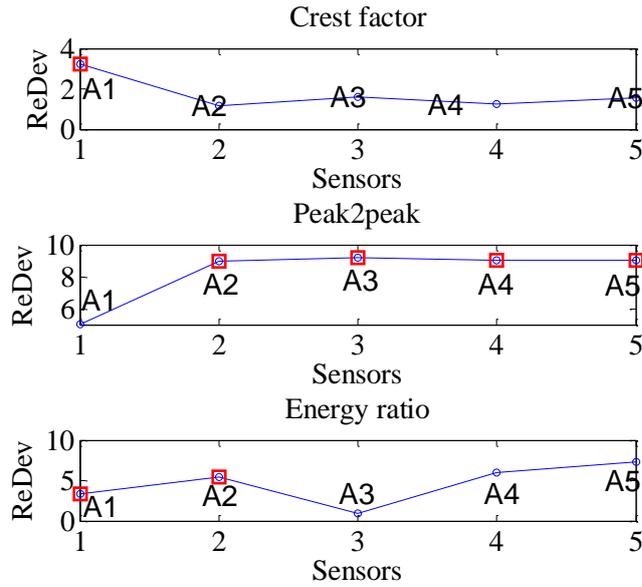


Figure 6-5 The changes of Re_Dev along the transmission path when the signals are from the diagnosis system (*sensors indicating the faulty system are highlighted in square shaped mark)

Table 6-1 Deviation ratio (IDev) and relative deviation ratio (Re_Dev) based on feature representatives from sensor chain

Features	Sensor	IDev (Normal)	IDev (Diagnosis)	Re_Dev (Disgnosis)
Crest factor	A1	2.3406±0.6023	7.5746±3.4380	3.2362
	A2	2.7046±0.8873	3.1256±1.0588	1.1557
	A3	3.5576±2.0509	5.6332±3.7526	1.5834
	A4	3.1285±1.7800	3.9444±2.4153	1.2608
	A5	3.0806±1.7391	4.7882±3.1536	1.5543
Peak2peak	A1	2.3098±1.4929	11.6090±9.0385	5.0261
	A2	1.7033±0.6558	15.2666±5.2243	8.9628
	A3	1.7497±0.8806	16.0499±8.1051	9.1729
	A4	1.7267±0.8479	15.6056±6.6677	9.0378
	A5	1.7385±0.8683	15.6677±6.9258	9.0121
Energy ratio	A1	2.2155±1.2793	7.3485±2.8031	3.3168
	A2	4.3239±3.2043	23.4715±16.2854	5.4283
	A3	8.8846±13.9156	8.7803±8.5731	0.9883
	A4	6.2454±5.3925	37.3870±35.2760	5.9863
	A5	22.7528±21.4871	166.4417±165.1747	7.3152

According to the crest factor, only sensor A1 reflects the abnormal state. We cannot decide whether the fault locates in the body unit or band antenna. While the feature based on peak to peak shows the obvious fault state, *Re_Dev* of sensors A2-

A3-A4-A5 are increasing. The potential fault is probably located around A2, A3 and A4 referring to Figure 5-6(a). Based on energy ratio, the potential fault is around sensors A1 and A2. Therefore, the potential fault is localized around sensor A2 and at the bottom of band antenna according to the proposed virtual beam based method. In this case, the potential fault is likely to be around sensor A2, which is also the situation found by SNAP method.

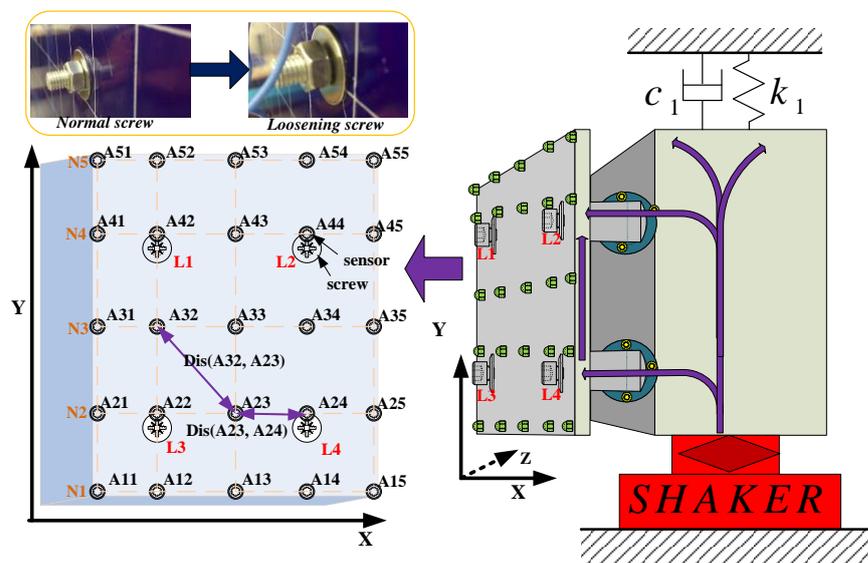


Figure 6-6 Sensor networked structure of solar panel

6.2.2. Fault Diagnosis on Solar Panel Structure

Solar panels are typical BBHSs in the satellite-like model. In this section, one of the solar panels was studied for the diagnosis of the loosening screw on that substructure. As shown in Figure 6-6, sensors distributed on the solar panel are mainly studied to monitor the health condition of this substructure. The signals for diagnosis are measured from the fault system with one loosening screw: loosening screw L2 or loosening screw L4.

Table 6-2 gives the $IDev$ and Re_Dev of the sensors from the virtual beams

obtained by the BAFS method. Two features (i.e. peak to peak and energy ratio) satisfying Constraint (5-6) are provided as representatives of the signals. The virtual beams obtained by optimization method are highlighted in Figure 6-7(b) and Figure 6-8(b), and the potential faults based on virtual beam-like structure method are pointed in Figure 6-7(d) and Figure 6-8 (d).

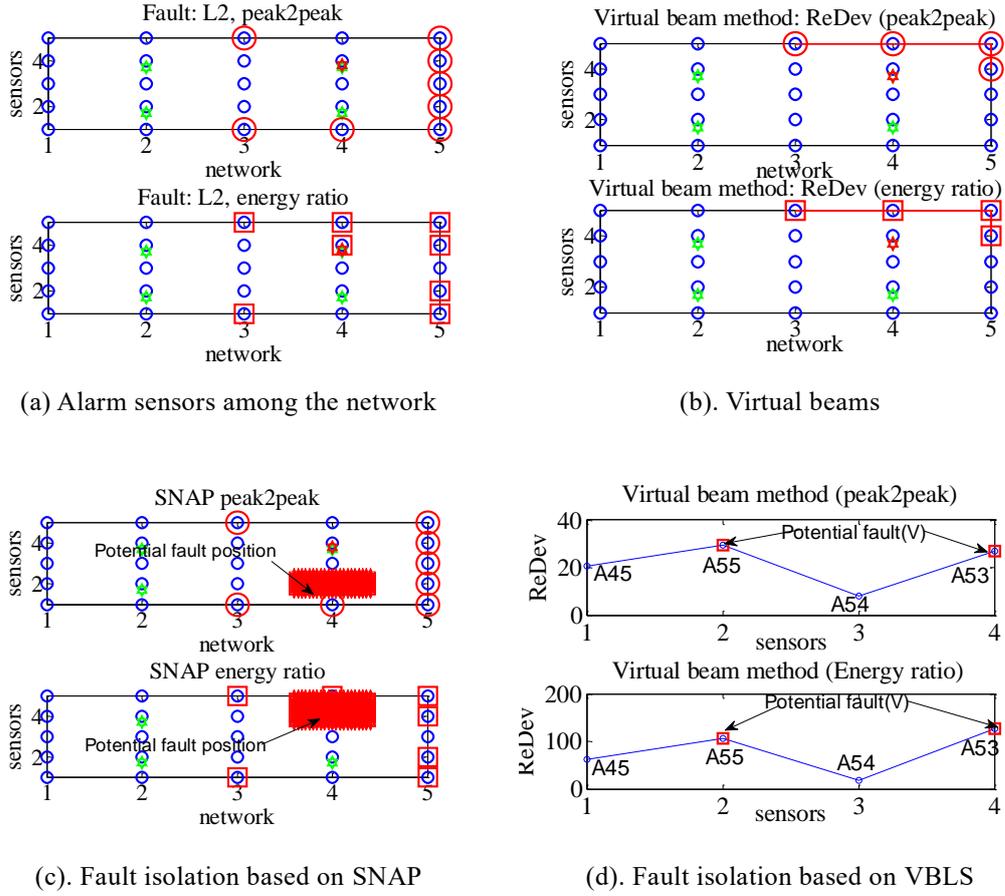


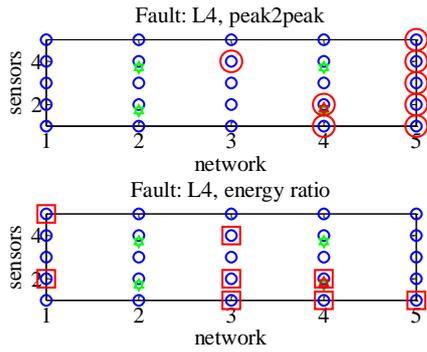
Figure 6-7 VBLS method and SNAP method for fault isolation when the diagnosis system is suffering from a loosening screw L2.

In the first diagnosis system, the virtual beams obtained using two features were the same: A45-A55-A54-A53. The trend of Re_Dev in Figure 6-7 (d) was similar to the case in Figure 5-6(c), so the potential fault was around sensors A55 and A53. Thus, the potential fault was localized in L2 screw, which was identical to the actual condition of the diagnosis system. Similarly, virtual beams were constructed using

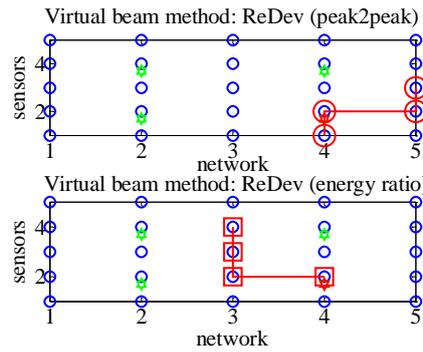
BAFS in second diagnosis system. The trend of Re_Dev based on the second virtual beam in Figure 6-8 (d) had indicated that the potential fault was in L4 screw (around sensor A23 and A24) or L2 screw (around sensor A43) — see Figure 5-6(f). The trend of Re_Dev based on the first virtual beam in Figure 6-8 (d) shows that the potential fault is L4 screw (around sensor A14 and A24) referring to Figure 5-6 (f). Therefore, the loosening screw L4 was isolated using the proposed virtual beam method.

Table 6-2 Results of feature-represented sensors based on optimal virtual beams (L2 and L4 fault)

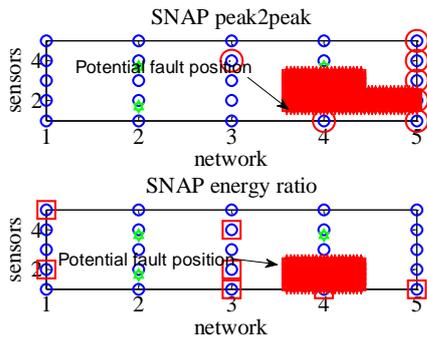
Feature	L2 Fault			
	Virtual Beam	$IDev$ (Normal)	$IDev$ (Diagnosis)	Re_Dev (Diagnosis)
Peak2peak	A45	0.1750±0.1297	3.6026±2.4374	20.5870
	A55	0.0733±0.0604	7.7350±2.8478	29.4732
	A54	0.1378±0.1046	1.0908±0.5258	7.9151
	A53	0.2010±0.1393	5.4318±1.8984	27.0194
Energy ratio	A45	0.0910±0.0743	5.6291±1.8934	61.8600
	A55	0.0733±0.0604	7.7350±2.8478	105.5254
	A54	0.0579±0.0346	1.0333±0.8878	17.8584
	A53	0.0659±0.0524	8.3404±2.4909	126.5724
Feature	L4 Fault			
	Virtual Beam	$IDev$ (Normal)	$IDev$ (Diagnosis)	Re_Dev (Diagnosis)
Peak2peak	A35	0.0877±0.0691	5.4717±2.1814	62.4085
	A25	0.0801±0.0664	4.3116±1.7888	53.8325
	A24	0.0607±0.0382	3.9737±1.9708	65.4589
	A14	0.0649±0.0383	4.6388±2.0632	71.4539
Energy ratio	A24	0.0806±0.0549	10.9196±6.7638	135.4077
	A23	0.0920±0.0641	9.9495±3.9524	108.0746
	A33	0.0563±0.0285	2.9290±0.8344	52.0444
	A43	0.0965±0.0961	23.1811±23.1811	240.1040



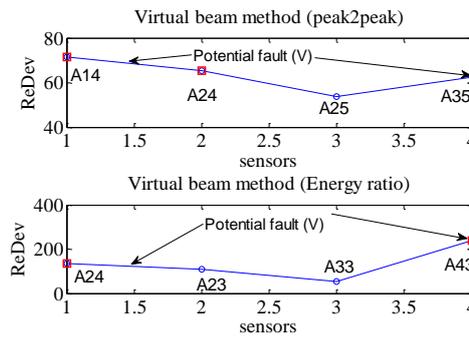
(a) Alarm sensors among the network



(b). Virtual beams



(c). Fault isolation based on SNAP



(d). Fault isolation based on VBLS

Figure 6-8 VBLS method and SNAP method for fault isolation when the diagnosis system is suffering from loosening screw of L4.

The distribution of alarm sensors are shown in Figure 6-7 (a) and Figure 6-8 (a) when the diagnosis system is suffering from the loosening screw of L2 and L4, respectively. Based on the distribution of alarm sensors, the potential fault positions detected by SNAP approach are isolated and highlighted in Figure 6-7 (c) and Figure 6-8 (c). According to the SNAP approach, the sensors indicating the abnormal state were mainly located at the right of solar panel. As shown in Figure 6-7 (c), the system was determined to have been beset by L4 fault according to peak to peak, but L2 fault was detected in terms of energy ratio. Thus, it could be difficult to distinguish between faults located in L2 screw and L4 screw, which provided larger potential fault region for further carefully study. In the second diagnosis system,

(see Figure 6-8 (c)) the L4 fault was isolated by both the feature representatives since the alarm sensors were mainly located nearby. Therefore, the fault decision for the second diagnosis system was considered to have been located at the loosening screw L4 using SNAP.

In comparison to SNAP approach, the proposed virtual beam-like structure method is more effective in isolating the loosening screw on solar panel by providing more accurate indications for fault localization and narrowing the potential fault region for further immediate remedies.

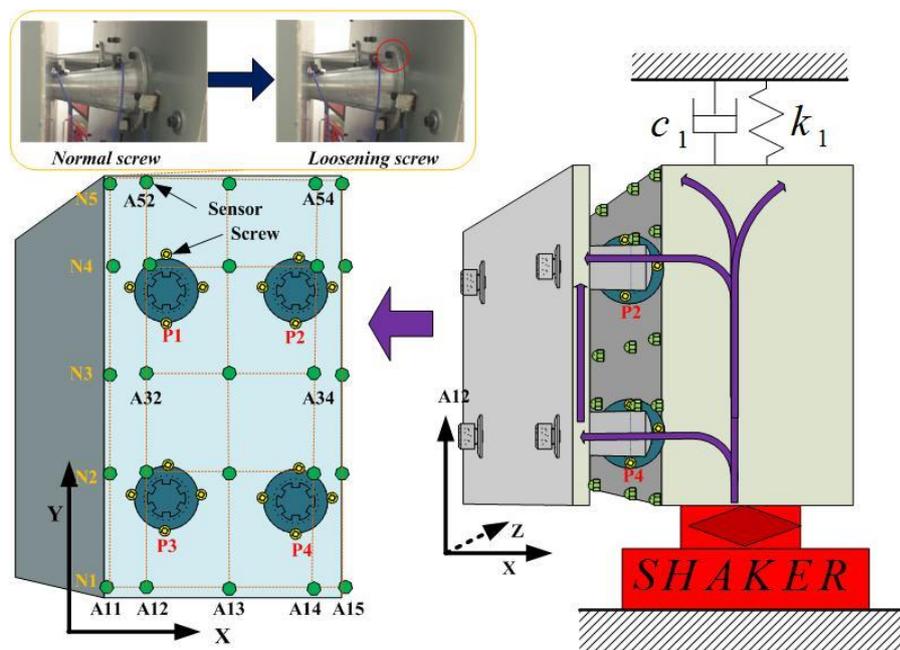


Figure 6-9 Sensors are networked on main body of satellite-like structure for virtual beam construction. *Four regions (i.e. P1, P2, P3 and P4) are frequently suffering from the loosening screw fault.

6.2.3. Fault Diagnosis on Body Structure

As shown in Figure 6-9, the main body of the satellite-like model is more complex in comparison to solar panel. The screws at the bottom of four cone shaped connectors are the only tools connecting the solar panel to the main body. Therefore,

loosening of those screws lead to rather serious safety problems and need to be detected and isolated at the very beginning.

Table 6-3 Results of feature-represented sensors based on optimal virtual beams (P1-P4 fault)

Feature	P1 Fault			
	Virtual Beam	IDev (Normal)	IDev (Diagnosis)	Re_Dev
Peak2peak	A22	0.0171±0.0153	0.0699±0.0365	2.1744
	A32	0.1084±0.1304	0.2356±0.1937	4.0928
	A42	0.0088±0.0073	0.0755±0.0229	8.5735
	A52	0.0310±0.0269	0.3838±0.0834	12.3835
Energy ratio	A41	0.0211±0.0157	0.2008±0.0567	9.5137
	A51	0.0237±0.0158	0.2397±0.1191	13.2868
	A52	0.0060±0.0030	0.2427±0.0549	40.3969
	A42	0.0620±0.0614	0.8238±0.2653	10.1163
Feature	P2 Fault			
	Virtual Beam	IDev (Normal)	IDev (Diagnosis)	Re_Dev
Energy ratio	A24	0.0581±0.0434	0.4232±0.1662	7.2870
	A34	0.0967±0.0665	0.5952±0.0903	6.1519
	A44	0.0919±0.0796	1.1293±0.1430	12.2862
	A54	0.1140±0.0834	1.6128±0.2069	14.1411
Feature	P3 Fault			
	Virtual Beam	IDev (Normal)	IDev (Diagnosis)	Re_Dev
Crest factor	A12	0.1378±0.1307	0.7181±0.3769	5.2118
	A22	0.0731±0.0695	0.7341±0.3300	10.0403
	A32	0.1633±0.1250	0.5403±0.2106	3.3096
	A42	0.0897±0.0774	0.8181±0.4341	9.1219
Peak2peak	A22	0.0666±0.0399	0.6482±0.2195	9.7384
	A32	0.1579±0.1310	1.1124±0.6299	7.0446
	A42	0.0673±0.0440	0.6619±0.2356	9.8287
	A52	0.2250±0.1828	1.7278±0.7503	7.6788
Energy ratio	A22	0.0862±0.1289	1.0500±0.5418	12.1821
	A32	0.3204±0.2403	2.3689±1.7827	7.3931
	A42	0.0807±0.0983	1.1618±0.6217	14.3971
	A52	0.0987±0.0710	1.5787±1.1465	15.9921
Feature	P4 Fault			
	Virtual Beam	IDev (Normal)	IDev (Diagnosis)	Re_Dev
Energy ratio	A24	0.0581±0.0434	0.6518±0.1090	11.2240
	A25	0.0468±0.0374	0.1695±0.0943	3.6236
	A35	0.0343±0.0284	0.2895±0.1099	8.4406
	A45	0.0283±0.0156	0.1030±0.0549	3.6396

Accelerometers used for fault diagnosis of main body were from multiple subnetworks since there were several adjacent substructures. As shown in Figure

6-9, sensors A22, A24, A42 and A44 assigned to the four cone shaped connectors were on the band of adjacent substructures (body unit and cone shaped connectors), which could be used for the construction of the optimal virtual beams using BAFS. The sensors were distributed evenly over the main body (5×5). Four positions, i.e., P1, P2, P3 and P4, were frequently suffering from the screw-loosening. Therefore, the diagnosis signals were measured when each of them was caused by a loosening screw, e.g., one loosening screw on P1 (or P2, P3, P4).

Table 6-3 gives the *Dev* and *Re_Dev* values of the sensors from the optimal virtual beams obtained by the BAFS. Only one feature (energy ratio) was used for fault diagnosis when the diagnosis system was suffering from P2 and P4 faults with respect to Constraint (5-7), while two features (energy ratio and peak to peak) and all the three features were used for fault diagnosis when the main body was suffering from P1 and P3 faults. The results of *Re_Dev* values are listed in Table 6-3 graphically presented in Figure 6-10 to Figure 6-13.

According to Figure 6-10, two virtual beams were obtained using two feature representatives: A22-A32-A42-A52, and A41-A42-A52-A51. As there were only two alarm sensors on the first virtual beam (sensors A42 and A52), the potential fault based on the first virtual beam was around sensors A42 and A52, i.e., fault P1. According to Figure 5-6(b), the potential fault was around sensor A41, A42 and A52 according to the trend of fault indicator based on the second virtual beam. Therefore, the fault decision based on virtual beam-like structure was fault P1. According to

the response of alarm sensors, the potential fault position isolated by SNAP was highlighted in Figure 6-10 (c), indicating that the potential fault was located at P1. The diagnosis signals measured from the P1 fault system could be detected correctly by both methods.

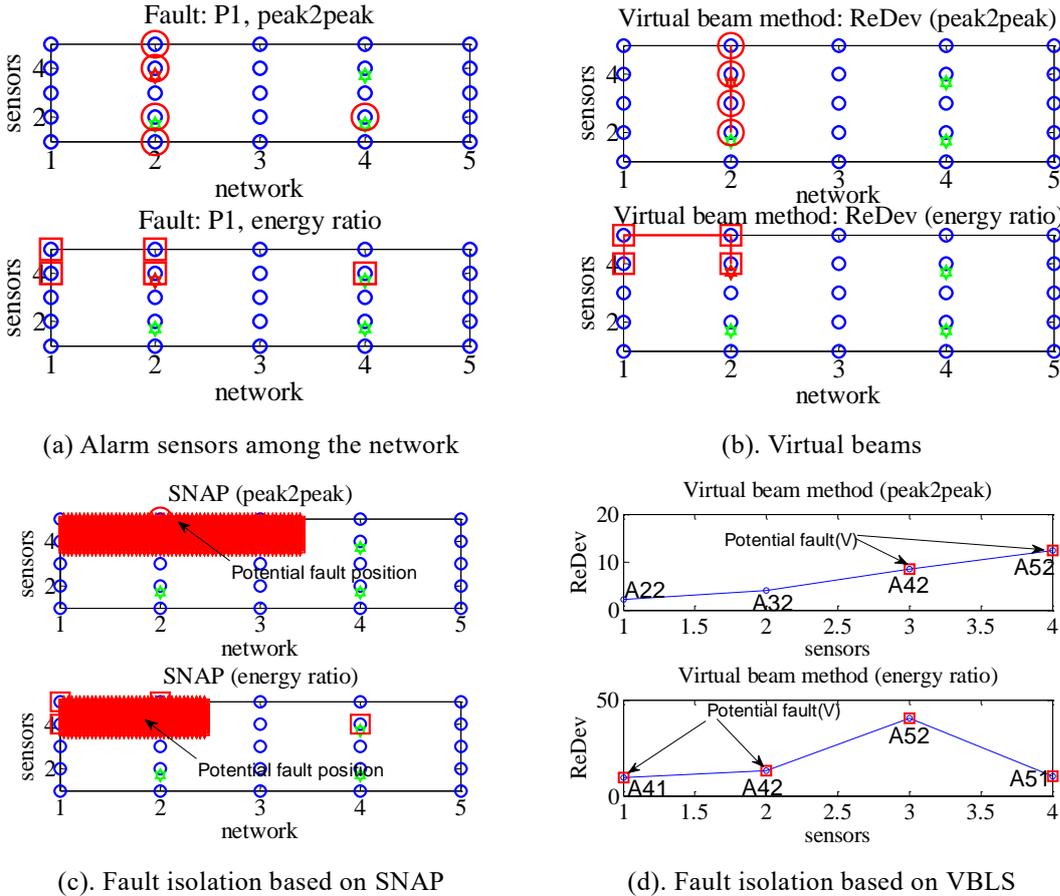


Figure 6-10 VBLs method and SNAP method for isolation of P1 fault.

Clearly, the diagnosis signals measured from the system with P2 fault could be detected correctly by virtual beam-like structures. As shown in Figure 6-11, only one virtual beam was obtained after conducting the optimization process by BAFS. According to Figure 5-6(e), P2 fault was detected and isolated by this virtual beam consisting of sensor chain: A24-A34-A44-A54. However, it was confusing to find different fault localizations when the SNAP method was used. Also, alarm sensors

were mainly located at the fourth column except for sensor A22. The location of P2 could not be identified by SNAP directly because of the distribution of sensor A22. It was obvious that the VBLS method was more effective while locating fault P2 since the construction of virtual beam had considered not only the sensors indicating the faulty system with large fault indicators but also the relative positions of the sensors on the structure.

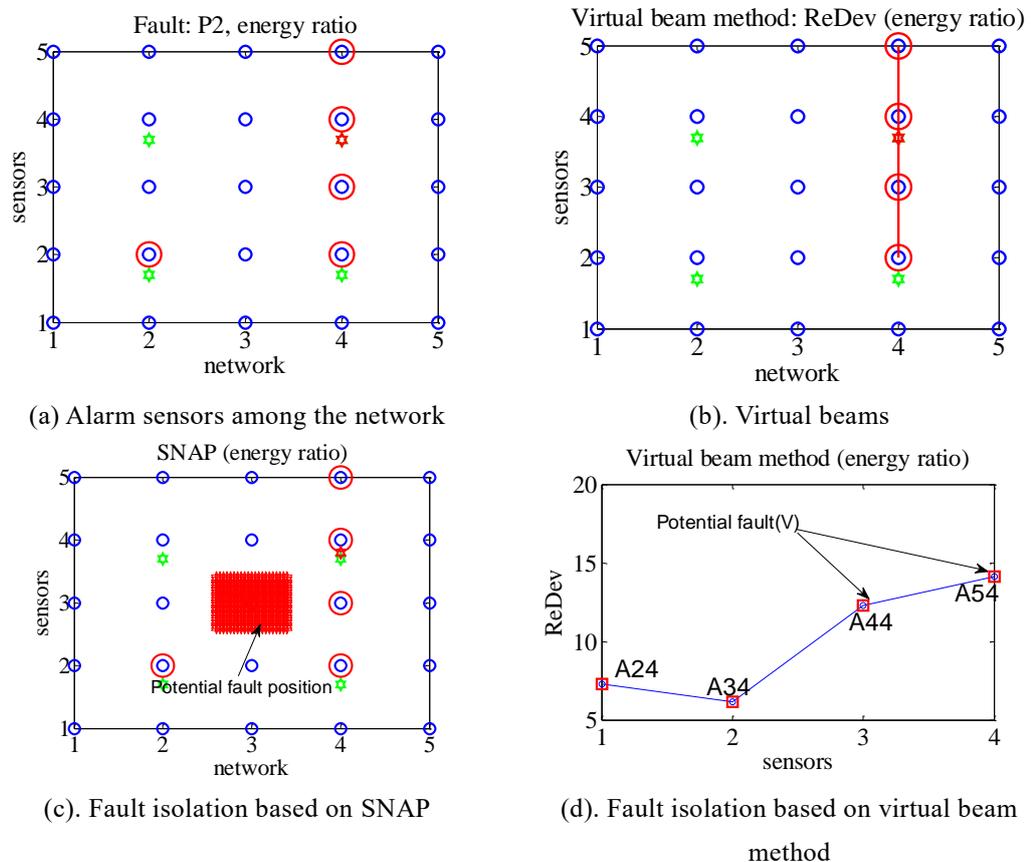
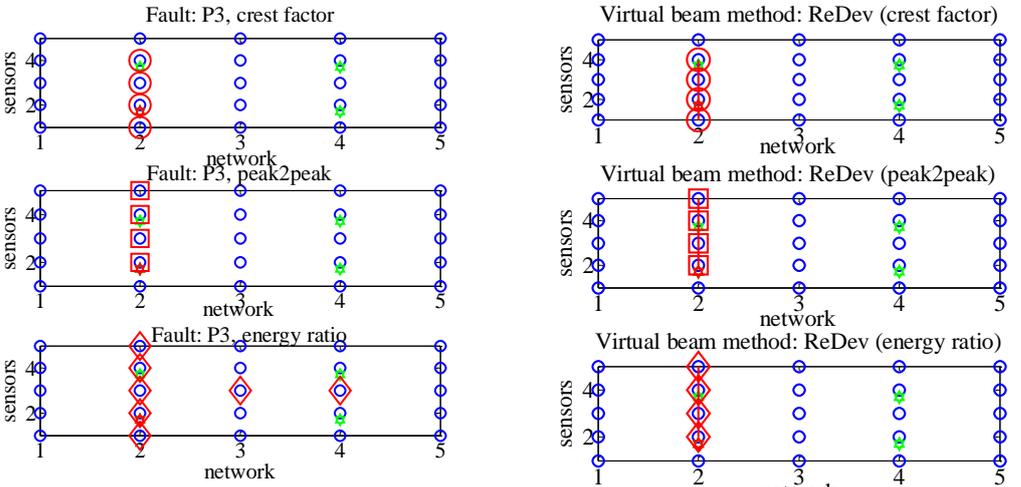


Figure 6-11 VBLs method and SNAP method for isolation of P2 fault.

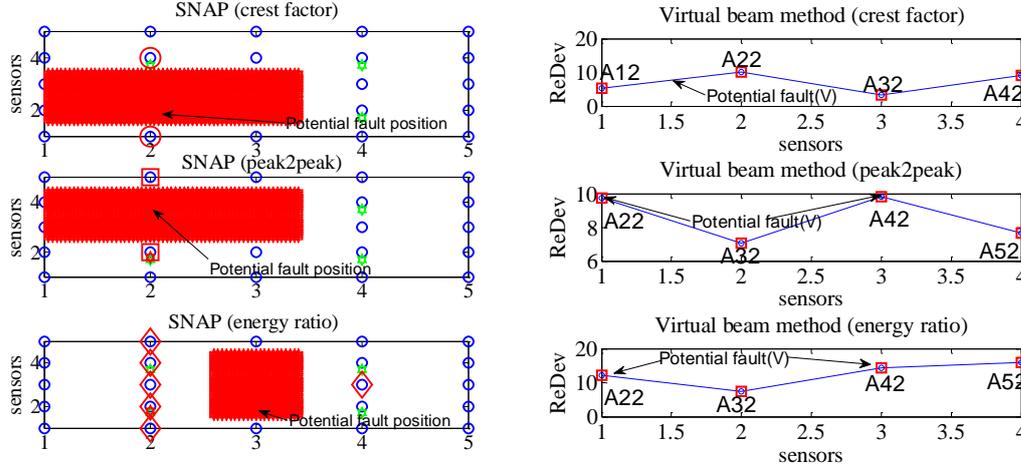
Figure 6-12(a) shows that the alarm sensors were located mainly at the second column of the sensor network when the main body is suffering from P3 fault, which is a typical case found in beam-like structures. According to the first virtual beam (i.e. A12-A22-A32-A42), the trend of fault indicator based on the sensor chain indicate the around sensors A12 and A22, i.e., P3 fault, in Figure 5-6(c). However,

the SNAP method based on the location of alarm sensors could not distinguish the faults of P1 and P3. The alarm sensors distributed in varying positions made it difficult for SNAP method to decide the accurate localization of potential fault.



(a) Alarm sensors among the network

(b). Virtual beams



(c). Fault isolation based on SNAP

(d). Fault isolation based on virtual beam method

Figure 6-12 VBLS method and SNAP method for isolation of P3 fault.

When the diagnosis signals were measured from the system suffering from P4 fault, a virtual beam was obtained using BAFS: A24-A25-A35-A45. According to Figure 6-13(d), the potential fault was positioned around sensors A24 and A35 according to the fault localization methods given in Figure 5-6(d). Considering that

the sensors located in the fourth row and fifth rows were not alarm sensors, the potential fault could not be P2. Thus, the fault was isolated around P4 using the proposed method. However, it was difficult for SNAP to isolate fault P4. As shown in Figure 6-13(a), five alarm sensors were available but distributed in varying positions. As a result, the potential fault position was localized at the region between P2 and P4. Not considering the vibration transmission path and relative positions of sensors had rendered the SNAP method less effective for fault localization.

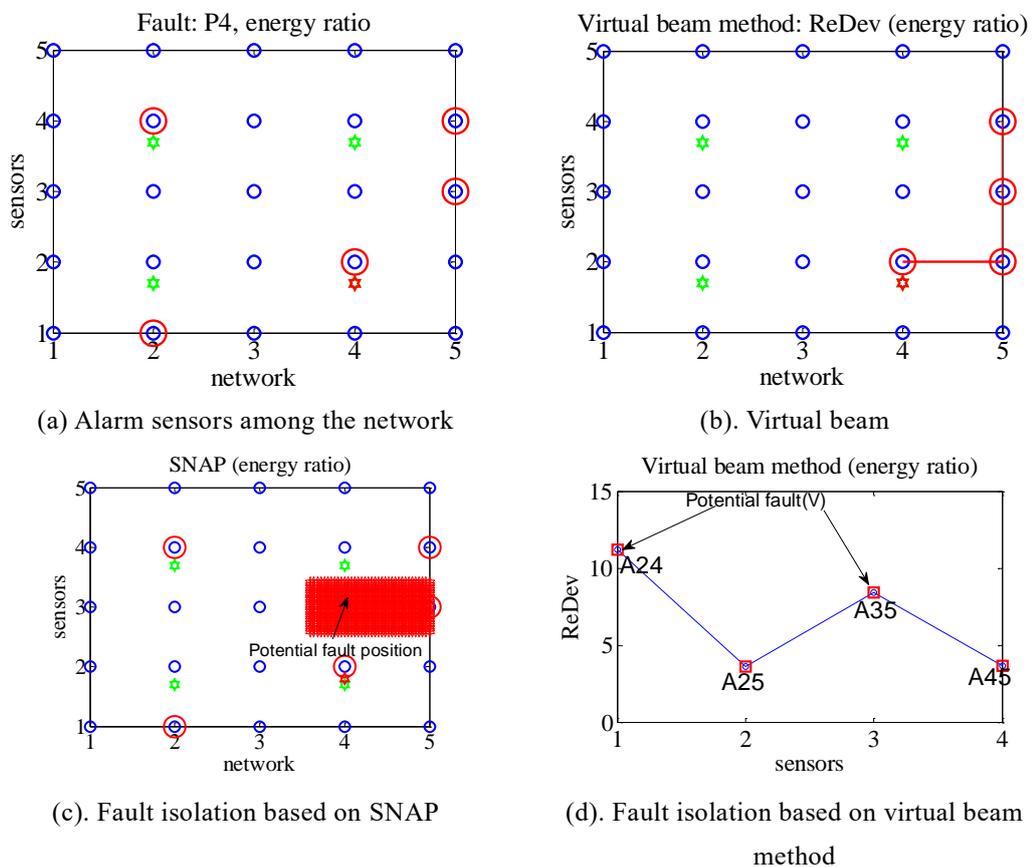


Figure 6-13 VBLS method and SNAP method for isolation of P4 fault.

6.3. Discussion and Conclusion

A virtual beam-like structure method has been investigated above. It is based on the finding that occurrence of the fault on the way of vibration transmission path

will be manifested in the response of sensors on that path. Though vibration transmission in complex structures is quite complicated, the vibration transmission paths consisting of vibration sensors can be found in a straightforward manner by BAFS from networks according to the response of sensors and relative positions between them.

Extensive experimental studies have indicated that this VBLS approach quite well for the considered complex satellite-like structures (e.g., solar panel, bonds, body module, band antenna), and has shown the advantages for fault diagnosis in comparison to the SNAP method, e.g. providing more accurate fault localization when the diagnosis system is detected in abnormal state. The vibration signals can be characterized as time domain features susceptible to uncertain environment noise. Sensors keeping a distance on vibration energy path could also reflect the abnormal condition of the system. Consequently, wrong or ambiguous diagnosis might be obtained using SNAP method.

In virtual beam-like structure method, the vibration transmission path, regarded as ‘virtual beam’, consisting of a chain of sensors (from sensor networks) can be obtained automatically by optimization algorithm, which makes it easier for fault localization by narrowing the region where the potential fault might be located. It may therefore be conclude that, in comparison to SNAP method, the proposed VBLS method is more reliable for fault localization.

Chapter 7. Fault Diagnosis of Complex Structures with Multiple Faults

Though VBLS approach for single-fault diagnosis is studied in Chapter 6, the threshold based fault detection method is adopted with the assumption that prior knowledge of the normal system is available. In this Chapter, the statistical approaches are combined with the adaptive threshold for fault detection, and the VBLS approach is validated for fault localization of both single and multiple fault localization without using the prior knowledge of faults and using the limited prior knowledge of normal operational conditions.

7.1. Problem Formulation

Occurrence of multiple faults indicates the presence of a set of single faults, simultaneously. The number of faults and faulty sites are unknown. The extent of faults is probably different so that the presence of some faults might be masked by others [200]. The occurrence of one fault (or main event) probably brings about the abnormal reflection on numerous sensors, and those sensors suffering from this main event (i.e. kernel fault with most serious influence on the system) might only reflect this fault and ignore the other potential faults. Though some sensors may indicate the potential faults, the response extent of those sensors cannot exceed the sensors distributed near closed to that main events, which contributes to the less detection capability of side faults. Those factors will increase the difficulties for

multi-fault isolation since other potential faults cannot be correctly isolated.



(a) Multiple faults on body

(b) Multiple faults on solar panel

(c) Multiple faults on both body and solar panel

Figure 7-1 Examples of multiple faults on satellite-like structure

Our study focuses on the fault diagnosis of complex structure with one or multiple faults on various components. Take a satellite-like structure as an example. As shown in Figure 7-1(a), the first case is the fault occurring on the bottom where the screws are working for the connection of a cone shaped-joint with the structure body. The second case is the fault located on the top of joint where there is only one screw for each joint to fix the solar panel on the hard, conic-like joint and to further connect to the structure body, which has been shown in Figure 7-1 (b). For multiple-fault cases, the faults might locate at varying substructures. For example, one fault occurs on the main body and one locates on the solar panel, as shown in Figure 7-1 (c). In this thesis, above mentioned three cases will be studied, and an effective and easy handling method based on virtual elements and feature characterization will be presented for fault diagnosis of structures with one fault or multiple faults.

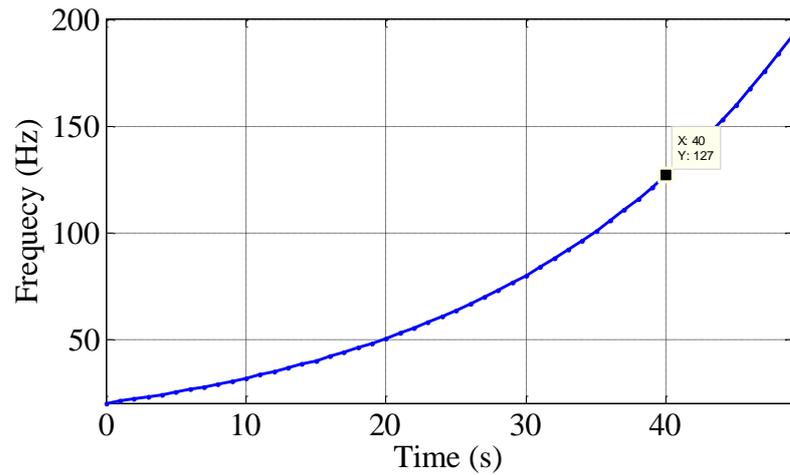


Figure 7-2 Correlation of time and frequency under sweeping-frequency excitation

At the very beginning, the reference signal is measured from the studied system which is in the normal state without any cracks or loosening bolts. Under the same input, another group of vibration signal referred as diagnosis signal is measured when the system is excited at the different moment. For benchmark study, bolt-loosening on the adjacent substructures is studied since it frequently occurs in various hanging structures with bolted-based connected sub-component and can easily implement for experimental study. As shown in Figure 7-2, the input is adopted here as a sweeping-frequency excitation from 20 Hz to 200 Hz with the exciting intensity 0.4g, the sampling frequency 8192 Hz, and velocity 4 oct/min.

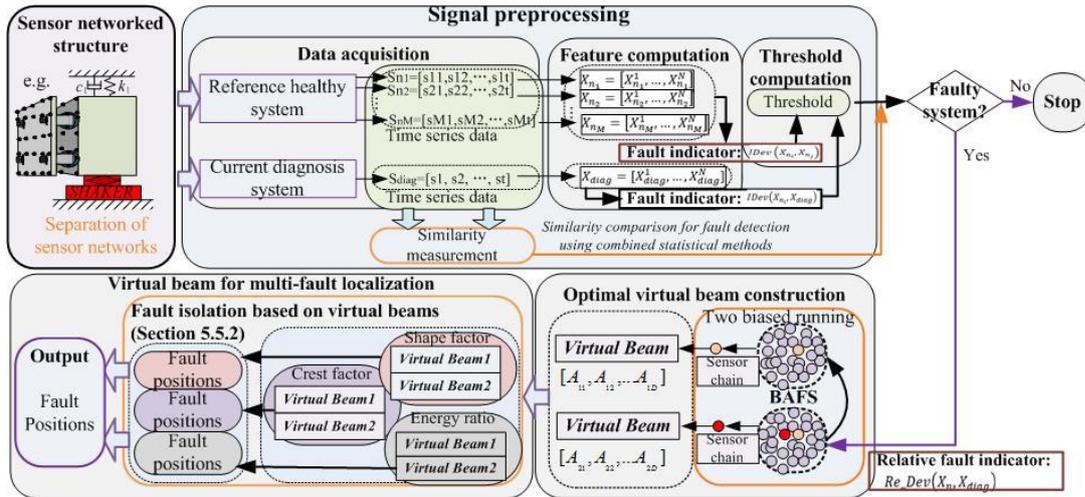


Figure 7-3 The overall framework of virtual beam-like structure approach (VBLS) for multi-fault diagnosis

7.2. Experimental Testing and Results

Different from the experimental tests in Chapter 6, in this section, the VBLS approach is applied to fault diagnosis of complex structures which might be suffering from more than one faults. Figure 7-3 shows the overall framework of the proposed method for multiple fault diagnosis. Biased running strategy is embedded in the optimization process to create multiple different virtual beams for a specific component with abnormal indicating sensors.

Based on SNAP, the decentralized SNAP (i.e. DSNAP) [171] was developed for multiple fault localization, in which all the alarm sensor nodes were the fusion center and tracking the response of their neighbors. Even so, the accuracy of fault localization is frequently decreased because of faulty sensors. To address the negative influence of faulty sensors, the recent Trust Index based Subtract on Negative Add on Positive (TI-SNAP) approach was developed and shown the superior for multiple fault localization in comparison to the DSNAP. In this section,

the proposed VBLS approach is compared with the TI-SNAP [173] to localize one or more faults using sensor network. The studied component like solar panel is divided into a grid g with 100×100 cells and grid resolution is $g = 1$. The radius of ROC is the length of 32 cells, i.e., $R_c = 32$ and $R_c = R_f$. More detail can be referred to [173].

7.3.1. Experimental Platform

To validate the effectiveness of the virtual beam-like structure approach for fault diagnosis of the complex structure which might be suffering from one or multiple faults, the satellite-like structure is adopted. Considering the bolt-loosening fault frequently occurs in BBHSs, the satellite-like structure with one or more loosening bolts is studied in this Chapter.

The same satellite-like structure adopted in previous Chapter (shown in Figure 6-2) is also studied for multi-fault diagnosis. The bolted-base connected structures like solar panel, main body on the satellite-like system are studied, and varying bolt-loosening faults on the satellite-like model are considered and intend to be isolated using the proposed virtual beam-like structure approach. Sensors are evenly (e.g., 5×5) distributed on the structure as shown in Figure 5-3. The sweeping-frequency excitation from 20 Hz to 200 Hz is adopted with the exciting intensity 0.5 g, sample frequency 8192 Hz, and velocity 0.4 oct/min.

7.3.2. Multi-Fault Diagnosis on Solar Panel

The solar panel is connected to the main body by four bolts, namely, L1, L2, L3 and L4. Referring to Figure 5-3, the sensor networks for the virtual beams include the sensors on the solar panel (25 sensors) and sensors on the adjacent connectors (8 sensors). Considering that the solar panel is connected to the main body by four cone-shaped connectors, we can say that the presence of a single loose bolt should have a negative impact on the functioning of the solar panel. Hence only two bolts are studied for early potential fault diagnosis. In the following, experiments on solar panel with one or two loosening bolts (see Figure 7-1(b)): L1 fault, L1&L2 faults, and L1&L3 faults

1). L1 fault condition

Figure 7-4 shows the responses from the sensors on the solar panel and the virtual beams as constructed by the optimization method while the diagnosis signals are measured from the system suffering from loosening bolt at L1. The point figures represent the sensors distributed in the solar panel. The sensors pointing to a faulty system are highlighted in Figure 7-4 (a)-(c). In Figure 7-4 (d)-(f), the sensor chain highlighted using circled marks represents the virtual beam in the first run and another sensor chain highlighted with squared marks in the second run (of the optimization process). The trends of fault indicator associated with the sensors on the virtual beams are also given in the figure below.

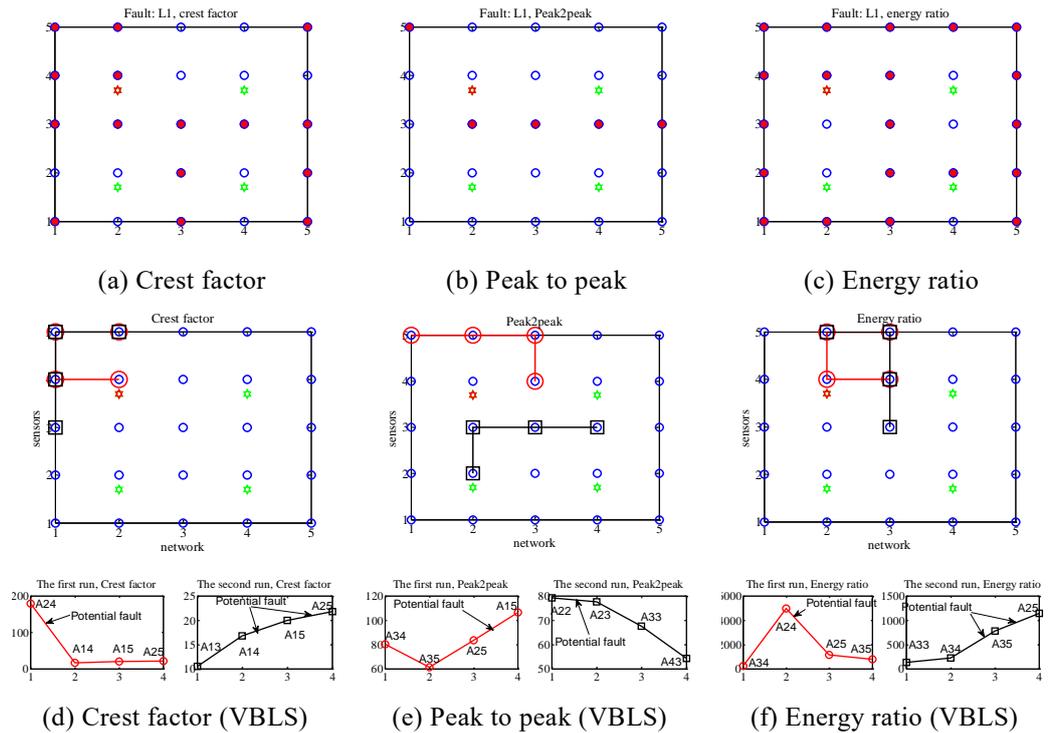


Figure 7-4 Response of sensors from the solar panel and virtual beams constructed by optimization method when diagnosis signals are measured from system suffering from loosening bolt at L1

Based on VBLs, two virtual beams with two sensor chains: A24-A14-A15-A25 and A13-A14-A15-A25 are obtained on the basis of crest factor. The first virtual beam indicates that the fault is around sensor A24 since the largest fault indicator value manifests in sensor A24 and it contributes to the construction of the kernel virtual beam in the first run. According to the second sensor chain (A13-A14-A15-A25), the fault is probably located around sensor A25. Thus, based on two virtual beams, fault L1 located on the top left of the structure is isolated by considering the crest factor.

Similarly, in Figure 7-4 (e), two virtual beams are constructed based on the peak to peak. The first virtual beam consisting of sensor chain A34-A35-A25-A15 indicates that the fault is likely to be located on the top left. Referring to Figure 5-6

where the trend of the second virtual beam is represented by the sensor chain A22-A23-A33-A43, the fault is around sensors A22 and A23. This implies that the fault is probably in L1 or L3 according to the corresponding ROCs. The virtual beams based on the energy ratio show that the fault is most likely located at the top left of the structure. According to the fault isolation rules described in Section 5.5.2, a fault is considered only if it is isolated by at least two virtual beams. As such, only L1 is isolated by the VBLS approach.

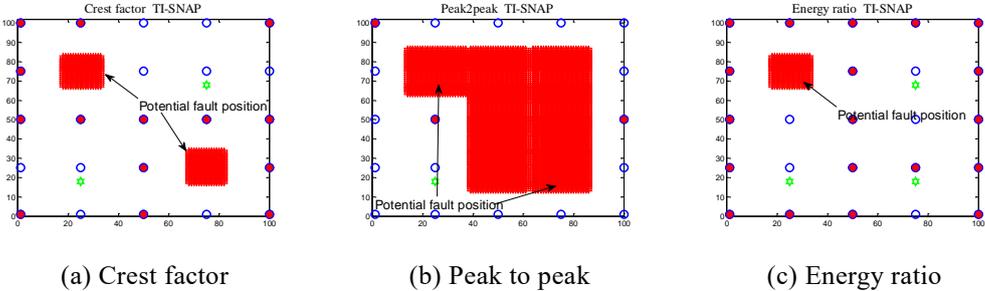


Figure 7-5 Fault localization using TI-SNAP when diagnosis signals are measured from system suffering from loosening bolt at L1

Figure 7-5 shows the potential fault events localized by TI-SNAP. The event L1 at the top left is isolated on the basis of all three features, but event L4 at the bottom right of the component is also identified by crest factor (see Figure 7-5 (a)) and peak to peak (see Figure 7-5 (b)). Thus, according to TI-SNAP, at least two events (L1 and L4) are isolated simultaneously by more than one features, which provides larger regions for potential fault isolation.

In the case of loose screw L1, the VBLS approach shows the advantage to isolate the single fault and provides the more accurate information for fault localization of the potential fault.

2). L1&L2 faults condition

In Figure 7-6, since all the three features have detected a faulty system, they are all considered to be valid and applicable for the purpose of virtual beam construction using the optimization method. The virtual beams obtained using the two biased optimization processes are shown in Figure 7-7.

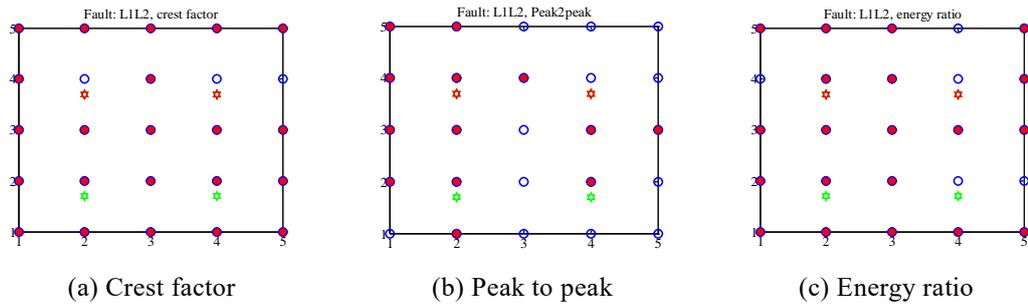


Figure 7-6 Response of sensors on solar panel when it is suffering from loosening bolts at L1&L2

According to in Figure 7-7 (a), two virtual beams A34-A35-A25-A15 and A35-A34-A44-A45, are obtained using the crest factor. According to the first sensor chain, a fault is probably located at the top left of the solar panel, i.e., L1, while the second sensor chain indicates that a fault is present at the top right of the structure, i.e., at L2. Thus, L1 and L2 are both isolated using the crest factor. However, based on the virtual beams on the basis of peak to peak, only L1 is localized as shown in Figure 7-7 while the virtual beams obtained according to the third feature (energy ratio) imply the possibility of multiple faults (L1 and L2). Specifically, the trend of the fault indicated by the first virtual beam (sensor chain A43-A33-A34-A35) indicates that the fault is around sensors A43, A34 and A35, so L2 is most likely to be isolated. Meanwhile, the second virtual beam consisting of the sensor chain A34-A24-A14-A15 indicates that the fault is located at L1. Since both L1 fault and L2 fault are isolated by crest factor and energy ratio simultaneously, it may be

concluded that the system contains two faulty sites, L1 and L2.

In contrast to the VBLS approach, the TI-SNAP approach focuses on the positions of alarm sensors which have been highlighted in Figure 7-8. Based on the regions containing the leader sensors reflecting the faulty system, the faulty sites are localized at L3 and L4 using crest factor, L1 using peak to peak, and L3 using energy ratio. Though the trust index is adopted to assign the weights for the sensor nodes based on their historic records in failing alarm, the harsh condition with limited available prior knowledge makes the poorer performance of TI-SNAP approach for fault localization.

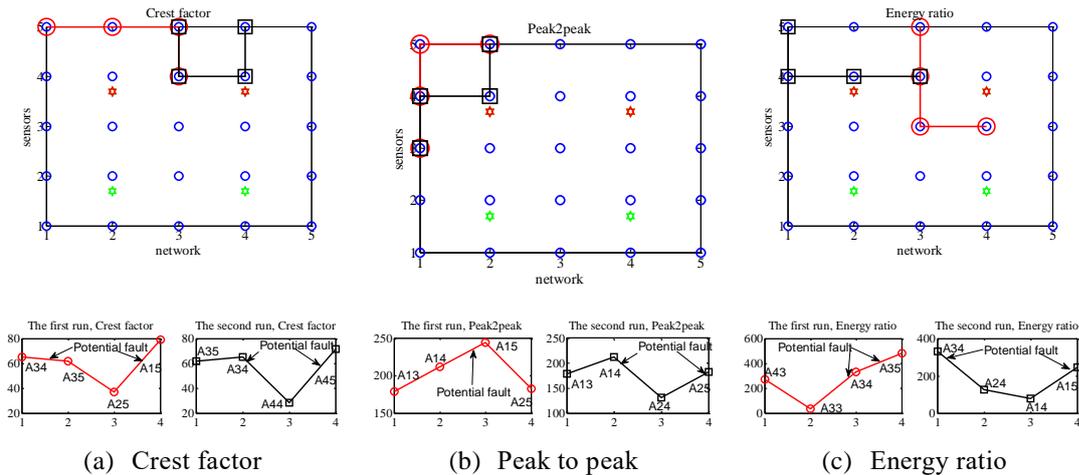


Figure 7-7 Virtual beams constructed using optimization method when the system is suffering from loosening bolts at L1&L2

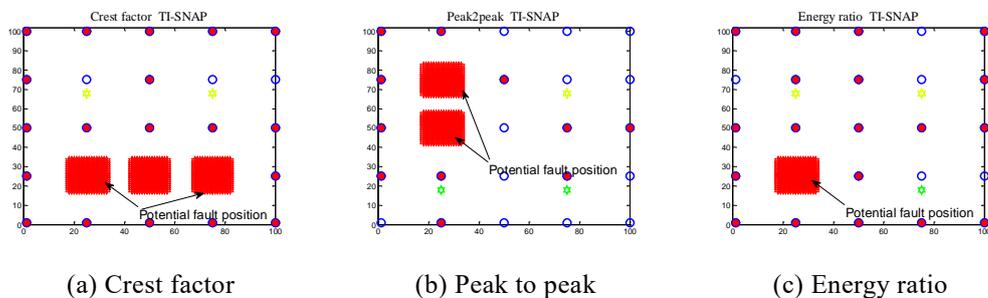


Figure 7-8 Fault localization using TI-SNAP when diagnosis signals are measured from system suffering from loosening bolt at L1&L2

3). L1&L3 faults condition

When the diagnostic signals are measured from the solar panel (owing to loosening of bolts at both L1 and L3), all the three features are capable of fault detection and the sensors indicating the faulty system are highlighted in Figure 7-9. While virtual beams constructed using the optimization method are shown in Figure 7-10. Based on the alarm sensors in Figure 7-9, the potential fault positions are highlighted in Figure 7-11 using TI-SNAP approach.

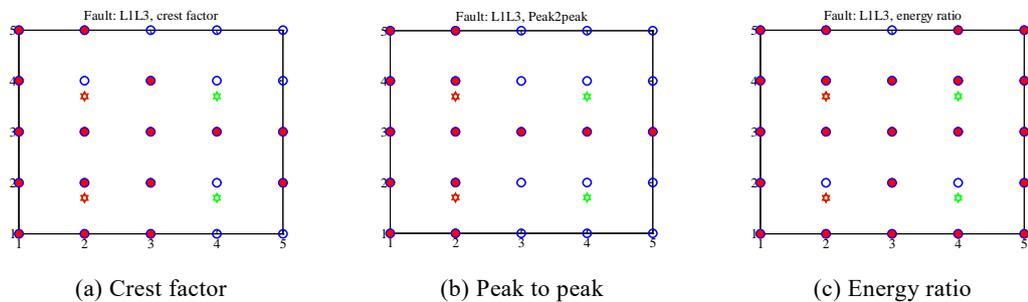


Figure 7-9 Response of sensors on solar panel when it is suffering from loosening bolts at L1&L3

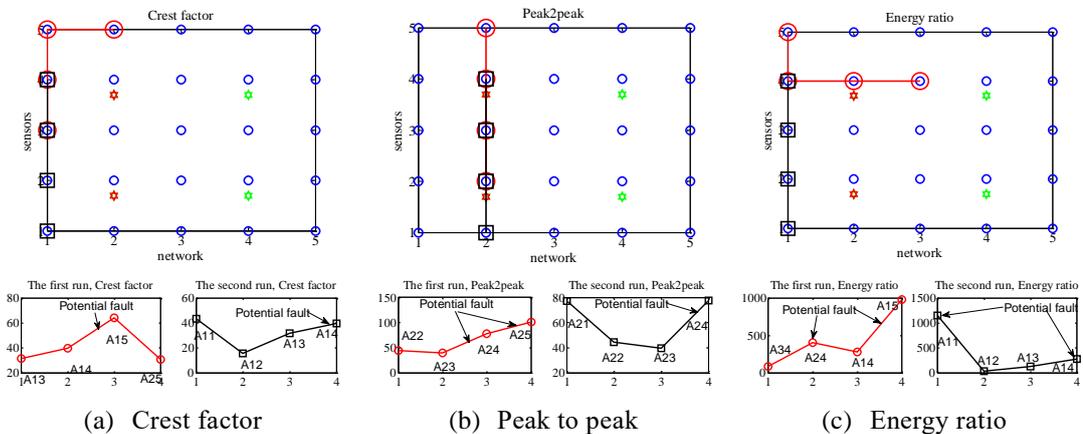


Figure 7-10 Virtual beams constructed using optimization method when the system is suffering from loosening bolts at L1&L3

Obviously, as shown in Figure 7-10, the virtual beams created according to the crest factor indicate that the potential fault of the system is on the left side of the panel, and the virtual beams obtained on the basis of peak-to-peak further lend support to the conclusion from the first feature, i.e., the fault is at L1 and/or L3. The

virtual beams based on energy ratio clearly indicate that there are two faults in the system and that they are L1 and L3. Thus, both L1 and L3 are correctly isolated by the proposed VBLS approach.

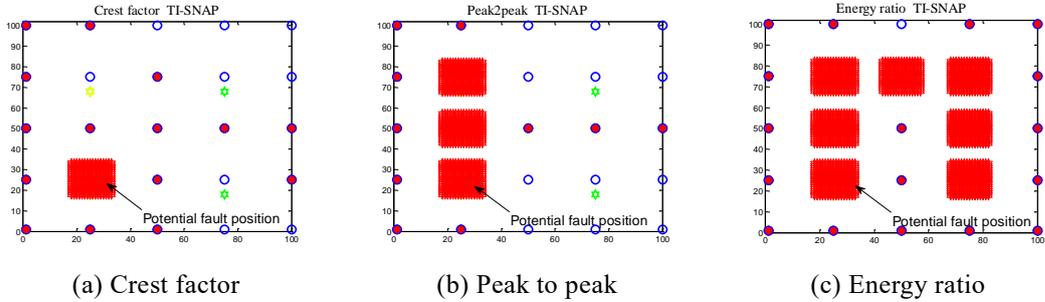


Figure 7-11 Fault localization using TI-SNAP when diagnosis signals are measured from system suffering from loosening bolt at L1&L3

According to the sensors responses based on fault indicator illustrated in Figure 7-9, Figure 7-11 gives the potential faulty events isolated by the TI-SNAP approach. Generally, the fault L3 is isolated by all three features, L1 is localized by the peak to peak and energy ratio. However, the loosen screws of L1 and L3 have brought almost all the sensors on the solar panel indicating the abnormal condition. As a result, all four screws have been isolated by TI-SNAP approach on the basis of the energy ratio.

7.3.3. Multi-Fault Diagnosis on Main Body

In this section, the main body in Figure 7-12 is studied for fault detection and fault localization. This study focuses on the health monitoring of the bolts located on the connection of main body and four connectors. To be convenient, as shown in Figure 7-12, four connectors are named as P1, P2, P3, and P4. Since each connector is fixed using four bolts, our study focuses on early fault detection, i.e., through the

one loosening bolt of four on the connector. For multi-fault cases, loosening bolts are created at two different connectors, e.g. P1&P2, P1&P3.

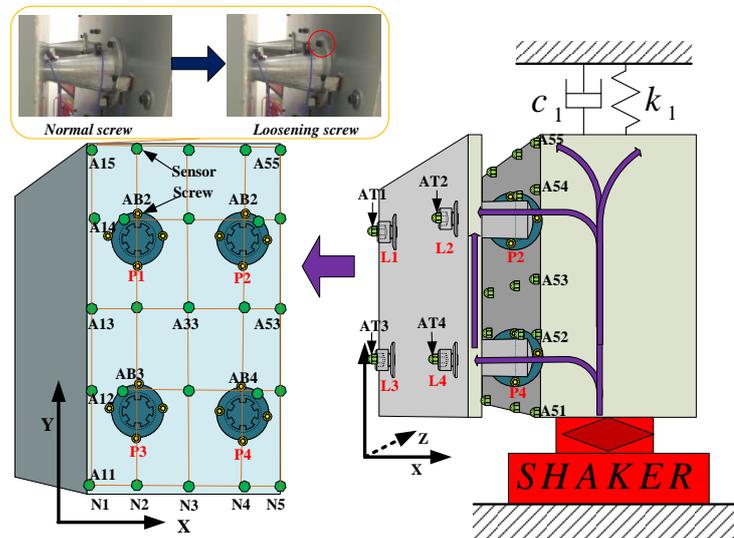


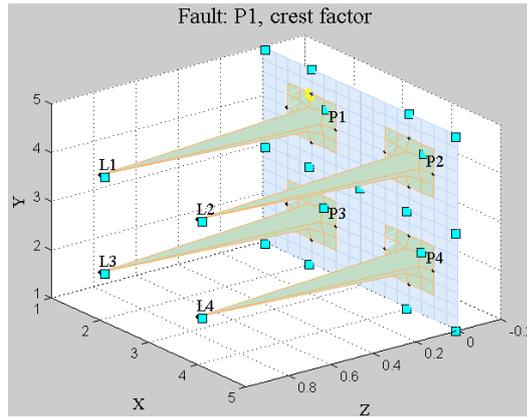
Figure 7-12 Distribution of sensor networks when main body is studied for health condition

Sensors on the main body and on the adjacent substructures (i.e., four connectors) are used for virtual beam construction. Eight sensors are assigned on the four connectors with one on the bottom and one on the top. Sensor on the bottom of the connector P_i is named as AB_i , and the sensor located on the top of the connect P_i is named as AT_i . The sensors on the main body for optimal virtual beams are reduced and distributed with three sensors evenly assigned in each column. In the first column, sensors A11, A13 and A15 are kept for the virtual beam construction, but sensors A14 and A12 are removed since they are closed to sensors AB_1 and AB_3 . For the same reason, the sensors on the third column (i.e. N3) are removed except A33, and sensors A54 and A52 are removed. As a result. There are 21 sensors in total for the construction of the virtual beams.

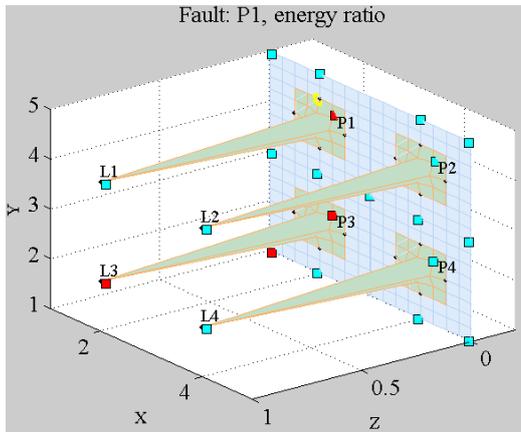
1). *P1 faults condition*

In Figure 7-13, the sensors that can detect the abnormal system are highlighted. Since no alarm sensors have been detected by the methods when the signals are characterized by crest factor, this feature will not be applied to virtual beam construction.

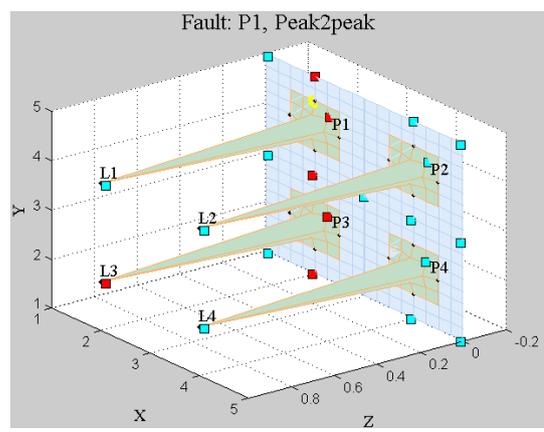
According to the peak-to-peak results illustrated in Figure 7-14 (a), two virtual beams are created consisting of sensor chains A23-AB1-A25-A15 and A13-A23-AB1-A25. Though the two virtual beams are different, both indicate that the fault is probably located at P1. From Figure 7-13 (b), the sensors detecting the abnormal state include A11, AB_3 (in Figure 7-12), AB_l (in Figure 7-12), and AT3 (i.e. the sensor on the top of connector P3). The two virtual beams based on energy ratio do not contain any alarm sensors. A virtual beam is regarded as valid for fault localization only if at least one sensor on the chain has been detected to be faulty according to statistical approaches. Thus, the virtual beams based on the energy ratio are inapplicable for fault localization. Since two further features cannot be used for fault localization, the virtual beams obtained from the peak-to-peak values are employed next. According to the potential fault position from this method, connector P1 is shortlisted for an examination of bolt loosening using VBLS.



(a) Crest factor



(b) Peak to peak



(c) Energy ratio

Figure 7-13 Alarm sensors among the sensor networks on solar panel when it is suffering from loosening bolt at connector P1

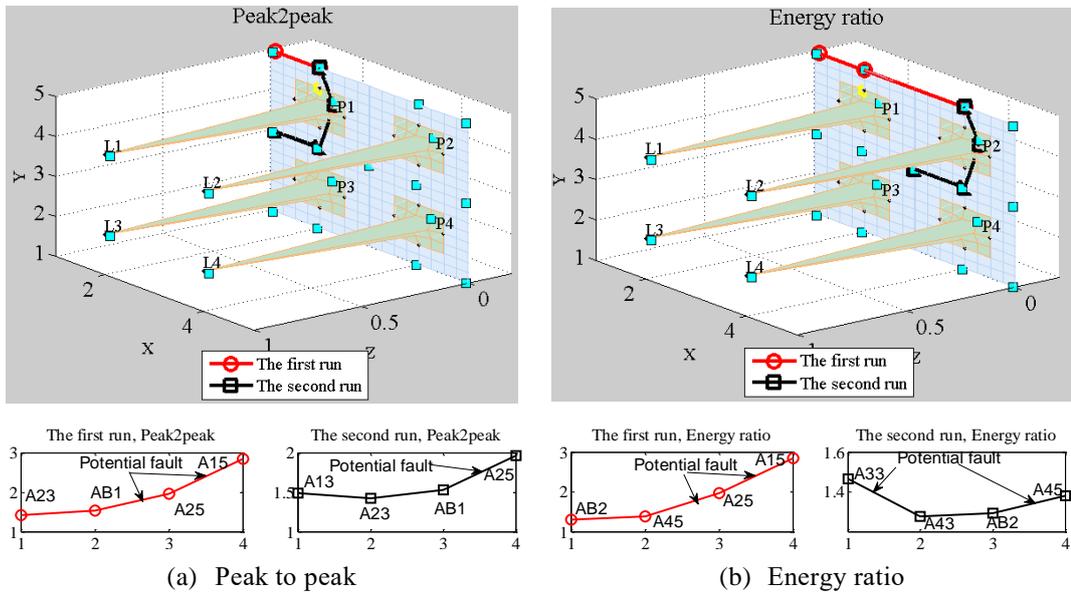
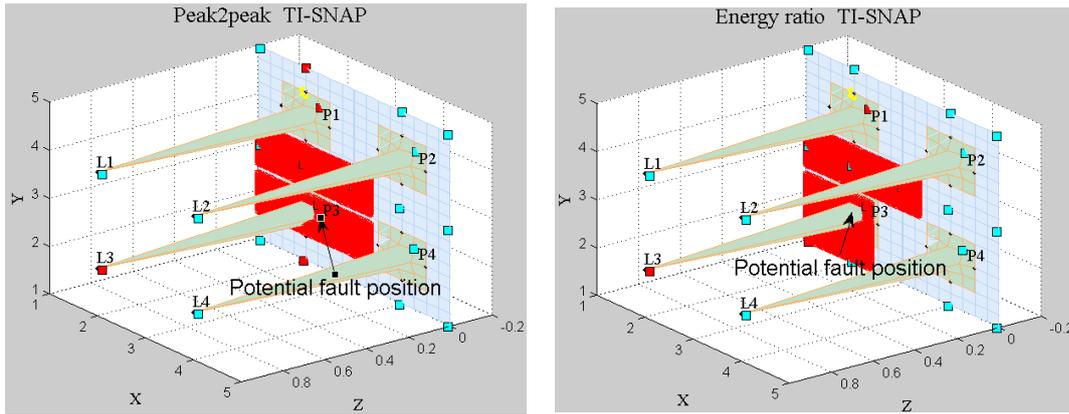


Figure 7-14 Virtual beams constructed using features when the system is suffering from loosening bolt at connector P1



(a) Peak to peak

(b) Energy ratio

Figure 7-15 Fault localization using TI-SNAP when the system is suffering from multiple loosening bolts at connectors P1

Based on TI-SNAP method, the potential faults are probably located at regions with more sensors indicating faulty system. According to the distribution of alarm sensors highlighted in Figure 7-13, the potential fault regions shown in Figure 7-15 suggest that P3 fault is more obvious than P1. Therefore, P3 is the potential faults to be detected for abnormality examination using TI-SNAP method. Compared to VBLS, the regions for the potential faults by TI-SNAP method are more confusing. It is more likely that we need to check a larger region for the potential faults. Instead, the fault localization by VBLS focuses on the regions with the most sensitive sensors while narrowing faulty sites by controlling the length of the sensor chain route.

2). P1&P2 faults condition

The sensors that can detect the abnormal system are highlighted in Figure 7-16 (a)-(c). Accordingly, the virtual beams obtained by “two biased running” are shown in Figure 7-16 (d)-(f). The first virtual beam, highlighted by square marks, are the same for the three features, which indicates the potential fault position of P1 because

the sensor A25 contains the largest value of fault indicator on this chain. Therefore, only P1 is isolated without the strategy of ‘two biased running’ embedded in the original VBLS approach. Based on the crest factor and energy ratio values, the second virtual beam (highlighted by circle marks) indicates that the potential fault is positioned at P2. Meanwhile, the second virtual beam based on the peak-to-peak values in Figure 7-16 (e) implies that the potential fault position is P3. According to the fault localization rules based on multiple features, connector P3 will not be considered as the potential fault. Thus, two connectors, P1 and P2, are isolated and shortlisted for further examination using VBLS.

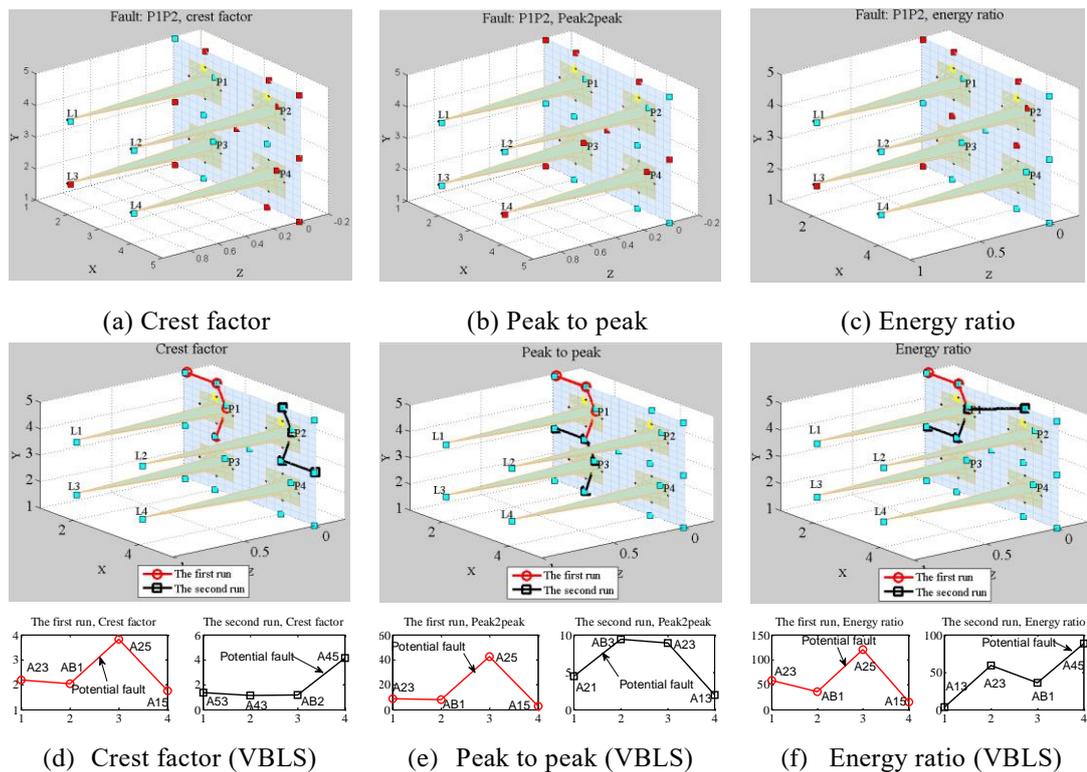
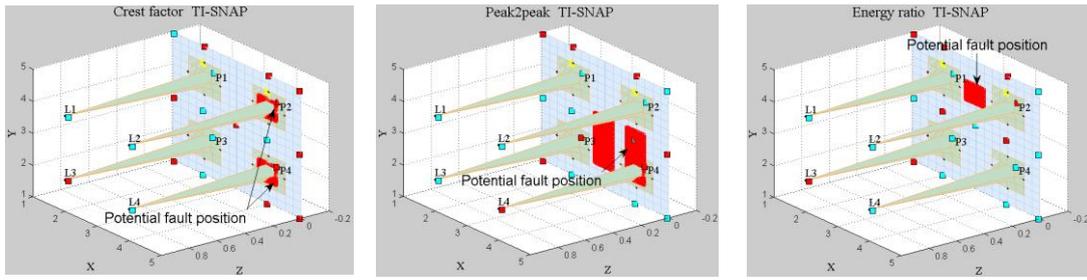


Figure 7-16 Response of sensors and virtual beams constructed using optimization method when the system is suffering from multiple loosening bolts at connectors P1&P2



(a) Crest factor

(b) Peak to peak

(c) Energy ratio

Figure 7-17 Fault localization using TI-SNAP when the system is suffering from multiple loosening bolts at connectors P1&P2

However, it is a challenge to isolate the potential faults using TI-SNAP approach on the basis of the distribution of alarm sensors. According to the alarm sensors in Figure 7-16 (a), connectors P2 and P4 are suspected as potentially faulty sites on the basis of crest factor (see Figure 7-17 (a)). Similarly, based on peak to peak, the region contains the bottom of connector P4 is isolated using TI-SNAP approach. The region isolated using energy ratio in Figure 7-17 (c) is valid. As a result, the potential faults located by TI-SNAP approach are at two regions (i.e., P2 and P4), which are completely identified with the real implement.

3) P1&P3 faults condition

Figure 7-18 show the virtual beams constructed by the candidate sensors from the sensors networks distributed both on the main body and adjacent substructures (the four connectors). Based on the crest factor and energy ratio values, there are two virtual beams: A23-AB1-A25-A15 and A23-AB3-A33-A43. The first sensor chain clearly indicates a faulty site at P1 while the second sensor chain points to a fault site at P3 in view of the significant response from sensor A22. Similarly, based on peak-to-peak value, we can create two virtual beams: AB3-A23-A25-A15 and A23-

AB1-A25-A15, which indicate that the potential fault might be located in connectors P1 and/or P3

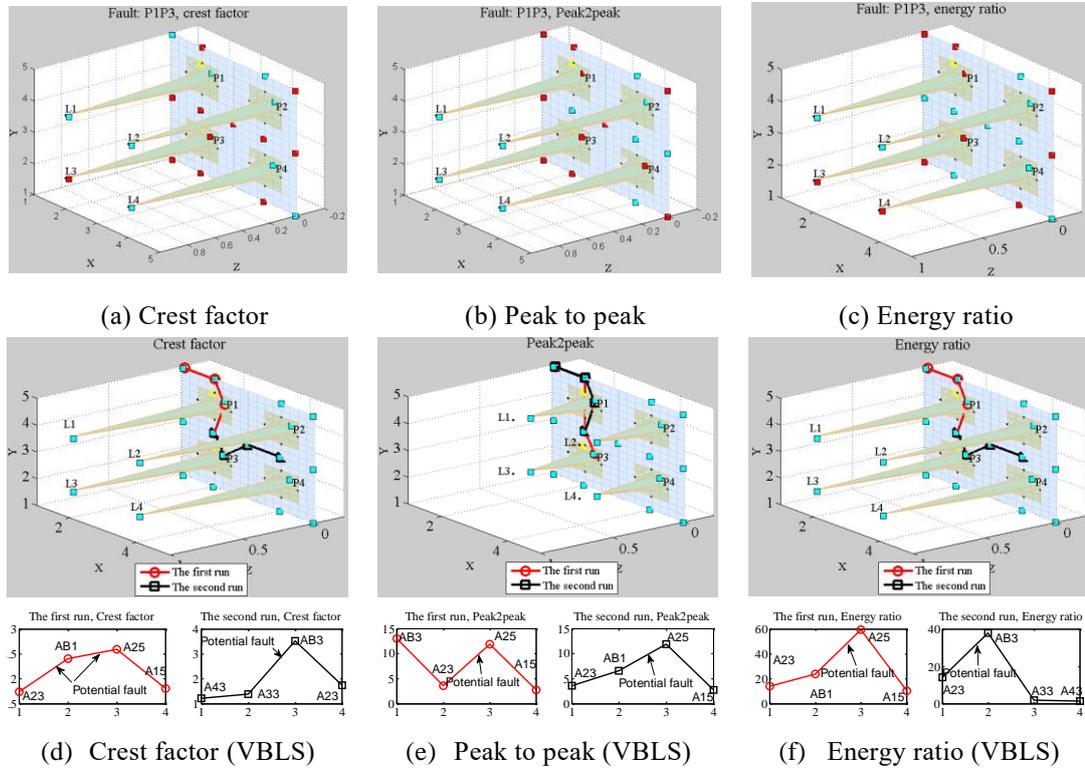


Figure 7-18 Response of sensors and virtual beams constructed by optimization method when the system is suffering from multiple loosening bolts at connectors P1&P3

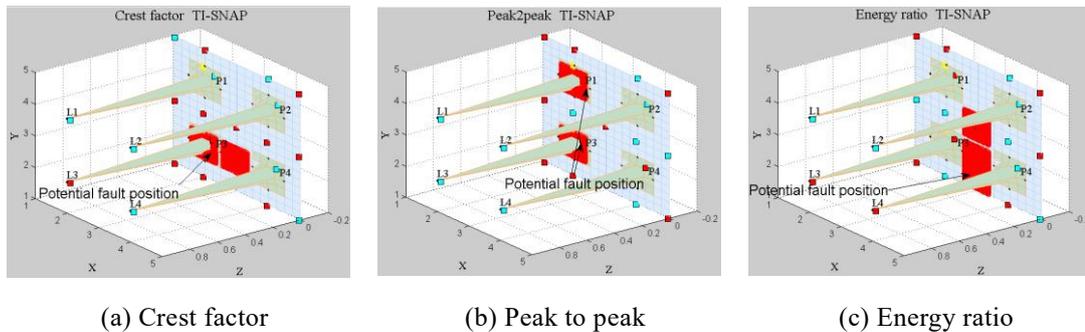


Figure 7-19 Fault localization using TI-SNAP when the system is suffering from multiple loosening bolts at connectors P1&P3

According to the main event virtual beam constructed during the first run (i.e., the original version of VBLS for single fault localization), only P1 is isolated. However, combined with three features and six two virtual beams (with two of them being obtained by each feature after the two biased running of optimization),

connectors P1 and P3 can both be isolated to examine potential faults by VBLS.

According to the distribution of sensors indicating the faulty system, the regions of containing the connectors P1 and P3 are isolated by TI-SNAP. Specifically, the event P3 is localized on the basis of crest factor and both P1 and P3 are localized using peak to peak, but the fusion regions are isolated using energy ratio. Thus, a system with multiple faults at P1 and P3 has also been correctly isolated by the TI-SNAP approach.

7.3.4. Multi-Fault Diagnosis on both Solar Panel and Main Body

To study the effectiveness of the proposed method for localizing multiple faults distributed over different substructures, the system suffering from the two faults located at L1 (one the solar panel) and P3 (on the main body) in Figure 7-1(c) is studied as an example. Similar to the previous two cases, sensors assigned on the connectors could be used for virtual beam construction in both solar panel and main body. Figure 7-20 show the virtual beams constructed on the basis of the sensors located on the main body and the adjacent substructures (connectors), while Figure 7-22 gives the virtual beams using the sensor networks on the solar panel since sensors from adjacent connectors have not been selected by the optimization method. All virtual beams based on the three features are valid and capable of fault localization. Using these virtual beams, it is easy to isolate the potential faulty sites at connectors P1 and/or P3.

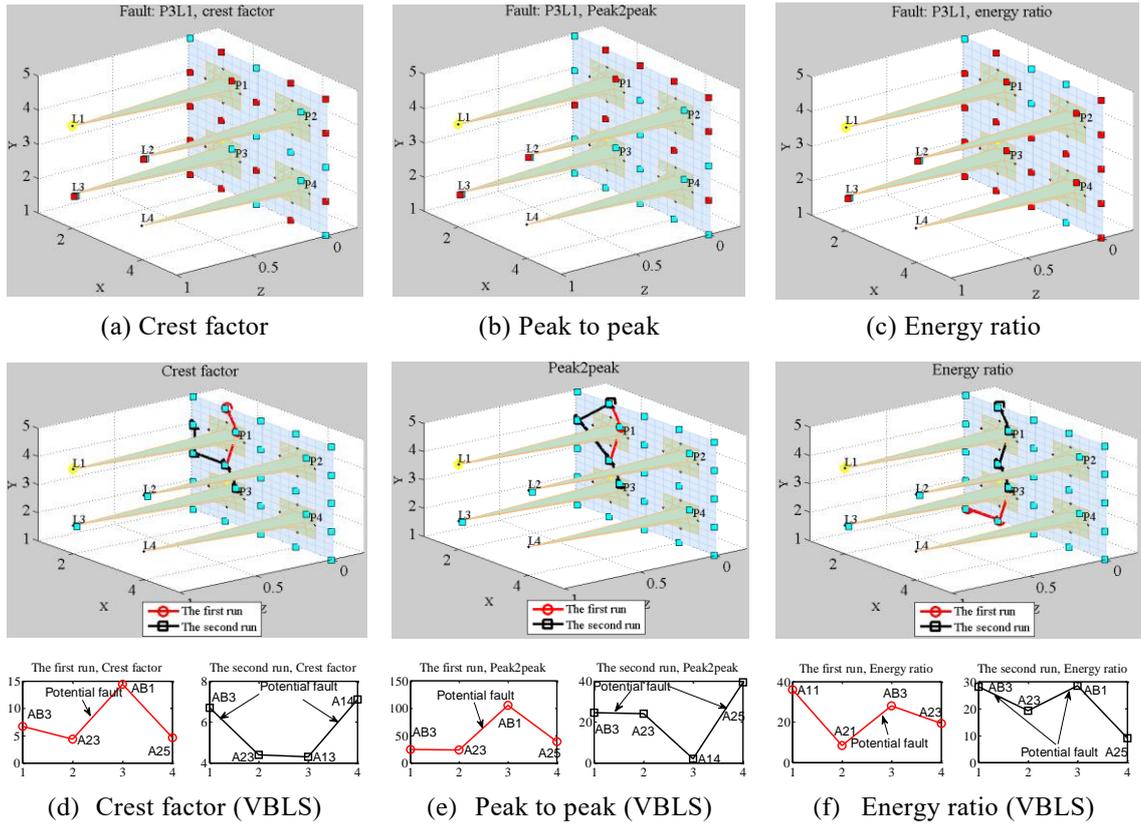


Figure 7-20 Response of sensors and virtual beams constructed using the sensors located on the main body and adjacent connectors

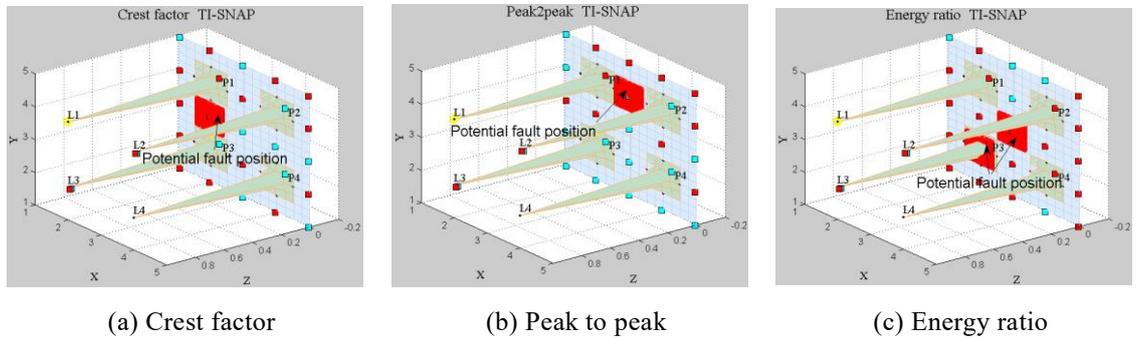


Figure 7-21 Fault localization using TI-SNAP when the system is suffering from L1&P3 faults

Since the alarm sensors are appeared on the solar panel, the sensors responsible for solar panel are employed for fault isolation using optimal virtual beams. Based on the crest factor (see Figure 7-22 (d)) and peak-to-peak (Figure 7-22 (f)), the fault is isolated around sensors A24, A14 and A23. According to the ROC, the event L1 is isolated. Each of the two virtual beams obtained on the basis of energy ratio has indicated a different fault region. Specifically, the kernel virtual beam consisting of

sensor chain A21-A22-A23-A24 reveals that the potential fault position is at L1 (since sensor A24 reflects large value of fault indicator), and the second virtual beam suggests that the potential fault position is L2. Since the detection of L2 is based on a minor virtual beam and no other virtual beam has given a similar fault indication of L2, it is not considered to be a potential fault. Therefore, only L1 are detected by the VBLS approach for the potential fault position using the sensors assigned on the solar panel.

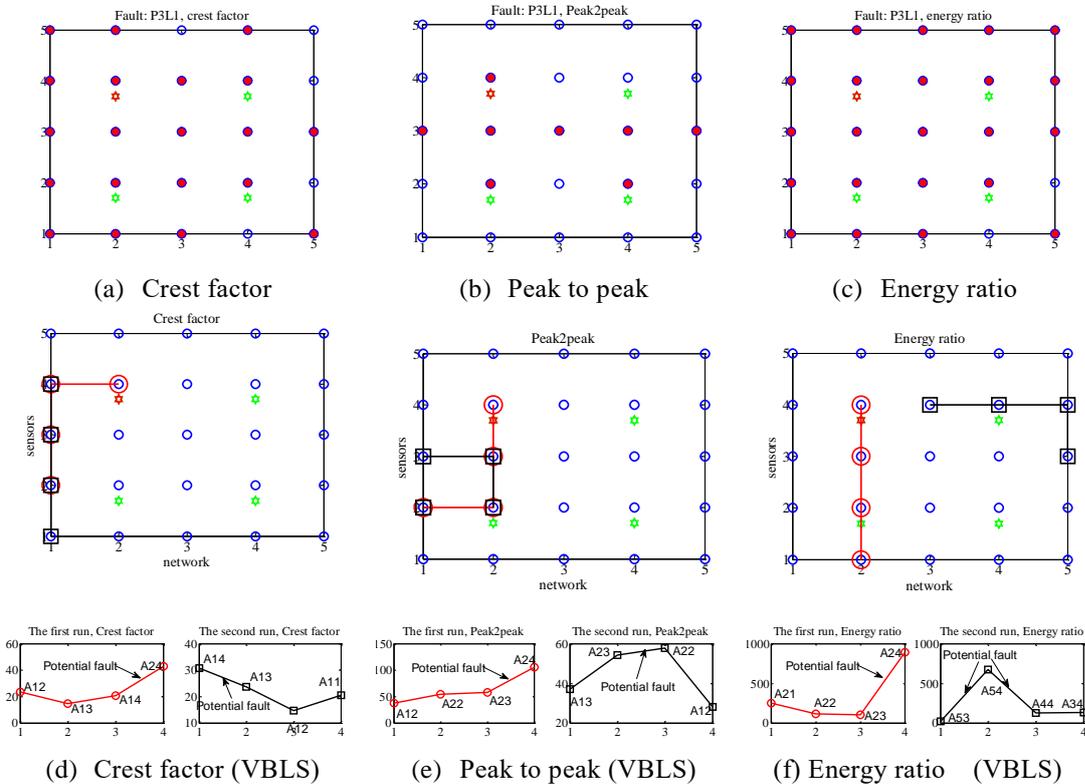


Figure 7-22 Response of sensors and virtual beams constructed using the sensors located on the solar panel

Combined with the information provided from solar panel and main body, the bolt L1 on the solar panel, and connectors P1 and P3 on the adjacent regions of connectors and main body are isolated using the proposed VBLS approach. Considering that bolt L1 is used to fix the solar panel with main body using

connector P1, bolt L1 and connector P1 could be considered to be the same element for the purpose of determining the potential fault position. In other words, the loose screws on the connector P3 and L1 have been correctly isolated by the proposed VBLS approach.

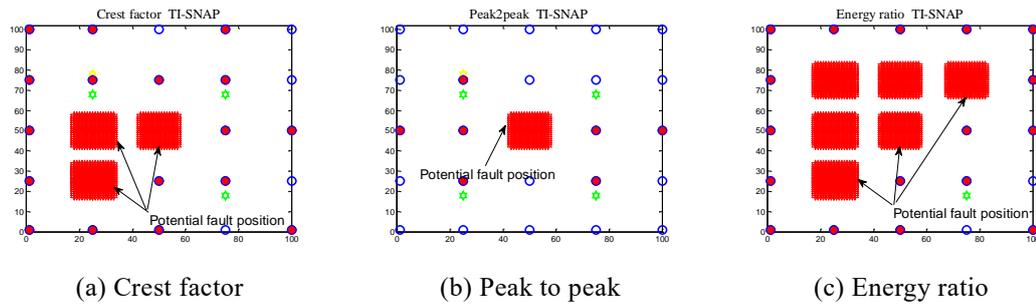


Figure 7-23 Fault localization using TI-SNAP when the system is suffering from L1&P3 faults

Similarly, according to the distribution of alarm sensors on main body shown in Figure 7-21, faulty events are localized at P3 by TI-SNAP approach in Figure 7-21. In addition, Figure 7-23 shows the potential faulty regions on the solar panel. Specifically, loose screw L3 is isolated using crest factor (see Figure 7-23 (a)), and loose screws L1, L2 and L3 are all isolated using energy ratio (Figure 7-23 (c)). As a result, the faults located at P3, L1 and L2 are all isolated by TI-SNAP approach providing wider regions for the potential fault examination.

In comparison to the bolt-loosening fault occurring in the main body, the fault on the solar panel is easier to be detected since bolt-loosening on the solar panel is a rather severe case. Whereas bolt-loosening at L1 can be detected by sensors located on the main body, the loosening screw (i.e., one of four screws) at the bottom of the connector closer to the main body cannot be detected by the sensors on the solar panel. This might be the reason why the loosening screw on connector P3 has not

been detected and isolated by the sensors on the solar panel.

7.3. Discussion

- *In comparison to the classical TI-SNAP approach for fault localization.* Though SNAP shows superior to other fault localization in sensor networks using binary data (e.g. CE, ML, FTML) for its computational efficiency and fewer parameter assumptions [170], it only focuses on the distribution of alarm sensors. The TI-SNAP approach introduces the strategy of trust index to decrease the negative influence of faulty sensors in a network and provides the more accurate localization of potential faults than SNAP using similar computational time [173]. However, it can be reliable for fault localization on the conditions that the prior knowledge or continuous signals from the system are available and the distance between two neighbor events should be far enough to avoid the interference with each other. The proposed VBLS approach considers the vibration transmission paths in a real implement structure and adopts the concept of ‘virtual beam’ to represent this transmission path. The occurrence of the faults could be manifested in the changes of the vibration energy on some transmission paths, which can be captured by some sensor chains. This vibration transmission paths represented by sensor chain are regarded as virtual beams could be automatically optimized by a heuristic algorithm. As a result, the proposed VBLS is easy to implement and computationally inexpensive for fault localization without requirement of prior knowledge of faults. The numerous

experimental results confirm that the VBLS approach is more reliable and effective for localization of potential faults in comparison to the TI-SNAP approach.

- *The effectiveness of sensor network separation.* The increased complexity of signal analysis brought by sensor numbers and sensor information (including the related feature characteristics, the knowledge of positions, and relative distances, etc.) will increase the computational complexity and decrease the accuracy of fault localization. In such a case, the occurrence of a fault in one substructure has little influence on another, the sensors on associated substructures need to be separated. Certain sensors assigned on the adjacent substructures, considered as overlap sensors, can be supplied for the optimal virtual beams construction since they might be capable of detecting nearby faults and then composed into potential vibration transmission paths. Generally, the factors for sensor network separation include: the underlying vibration transmission paths, the influence regions of events, and the response of sensors for the potential faults. Each component is related to only one sensor network, but a sensor might belong to more than one sensor networks. Once a faulty component is detected, the sensors responsible of corresponding component are separated and used as the candidates to construct the optimal virtual beams for the fault localization. The selection of sensor networks based on components could greatly decrease the computational complexity and improve the accuracy

of fault localization, which imposes the proposed fault diagnosis method could be done for online detection (within 30 seconds) and fault localization in a reasonable offline time (no more than 3 minutes in localization for the satellite-like system).

- *The integrated fault detection.* Considering the unavailability of prior knowledge of faults and limited prior knowledge of normal operational conditions, the statistical tests are combined with the adaptive threshold for fault detection. With limited prior knowledge of normal operational condition, the threshold value is difficult to obtain. To increase the reliability, the threshold value is to be updating with the coming of the new data. Additionally, statistical tests are adopted for fault detection by similarity comparison of two datasets to address the harsh condition that only small datasets (e.g. only two or three datasets) are available initially. Therefore, an integrated method is presented in this thesis to take advantage of threshold based method and statistical based methods for fault detection when the prior knowledge of the system is limited.
- *The assumptions of the proposed VBLS approach.* The proposed sensor network based method is applicable to more complex structures if the following assumptions are satisfied. (1) Sensors are distributed on the structure surface ideally with a tree-like topology rooted from the excitation source, and over 80% of them work stably for health monitoring; (2) The topologic information of the sensor networks is available such as the relative positions of each sensor with

respect to others; (3) The assembling information from the experts are roughly known such as the region of influence (of event), region of coverage (of sensors), region of neighbors, and the position of vibration source to identify some basis of vibration transmission paths; (4) At least one reference signal from the healthy operational condition is available working as benchmark representative for fault detection.

7.4. Conclusion

In this Chapter, the VBLS approach was validated for multi-fault diagnosis in the absence of prior knowledge concerning the faults in the system. The fault detection algorithm used in this Chapter adopted the statistical methods (T-test, rank-sum, and K-S) and adaptive threshold to address the small sample (i.e. historic data) limitation. Based on time domain statistical features, the combined fault detection method is computationally inexpensive, which can be implemented for on-line condition monitoring. Different from the classical localization methods concentrating on the distribution of alarm sensors, the proposed VBLS approach utilizes the changes of energy caused by the occurrence of faults for fault localization. Through validation in the contexts of three different abnormal cases, it shown that the proposed approach is more accurate than the classical TI-SNAP approach. It has been demonstrated to be a promising and easy to implement model-free method for multi-fault detection and localization of complex structures with limited or little prior knowledge.

Chapter 8. Procedures and Computational Complexity Evaluation for Fault Diagnosis System

In this Chapter, a detailed procedure for applying the proposed virtual beam-like structure approach to a hierarchic diagnostic system is presented. To establish a comprehensive on-line fault diagnosis system, the computational complexity, knowledge requirements, as well as the applications of the proposed fault diagnosis method are also studied and discussed.

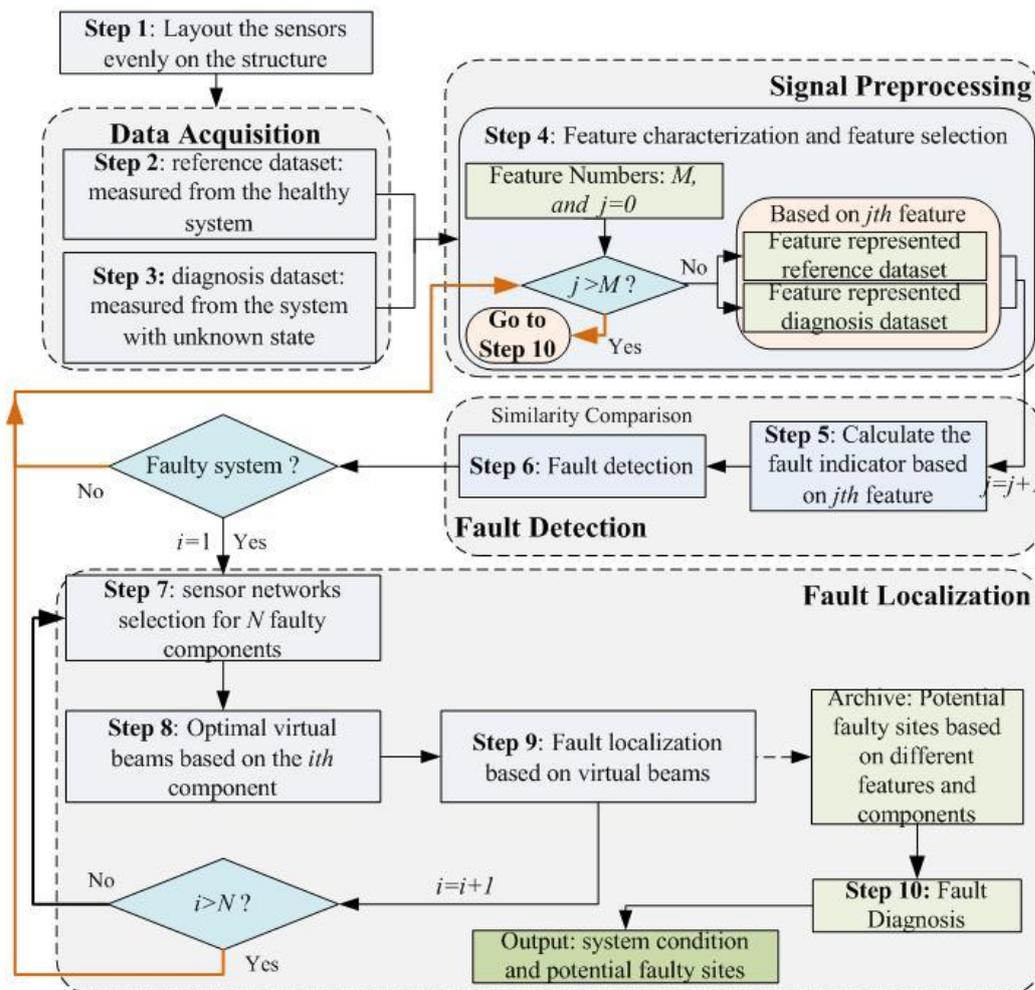


Figure 8-1 Procedures of the proposed fault diagnosis system

8.1. Procedures for Fault Diagnosis

Since the main objective of this study is to develop a novel fault diagnosis with less or little prior knowledge, the data for references are not complete. For that, the method for fault diagnosis of the harshest case using only two datasets (one measured from normal system and one measured from diagnosis system) is presented with the assumption that the reference dataset is reliable and representative for online monitoring. Figure 8-1 shows the overall flow chart of the technical procedures. To apply virtual beam approach to fault diagnosis, the detailed illustration of these technical procedures is provided as follows step by step:

Step 1. Layout the vibration sensors networks.

- 1.(a). If the structure is relatively small in size with no more than 5 sensors:
develop potential effective vibration transmission path (from the vibration source) and assign the sensors along that path for fault diagnosis of structure;
- 1.(b). Otherwise, the sensors are evenly assigned (as networks) in the complex structure (referred to Section 5.2).

Step 2. Measure a group of signals (recorded as reference dataset) when the system is in the normal state.

Step 3. Measure another group of signals (recorded as diagnosis dataset) from the same system with the unknown condition.

Step 4. Feature characterization (referred to Table 4-1) and feature selection (using PSO-FCM, Section 4.1.3).

- 4.(a).** Feature characterization: Signals measured in Step 2 and Step 3 are represented by features (i.e. time domain statistical functions in Table 4-1. The features used for feature characterization could be in the time domain, frequency domain, or time-frequency domain).
- 4.(b).** Feature selection: unsupervised clustering method PSO-FCM is adopted to select the most informative features to represent the time series data. It is assumed that the maximum number of features to be selected is M . Let parameter $j = 0$.
- 4.(c).** Represented by the j^{th} selected statistical feature, both reference and diagnosis datasets are provided for data analysis. Let $j=j+1$.
- 4.(d).** If all the statistical features have been used for data analysis (i.e. $j > M$), then data analysis is terminated; otherwise, turn to Step 5.

Step 5. Calculate the fault indicator.

- 5.(a).** If there are only two datasets, deviation ratio (Dev) is calculated using Eq. (4-10)
- 5.(b).** Otherwise, deviation ratio is calculated according to Eq. (4-11), and relative deviation ratio is calculated using Eq. (4-12)

Step 6. Fault detection using combined statistical methods. (It is assumed that the reference dataset is reliable and representative. *if the knowledge of the

system is available, the combination of threshold method and statistical tests is applicable to fault detection)

6.(a). The faulty state is considered if the abnormal system is decided by any one of statistical tests or threshold method;

6.(b). If the faulty state is detected, then turn to Step 7; otherwise, the system is considered to be in normal condition by j^{th} feature, turn back to Step 4.(c) and make a record in 'Archive'.

Step 7. Sensor networks selection.

7.(a). Fault localization focuses on the components containing the sensors indicating the abnormal state. It is assumed that there are N components in the system containing the faulty information. Let $i=1$.

7.(b). Select the appropriate sensor networks for the i^{th} components (referred to Section 5.2).

7.(c). If $N > 1$, more than one components are applied and considered for virtual beam construction. Otherwise, fault localization focuses on one component.

7.(d). Based on a faulty component, the number of subnetworks is calculated using Eq. (5-7) to determine the iteration times embedded in the 'biased running'(referring to Pseudo-code 5-2) for virtual beams.

Step 8. Construct the optimal virtual beams using appropriate sensor networks in Step 7.(b).

8.(a). Based on the i^{th} failing component, optimal virtual beams (consisting of sensor-based paths) are created using optimization method (referred to Section 5.3). *If the structure is relatively small in size with no more than 5 sensors: develop potential effective vibration transmission path (from the vibration source) and assign the sensors along that path for fault diagnosis of the structure.

Step 9. Fault localization using virtual beams (referred to Section 5.5).

9.(a). Based on the j^{th} feature, the potential faulty sites are isolated according to the two virtual beams obtained on the basis of that feature. The faulty sites associated with the specific feature and component are kept into ‘Archive’.

9.(b). Let: $i=i+1$.

9.(b). If $i > N$, then turn back to Step 4.(c), otherwise, go to Step 7.(b).

Step 10. Fault diagnosis based on an outlier ‘Archive’ (which is used to record the health condition of the system based on feature representatives and associated potential faulty sites).

10.(a). Fault decision: Based on M selected statistical features, any one of the feature indicates the faulty system, the system is suspected for the potential fault(s). If a faulty system is suspected, then go to Step 10.(b); otherwise, the fault diagnosis process is terminated since no faulty elements is detected.

10.(b). Fault localization (referred to Section 5.5). Based on a specific component:

compare the fault positions localized using different features. The faulty sites are considered for a further examination if they meet any one of following cases:

(i). A fault is isolated by at least two valid virtual beams, simultaneously;

(ii). A fault is isolated by only one feature, but it is localized by kernel virtual beam (i.e., obtained from the first run of the optimization process).

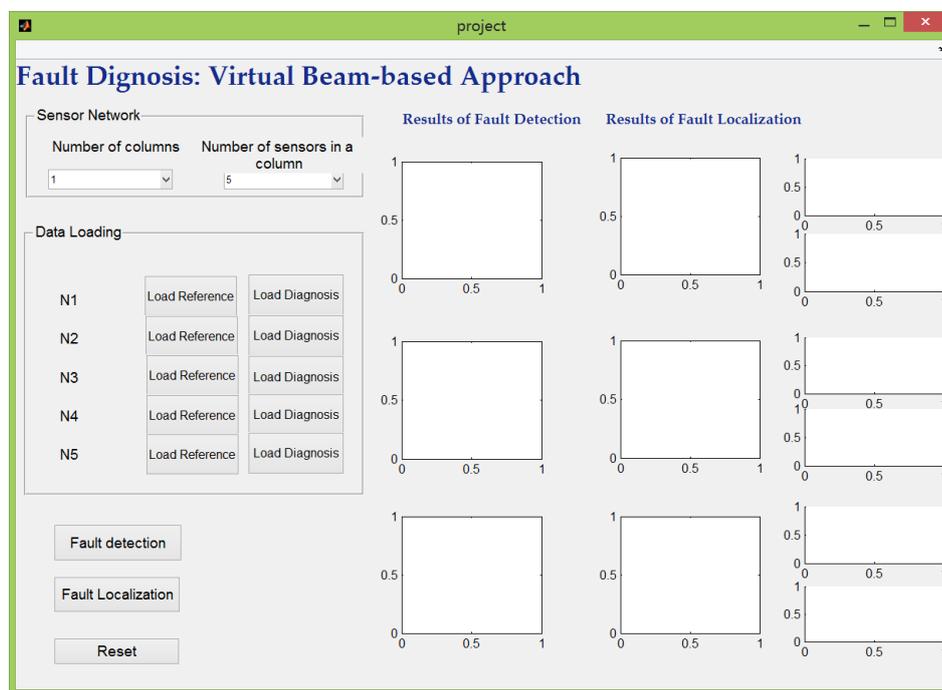


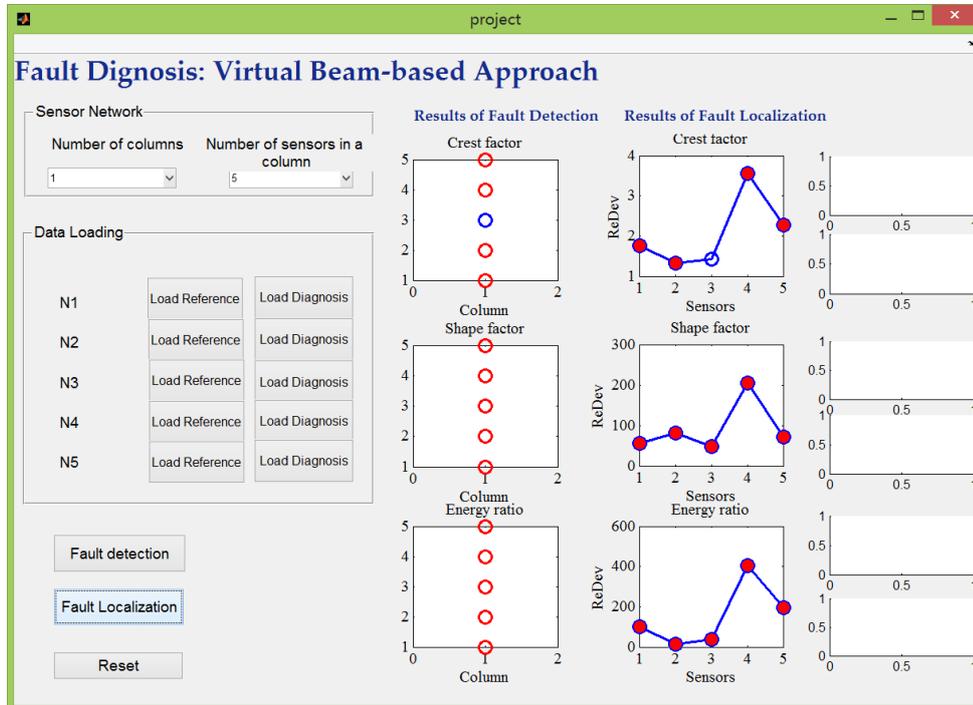
Figure 8-2 The fault diagnosis system

To make the diagnosis system easy to implement in real-world sensor-networked structures, as shown in Figure 8-2, an interface is designed in MATLAB program with GUI front. At the first stage, two parameters related to the information of the studied sensor-networked structures are initialized, i.e. number of columns in a sensor network, and the number of sensors in a column. The second stage is to load the signals measured by the sensors from the network. The original raw time series data are loaded by pressing the buttons: ‘Load

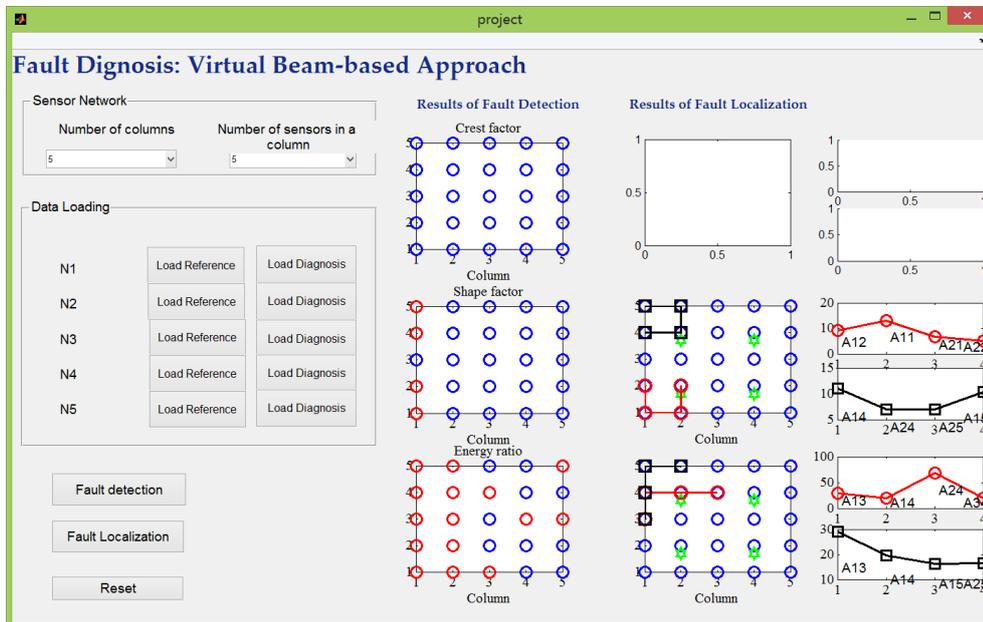
Reference' (the data used as the benchmarks and measured from the healthy system), and 'Load Diagnosis' (the data to be measured), where N represents the column of the sensor network. For example, N1 is the first column. At the third stage, the health condition of the system is to be detected by pressing the button 'Fault detection'. The results of fault detection are shown in the first column of the right region using three figures. The sensors indicating the abnormal state are highlighted in red. Once the system is suspected in the abnormal state, the last stage is to isolate the potential fault by pressing the button 'Fault localization'. The results of fault localization are shown in the remaining columns of the right areas. If the number of columns is more than one column, the virtual beams are constructed and the response of sensors from the virtual beams are shown in the last column of the interface.

Two examples are provided and shown in Figure 8-3. Figure 8-3 (a) shows the results of fault detection and fault localization when the sensor network consists of sensors assigned only in one column. The abnormal condition is detected by both three features since the abnormal state is reflected by most sensors which are highlighted in red. According to the sensors from that column, the potential fault is finally isolated around the fourth sensor according to the rules predefined in Section 5.5. Similarly, Figure 8-3(b) shows the results of fault diagnosis when the sensors are evenly assigned in five columns (i.e. 5×5 matrix). According

to the trends of fault indicators related to the sensors from the virtual beams, the potential faults could be isolated at the right side of the studied structure.



(a) An example of fault diagnosis with only *one* column



(b) an example of fault diagnosis with *multiple* columns

Figure 8-3 Examples of fault diagnosis using the proposed fault diagnosis system

Generally, the interface makes the convenience of fault diagnosis avoiding the complexity of implementation, preventing the error operations, and reducing the capital or time cost for executions to train the special experts to understand the nature of the actions.

8.2. Evaluation of Computational Complexity

Based on the procedures for fault diagnosis of sensor networked-structures, the computational complexity of each step is evaluated according to the CPU record of the PC utilized. In this study, the input consisted of a sweeping-frequency excitation (20 to 200Hz) with an exciting intensity of 0.4g. The sample frequency was 8192Hz and the velocity 4 oct/min. To eliminate noise disturbance at the beginning of measurement, the first 40% series were removed and the remaining 60% (ranging approximately between 80Hz and 200Hz) were used for data analysis. Thus, the length of the time series was around 401472 and the length of data used for analysis about 240883. If the time interval window used for feature characterization was 1 second, the dimensionality of features based on each time domain feature function was around 20, which greatly decreased the computational complexity for data analysis.

It is assumed that the number of feature points is S , a complex structure consists of N components, and there are H accelerometers mounted on each component. Thus, in the worst case, the asymptotic complexity is $O(HMNS)$ using the classical sensor network based localization methods (e.g., SNAP). However, the upper bound is

improved to $O(HMS)$ in the proposed VBLS approach since the fault localization in the proposed method is fulfilled using virtual beams in a faulty component rather than all the sensors.

Table 8-1 Time complexity of the proposed fault diagnosis system

No.	Procedures	Computer run time (second)	Accumulated time (second)	Notes
1	Data acquisition	0	0	80Hz-200Hz; Length of raw data is 240883
2	Feature characterization	2.5663	2.5663	three features, 25 sensors
3	<i>Fault detection</i>	0.2411	<i>2.8074</i>	Statistical methods
4	Optimal virtual beams	48.3418	51.1492	The first feature
5	Faulty sites based on the first feature	0.0001	51.1493	The first feature
6	Optimal virtual beams	48.5011	99.6504	The second feature
7	Faulty sites based on the second feature	0.0001	99.6505	The second feature
8	Optimal virtual beams	48.4487	148.0992	The third feature
9	Faulty sites based on the third feature	0.0001	148.0993	The third feature
10	<i>Output the potential faulty sites</i>	0.0010	148.1003	Considering the combined features

Take the satellite-like structure as an example. There are 25 sensors evenly distributed on the structure for health monitoring, and three time domain features are employed to represent the high-dimensional raw data. Only one substructure (i.e., solar panel) contains one or more faults. The tests were carried out under a 64-bit MATLAB 2012b environment with a computer of 3.4-GHz CPU and 8-G RAM. Table 8-1 gives the average time needed for each stage of the proposed fault diagnosis system over 30 independent runs. Since an advantage of time domain features is its simplicity and interpretability (these avoid the complexity of the

preprocessing), the method is inexpensive in terms of time despite the complexity of feature characterization. As a result, the time used for fault detection is *no more than 3 seconds*; most of time is spent on feature characterization. Thus, the proposed method can be used for on-line fault detection in sensor networks.

If no fault is found by the combined statistical methods and threshold, then system is considered in the normal state and it is not necessary to use the virtual beams for fault localization (after Stage 3). Once the abnormal condition is detected by sensor network, the potential faults are localized using virtual beams obtained by applying the optimization method described earlier. Though the availability of information on more features could greatly enhance the accuracy of fault detection, the computational complexity is increased at the same time. From the data on No. 4, No.6 and No.8 procedures, we can see that the entry of one single feature leads to additional time consumption (around 48seconds) for the optimal virtual beams. To isolate the potential faults within 3 minutes, no more than three features are suggested to represent the time domain signals.

8.3. Application of the VBLS Approach

The virtual beam-like structure (VBLS) approach currently focuses on the fault diagnosis of the bolted-base hanging structures with one or more bolt-loosening faults since they frequently occur in various hanging structures with bolt-base connected with the other structures.

The proposed VBLS approach is applicable to fault diagnosis of complex

structures (e.g. bridge or concrete structures) provided that the following are available:

- 1). Sensors are evenly distributed on the structure, and over 90% of them work stably for health monitoring.
- 2). The topologic information of the substructures and sensor network on each sub-structure needs to be available, for example, the rough geometric position between two sensors on that sub-structure.
- 3). The assembling information from the experts is roughly known such as the region of influence (of event), region of coverage (of sensors), region of neighbors, and the position of vibration source to identify some basis of vibration transmission path.
- 4). At least one reference signal from the healthy operational condition is available and representative for fault detection.

The performance of the proposed method for fault detection greatly relies on the reliability information contained in signals measured before and after. Thus, smaller collection gap between the two measurements can help on-line detection or fault diagnosis with limited priori.

8.4. Advantages of the Proposed Fault Diagnosis Method

The virtual beam-like structure approach (VBLS) has been validated for multi-fault detection and localization of bolt-based hanging structures (BBHSs) while

demanding limited or no prior knowledge concerning the faults present. The advantages of the proposed VBLS approach are summarized below.

1) *In comparison to model-based methods:* As demonstrated earlier, the reliability of model-based methods for fault diagnosis is determined by multiple physical parameters or mode shapes arising under different operation conditions. However, since the system is rather complex with internal state variables that are inaccessible to measurement by sensors commonly used in time-varying dynamic systems, it is not appropriate to apply model-based approaches to fault diagnosis [3]. However, as a model-free method, the method proposed in this thesis has no such limitation; it is also applicable to the fault diagnosis of structures without prior knowledge of faulty system.

2) *In comparison to sensor network-based fault diagnosis methods:* Sensor networks have been used widely to detect the abnormal vibrations and thus prevent structural damage. The methods developed using sensor networks have been used mostly for estimating the parameters associated with the modeling of structures for fault diagnosis [201-204]. The major limitations of these methods include high computational cost and requirement of system off-line during fault diagnosis. Though the binary methods like CE, ML, FTML, and SNAP have shown the superiority in terms of computational efficiency and parameter assumptions [170], they only focus on the distribution of alarm sensors. As a result, the accuracy of fault localization would greatly depend on the reliability of sensors mounted on the

studied structures. Instead of modeling the structures or monitoring the alarm sensors, the method proposed in this thesis selects sensors with large fault indicator values to capture the most sensitive changes on the vibration transmission using the optimization methodology. The machine learning based optimization method embedded is robust and capable of selecting sensor nodes with limited human intervention, which leads to rather low computational cost, and it is more reliable and effective for localization of potential faults.

3) *In comparison to feature based methods:* Since the advantages of time domain features have already been discussed in Chapter 2, further discussions of frequency domain as well as time-frequency domain feature based methods will not be engaged. Regardless of frequency range, time domain features selected need to be sensitive to damage. A typical example of fault diagnosis using time domain features has been presented in [19], in which a multi-sensor system based on feature-level fusion method is proposed for fault diagnosis rotating machinery. Though [19] had adopted multiple sensors, the sensors were analyzed independently using a support vector machine (SVM). Compared to the multi-sensor feature-level fusion method, the method proposed in this thesis is more capable of fault diagnosis by utilizing comprehensive knowledge of virtual beam consisting of sensors along the vibration transmission path. Further, fault localization is also realized by analyzing the virtual beams selected. Literature contains many more studies based on time domain features fault diagnosis [14, 18, 20, 21]. However, a common issue in these methods

concerns the need for classification while classification methods such as SVM require training samples which are not always available in real application. By contrast, the method proposed in this thesis requires no prior knowledge of the fault system, and the object to fault diagnosis is the general structure rather than a specific one, which opens the door to extensive application of the proposed method.

Chapter 9. Conclusion and Prospect

9.1. Conclusion

In this study, a novel method for the fault diagnosis of complex structures based on an optimized virtual beam-like structure approach has been presented for situations in which the system is suffering from one or more faults. This approach provides a new idea of fault detection and fault isolation by constructing the virtual beam-like structures utilizing vibration sensors placed on the complex structure along the vibration transmission path.

In the preprocessing stage, fault indicator-based adaptive threshold method and statistical tests are combined for fault detection. Threshold method is applicable to fault detection when the prior knowledge of the healthy system is available. Meanwhile, to address the issue of limited knowledge of the normal operational conditions, statistical tests (T-test K-S, and rank-sum) have been adopted for fault detection by similarity comparison of two datasets (i.e., the reference dataset and the estimated dataset) since these are sensitive to minor changes in the data series. The time series signals are segmented and the interval features are characterized by time domain statistical features, which enable more exact fault information to be obtained for on-line fault detection.

Once the abnormal condition is detected in a component, a bacterial based intelligent algorithms is provided for the construction of the effective virtual beams

automatically. To address multi-fault diagnosis, biased running is developed in the optimization methodology by constructing the multiple biased virtual beams so as to isolate the potential faults in a faulty component. Further, separation of sensor networks based on components and search space has been considered to arrive at optimal solutions, minimize the computational complexity and maximize the fault localization accuracy.

The rules of fault localization are based on virtual beam stemmed from the findings about beam-like structures: (1). Cracks located closer to the vibration source brings about larger relative deviation ratio of all sensors. (2). Cracks located far away from the vibration source result in large relative deviation ratios of sensors close to the crack or on the transmission path. (3). Sensors located close to the vibration source are less capable of detecting cracks located far away from the vibration source. These findings related to the response of the sensors on the associated virtual beam are taken to represent general rules for fault localization.

Following validation using a satellite-like structure with both single fault and multiple faults (Typical faults occurring more frequently in complex structures, i.e., bolt loosening around the connecting rods), the proposed method has been confirmed to be effective for multi-fault localization through the selection of the most sensitive sensor chains and narrowing regions for potential faults. The findings in the beam-like structure are also applicable to virtual beams embedded in complex structures. Compared with classical sensor network based fault localization method,

the proposed virtual beam-like structure method has proved to be more reliable for fault localization. It has been demonstrated to be a promising and easily implemented data-driven method for multiple-fault detection and localization of complex structures with less prior knowledge requirement.

The main advantages of the proposed methods are: (1) it is capable of both single fault and multiple-fault diagnosis of complex structures without using prior knowledge of the faults, and using only limited prior knowledge of normal operational conditions; (2) it presents an optimal fault localization method based on the concept of the virtual beam, which is computationally inexpensive and applicable to on-line condition monitoring using sensor network; (3) it is more practical and easy to implement in real-world sensor-networked structures with less parameter assumptions and no limitation on stationary data.

9.2. Prospect

This study focuses on the fault diagnosis of multiple loosening bolts in a complex system with numerous bolt-based hanging structures. Though it has been shown to be capable of fault diagnosis of bolt-based hanging structures with one or more loosening bolts (no more than two loosening bolts in a component), there are still several issues worthy of further explorations as follows.

- Studies on different natures of faults.

This study focuses on the fault diagnosis of multiple loosening bolts in a complex system, future work will focus on the development of the method

for fault diagnosis of complex structures with different natures of faults like fatigue cracks [34, 205], fractures [1, 41], and hybrid faults (e.g. both cracks and loosening bolts) [166, 174].

- Explorations on the nonlinear features in frequency domain.

The time domain features only are considered in this study for fault diagnosis because of computational inexpensive. In the further studies, nonlinear features in the frequency domain [206, 207] can also be adopted and cooperate with time domain features to capture the fault information in terms of frequency bands.

- Health monitoring of faulty sensors.

The sensors in this thesis are mounted on the studied structures evenly according to the factors like the region of coverage (ROC) associated with the sensors and the region of influence (ROI) associated with the events. The accuracy of fault localization might be decreased with the occurrence of faulty sensors. Thus, the strategies considering the stability of sensors should be developed in further studies.

- Automatic fault diagnosis system without human intervention.

Though the machine learning based optimization method embedded is capable of selecting sensor nodes with limited human intervention, in the further, an advanced control interface without human involved is to be designed to realize the operational convenience.

References

- [1] S. M. Li, M. Dan, and L. Yang, "An Airplane Health Management approach for civil aviation," in *2011 Prognostics and System Health Management Conference, PHM-Shenzhen 2011*, 2011.
- [2] C. Lo, J. P. Lynch, and M. Liu, "Distributed model-based nonlinear sensor fault diagnosis in wireless sensor networks," *Mechanical Systems and Signal Processing*, 2014.
- [3] J. Liu, W. Wang, F. Ma, Y. B. Yang, and C. S. Yang, "A data-model-fusion prognostic framework for dynamic system state forecasting," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 4, pp. 814-823, 2012.
- [4] P. K. Vemulapalli, V. Monga, and S. N. Brennan, "Robust extrema features for time-series data analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1464-1479, 2013.
- [5] R. T. Olszewski, "Generalized feature extraction for structural pattern recognition in time series data," Department of Philosophy, Carnegie Mellon University, 2001.
- [6] M. S. Nixon, X. U. Liu, C. Direkoğlu, and D. J. Hurley, "On using physical analogies for feature and shape extraction in computer vision," *Computer Journal*, vol. 54, no. 1, pp. 11-25, 2011.
- [7] G. J. Scott, M. N. Klaric, C. H. Davis, and C. R. Shyu, "Entropy-balanced bitmap tree for shape-based object retrieval from large-scale satellite imagery databases," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 5, pp. 1603-1616, 2011.
- [8] X. Tan, and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635-1650, 2010.
- [9] A. Samal, S. Bhatia, P. Vadlamani, and D. Marx, "Searching satellite imagery with integrated measures," *Pattern Recognition*, vol. 42, no. 11, pp. 2502-2513, 2009.
- [10] R. Das Gupta, J. K. Dash, and M. Sudipta, "Rotation invariant textural feature extraction for image retrieval using eigen value analysis of intensity gradients and multi-resolution analysis," *Pattern Recognition*, vol. 46, no. 12, pp. 3256-3267, 2013.
- [11] S. Y. Lee, K. D. Nguyen, T. C. Huynh, J. T. Kim, J. H. Yi, and S. H. Han, "Vibration-based damage monitoring of harbor caisson structure with damaged foundation-structure interface," *Smart Structures and Systems*, vol. 10, no. 6, pp. 517-546, 2012.
- [12] J. Dai, M. Qi, J. Wang, J. Dai, and J. Kong, "Robust and accurate moving shadow detection based on multiple features fusion," *Optics and Laser*

- Technology*, vol. 54, pp. 232-241, 2013.
- [13] O. Háva, M. Skrbek, and P. Kordík, “Supervised two-step feature extraction for structured representation of text data,” *Simulation Modelling Practice and Theory*, vol. 33, pp. 132-143, 2013.
- [14] Z. Liu, J. Qu, M. J. Zuo, and H. B. Xu, “Fault level diagnosis for planetary gearboxes using hybrid kernel feature selection and kernel Fisher discriminant analysis,” *International Journal of Advanced Manufacturing Technology*, vol. 67, no. 5-8, pp. 1217-1230, 2013.
- [15] M. L. Hoseini, Z. L.; Tao, T.; Zuo, M. J., *Original feature used for condition monitoring of planetary gearboxes*, Department of Mechanical Engineering, University of Alberta, Canada, 5, May, 2011.
- [16] K. Javed, R. Gouriveau, N. Zerhouni, and P. Nectoux, “Enabling health monitoring approach based on vibration data for accurate prognostics,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 647-656, 2015.
- [17] W. Dargie, and M. K. Denko, “Analysis of error-agnostic time- and frequency-domain features extracted from measurements of 3-D accelerometer sensors,” *IEEE Systems Journal*, vol. 4, no. 1, pp. 26-33, 2010.
- [18] D. Tkach, H. Huang, and T. A. Kuiken, “Study of stability of time-domain features for electromyographic pattern recognition,” *J Neuroeng Rehabil*, vol. 7, pp. 21, 2010.
- [19] L. L. Jiang, H. K. Yin, X. J. Li, and S. W. Tang, “Fault diagnosis of rotating machinery based on multisensor information fusion using SVM and time-domain features,” *Shock and Vibration*, vol. 2014, 2014.
- [20] C. C. Wang, Y. Kang, and C. C. Liao, “Gear fault diagnosis in time domains via bayesian networks,” *Transactions of the Canadian Society for Mechanical Engineering*, vol. 37, no. 3, pp. 665-672, 2013.
- [21] B. Samanta, and K. R. Al-Balushi, “Artificial neural network based fault diagnostics of rolling element bearings using time-domain features,” *Mechanical Systems and Signal Processing*, vol. 17, no. 2, pp. 317-328, 2003.
- [22] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, “Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks,” *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3398-3407, 2013.
- [23] A. Moosavian, H. Ahmadi, B. Sakhaei, and R. Labbafi, “Support vector machine and K-nearest neighbour for unbalanced fault detection,” *Journal of Quality in Maintenance Engineering*, vol. 20, no. 1, pp. 65-75, 2014.
- [24] L. Wei, Z. Kun, and C. Xi, “A new fault diagnosis method for cold test based on frequency analysis,” *International Journal of Modelling, Identification and Control*, vol. 18, no. 2, pp. 146-157, 2013.
- [25] A. Bouzida, O. Touhami, R. Ibtouen, A. Belouchrani, M. Fadel, and A. Rezzoug, “Fault diagnosis in industrial induction machines through discrete wavelet transform,” *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4385-4395, 2011.

- [26] R. Liao, K. Wang, L. Yang, J. Li, S. Nie, and L. Yuan, "Binary particle swarm optimization selected time-frequency features for partial discharge signal classification," *International Review of Electrical Engineering*, vol. 7, no. 5, pp. 5905-5917, 2012.
- [27] X. Z. Gao, H. Xu, X. Wang, and K. Zenger, "A study of negative selection algorithm-based motor fault detection and diagnosis," *International Journal of Innovative Computing, Information and Control*, vol. 9, no. 2, pp. 875-901, 2013.
- [28] J. Flett, and G. M. Bone, "Fault detection and diagnosis of diesel engine valve trains," *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 316-327, 2016.
- [29] A. Beghi, L. Cecchinato, C. Corazzol, M. Rampazzo, F. Simmini, and G. A. Susto, "A one-class SVM based tool for machine learning novelty detection in HVAC chiller systems." *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 1953-1958, 2014.
- [30] J. Toivola, M. A. Prada, and J. Hollmen, "Novelty Detection in Projected Spaces for Structural Health Monitoring," *Advances in Intelligent Data Analysis IX, Proceedings*, vol. 6065, pp. 208-215, 2010.
- [31] A. Diez, N. L. D. Khoa, M. Makki Alamdari, Y. Wang, F. Chen, and P. Runcie, "A clustering approach for structural health monitoring on bridges," *Journal of Civil Structural Health Monitoring*, vol. 6, no. 3, pp. 429-445, 2016.
- [32] F. Di Maio, E. Zio, M. Pecht, P. Tse, and K. Tsui, "Ensemble of unsupervised fuzzy C-means classifiers for clustering health status of oil sand pumps." *Esrel*, pp. 419-427, 2011.
- [33] X. Y. Wang, and V. Makis, "Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov-Smirnov test," *Journal of Sound and Vibration*, vol. 327, no. 3-5, pp. 413-423, Nov 13, 2009.
- [34] C. Kar, and A. R. Mohanty, "Multistage gearbox condition monitoring using motor current signature analysis and Kolmogorov-Smirnov test," *Journal of Sound and Vibration*, vol. 290, no. 1-2, pp. 337-368, 2006.
- [35] F. Cong, J. Chen, and Y. Pan, "Kolmogorov-Smirnov test for rolling bearing performance degradation assessment and prognosis," *JVC/Journal of Vibration and Control*, vol. 17, no. 9, pp. 1337-1347, 2011.
- [36] C. G. Gonzalez, S. Da Silva, M. J. Brennan, and V. Lopes, "Structural damage detection in an aeronautical panel using analysis of variance," *Mechanical Systems and Signal Processing*, vol. 52-53, no. 1, pp. 206-216, 2015.
- [37] M. C. Carnero, R. González-Palma, D. Almorza, P. Mayorga, and C. López-Escobar, "Statistical quality control through overall vibration analysis," *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1138-1160, 2010.
- [38] Y. Shao, and C. K. Mechefske, "Gearbox vibration monitoring using extended Kalman filters and hypothesis tests," *Journal of Sound and*

- Vibration*, vol. 325, no. 3, pp. 629-648, 2009.
- [39] Y. M. Zhan, and C. K. Mechefske, "Robust detection of gearbox deterioration using compromised autoregressive modeling and Kolmogorov-Smirnov test statistic. Part II: Experiment and application," *Mechanical Systems and Signal Processing*, vol. 21, no. 5, pp. 1983-2011, Jul, 2007.
- [40] D. Gorinevsky, "Fault Isolation in Data-Driven Multivariate Process Monitoring," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1840-1852, Sep, 2015.
- [41] Y. Ling, and S. Mahadevan, "Integration of structural health monitoring and fatigue damage prognosis," *Mechanical Systems and Signal Processing*, vol. 28, pp. 89-104, Apr, 2012.
- [42] A. Willersrud, M. Blanke, L. Imsland, and A. Pavlov, "Fault diagnosis of downhole drilling incidents using adaptive observers and statistical change detection," *Journal of Process Control*, vol. 30, pp. 90-103, Jun, 2015.
- [43] A. Willersrud, M. Blanke, L. Imsland, and A. Pavlov, "Drillstring Washout Diagnosis Using Friction Estimation and Statistical Change Detection," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1886-1900, Sep, 2015.
- [44] J. Kullaa, "Detection, identification, and quantification of sensor fault in a sensor network," *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 208-221, Oct, 2013.
- [45] R. Gonzalez, B. Huang, and E. Lau, "Process monitoring using kernel density estimation and Bayesian networking with an industrial case study," *ISA Transactions*, vol. 58, pp. 330-347, 2015.
- [46] T. Wang, J. Wei, W. Zhang, H. Zhong, and T. Huang, "Workload-aware anomaly detection for web applications," *Journal of Systems and Software*, vol. 89, no. 1, pp. 19-32, 2014.
- [47] A. Picot, Z. Obeid, J. Régner, S. Poignant, O. Darnis, and P. Maussion, "Statistic-based spectral indicator for bearing fault detection in permanent-magnet synchronous machines using the stator current," *Mechanical Systems and Signal Processing*, vol. 46, no. 2, pp. 424-441, 2014.
- [48] S. Chatterjee, S. Sadhu, and T. K. Ghoshal, "Improved estimation and fault detection method for a class of nonlinear hybrid systems using self switched sigma point filter," in *2014 International Conference on Control, Instrumentation, Energy and Communication, CIEC 2014*, 2014, pp. 578-582.
- [49] R. Liu, and H. Zuo, "Research on electrostatic monitoring method of rolling bearings with injected fault under variable operating conditions," *Yi Qi Yi Biao Xue Bao/Chinese Journal of Scientific Instrument*, vol. 35, no. 10, pp. 2348-2355, 2014.
- [50] Y. Wang, and D. Infield, "Supervisory control and data acquisition data-based non-linear state estimation technique for wind turbine gearbox condition monitoring," *IET Renewable Power Generation*, vol. 7, no. 4, pp.

350-358, 2013.

- [51] A. Kumar, A. Srivastava, N. Goel, and A. Nayak, "Model-free approach and methodology for data anomaly detection for real time diagnostic solution." *Electrical & Computer Engineering*, pp. 1-4, 2010.
- [52] V. Vakharia, V. K. Gupta, and P. K. Kankar, "A comparison of feature ranking techniques for fault diagnosis of ball bearing," *Soft Computing*, vol. 20, no. 4, pp. 1601-1619, 2016.
- [53] P. Abreu, J. Moreira, I. Costa, D. Castelão, L. Reis, and J. Garganta, "Human versus virtual robotics soccer: A technical analysis," *European Journal of Sport Science*, vol. 12, no. 1, pp. 26-35, 2012.
- [54] M. Jiang, M. A. Munawar, T. Reidemeister, and P. A. S. Ward, "Efficient fault detection and diagnosis in complex software systems with information-theoretic monitoring," *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 4, pp. 510-522, 2011.
- [55] M. Jiang, M. A. Munawar, T. Reidemeister, and P. A. S. Ward, "System monitoring with metric-correlation models," *IEEE Transactions on Network and Service Management*, vol. 8, no. 4, pp. 348-360, 2011.
- [56] M. Jiang, M. A. Munawar, T. Reidemeister, and P. A. S. Ward, "Automatic fault detection and diagnosis in complex software systems by information-theoretic monitoring." *IEEE/IFIP International Conference on Dependable Systems & Networks*, pp. 285-294.
- [57] F. Calitz, "An Alternative to the Kolmogorov-Smirnov Test for Goodness of Fit," *Communications in Statistics-Theory and Methods*, vol. 16, no. 12, pp. 3519-3534, 1987.
- [58] F. Y. Cong, J. Chen, and G. M. Dong, "Performance degradation assessment by Kolmogorov-Smirnov test and prognosis based on AR model," *Zhendong yu Chongji/Journal of Vibration and Shock*, vol. 31, no. 10, pp. 79-82, 2012.
- [59] J. J. Qian, J. Yang, and G. W. Gao, "Discriminative histograms of local dominant orientation (D-HLDO) for biometric image feature extraction," *Pattern Recognition*, vol. 46, no. 10, pp. 2724-2739, Oct, 2013.
- [60] Y. Feng, J. C. Ren, J. M. Jiang, M. Halvey, and J. M. Jose, "Effective venue image retrieval using robust feature extraction and model constrained matching for mobile robot localization," *Machine Vision and Applications*, vol. 23, no. 5, pp. 1011-1027, Sep, 2012.
- [61] Y. S. Dong, and J. W. Ma, "Feature extraction through contourlet subband clustering for texture classification," *Neurocomputing*, vol. 116, pp. 157-164, Sep 20, 2013.
- [62] V. Sugumaran, and K. I. Ramachandran, "Effect of number of features on classification of roller bearing faults using SVM and PSVM," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4088-4096, 2011.
- [63] S. Narasimhan, R. Rengaswamy, and R. Vadigepalli, "Structural properties of gene regulatory networks: Definitions and connections," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 6, no. 1,

- pp. 158-170, 2009.
- [64] E. S. García-Treviño, and J. A. Barria, "Structural generative descriptions for time series classification," *IEEE Transactions on Cybernetics*, vol. 44, no. 10, pp. 1978-1991, 2014.
- [65] S. Güneş, M. Dursun, K. Polat, and Ş. Yosunkaya, "Sleep spindles recognition system based on time and frequency domain features," *Expert Systems with Applications*, vol. 38, no. 3, pp. 2455-2461, 2011.
- [66] Y. G. Lei, M. J. Zuo, Z. J. He, and Y. Y. Zi, "A multidimensional hybrid intelligent method for gear fault diagnosis," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1419-1430, Mar, 2010.
- [67] B. Boashash, N. A. Khan, and T. Ben-Jabeur, "Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review," *Digital Signal Processing: A Review Journal*, vol. 40, no. 1, pp. 1-30, 2015.
- [68] M. Pirra, "Advanced techniques for aircraft bearing diagnostics," Mechanical Engineering, Politecnico di Torino, 2013.
- [69] Y. Zhou, S. Wang, and P. Zhao, "Fan fault analysis based on time domain features and improved k-means clustering algorithm," *Zhendong Ceshi Yu Zhenduan/Journal of Vibration, Measurement and Diagnosis*, vol. 32, no. 3, pp. 437-440, 2012.
- [70] R. Puche-Panadero, M. Pineda-Sanchez, M. Riera-Guasp, J. Roger-Folch, E. Hurtado-Perez, and J. Perez-Cruz, "Improved resolution of the MCSA method via Hilbert transform, enabling the diagnosis of rotor asymmetries at very low slip," *IEEE Transactions on Energy Conversion*, vol. 24, no. 1, pp. 52-59, 2009.
- [71] A. Bellini, A. Yazidi, F. Filippetti, C. Rossi, and G. A. Capolino, "High frequency resolution techniques for rotor fault detection of induction machines," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4200-4209, 2008.
- [72] J. Zhou, Y. Qin, L. Kou, M. Yuwono, and S. Su, "Fault detection of rolling bearing based on FFT and," *Journal of Advanced Mechanical Design, Systems and Manufacturing*, vol. 9, no. 5, 2015.
- [73] W. Hu, and J. Bao, "The outlier interval detection algorithms on astronomical time series data," *Mathematical Problems in Engineering*, vol. 2013, 2013.
- [74] R. Wald, T. Khoshgoftaar, and J. C. Sloan, "Fourier transforms for vibration analysis: A review and case study." in *IEEE International Conference on Information Reuse & Integration*, pp. 366-371, 2011.
- [75] I. Howard, "A review of rolling element bearing vibration 'detection, diagnosis and prognosis'," *DSTO Aeronautical and Maritime Research Laboratory, GPO Box 4331 Melbourne Victoria 3001 Australia*, 1995.
- [76] N. S. Bt Ahmad, A. R. Abdullah, N. Bt Bahari, and M. A. A. Hassan, "Switches faults analysis of voltage source inverter (VSI) using short time fourier transform (STFT)," *International Review on Modelling and Simulations*, vol. 7, no. 3, pp. 409-415, 2014.

- [77] S. R. Samantaray, R. K. Dubey, and B. C. Babu, "A novel time-frequency transform based spectral energy function for fault detection during power swing," *Electric Power Components and Systems*, vol. 40, no. 8, pp. 881-897, 2012.
- [78] J. Cusidó, L. Romeral, J. A. Ortega, J. A. Rosero, and A. G. Espinosa, "Fault detection in induction machines using power spectral density in wavelet decomposition," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 2, pp. 633-643, 2008.
- [79] N. Baydar, and A. Ball, "A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution," *Mechanical Systems and Signal Processing*, vol. 15, no. 6, pp. 1091-1107, 2001.
- [80] V. V. Polyshchuk, F. K. Choy, and M. J. Braun, "Gear fault detection with time-frequency based parameter NP4," *International Journal of Rotating Machinery*, vol. 8, no. 1, pp. 57-70, 2002.
- [81] X. Lv, M. Xing, S. Zhang, and Z. Bao, "Keystone transformation of the Wigner-Ville distribution for analysis of multicomponent LFM signals," *Signal Processing*, vol. 89, no. 5, pp. 791-806, 2009.
- [82] Y. Wang, G. Xu, L. Liang, and K. Jiang, "Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 54, pp. 259-276, 2015.
- [83] Z. Feng, X. Chen, M. Liang, and F. Ma, "Time-frequency demodulation analysis based on iterative generalized demodulation for fault diagnosis of planetary gearbox under nonstationary conditions," *Mechanical Systems and Signal Processing*, vol. 62, pp. 54-74, 2015.
- [84] W. He, Y. Zi, B. Chen, F. Wu, and Z. He, "Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform," *Mechanical Systems and Signal Processing*, vol. 54, pp. 457-480, 2015.
- [85] Q. He, "Time-frequency manifold for nonlinear feature extraction in machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 200-218, 2013.
- [86] D. P. Jena, S. N. Panigrahi, and R. Kumar, "Multiple-teeth defect localization in geared systems using filtered acoustic spectrogram," *Applied Acoustics*, vol. 74, no. 6, pp. 823-833, 2013.
- [87] N. D. Kim, and S. Udpa, "Texture classification using rotated wavelet filters," *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans.*, vol. 30, no. 6, pp. 847-852, 2000.
- [88] T. M. Lai, L. A. Snider, E. Lo, and D. Sutanto, "High-impedance fault detection using discrete wavelet transform and frequency range and RMS conversion," *IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. 397-407, 2005.

- [89] T. Sun, T. S. Huang, L. M. Kan, M. Li, and J. D. Xiang, "Fault detection and isolation capability improvement of wavelet neural network basing on wavelet packet transform de-noising," *Harbin Gongye Daxue Xuebao/Journal of Harbin Institute of Technology*, vol. 37, no. 4, pp. 561-564, 2005.
- [90] B. P. Tang, W. Y. Liu, and T. Song, "Wind turbine fault diagnosis based on Morlet wavelet transformation and Wigner-Ville distribution," *Renewable Energy*, vol. 35, no. 12, pp. 2862-2866, Dec, 2010.
- [91] N. Li, J. Yang, R. Zhou, and C. Liang, "Determination of knock characteristics in spark ignition engines: An approach based on ensemble empirical mode decomposition," *Measurement Science and Technology*, vol. 27, no. 4, 2016.
- [92] Y. Chen, and M. Q. Feng, "A technique to improve the empirical mode decomposition in the Hilbert-Huang transform," *Earthquake Engineering and Engineering Vibration*, vol. 2, no. 1, pp. 75-85, 2003.
- [93] L. Xu, "Study on fault detection of rolling element bearing based on translation-invariant denoising and hilbert-huang transform," *Journal of Computers*, vol. 7, no. 5, pp. 1142-1146, 2012.
- [94] Z. K. Peng, P. W. Tse, and F. L. Chu, "A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing," *Mechanical Systems and Signal Processing*, vol. 19, no. 5, pp. 974-988, 2005.
- [95] G. Karakaya, S. Galelli, S. D. Ahipasaoglu, and R. Taormina, "Identifying (Quasi) Equally Informative Subsets in Feature Selection Problems for Classification: A Max-Relevance Min-Redundancy Approach," *IEEE Transactions on Cybernetics*, 2015.
- [96] M. Han, and X. X. Liu, "Feature selection techniques with class separability for multivariate time series," *Neurocomputing*, vol. 110, pp. 29-34, Jun 13, 2013.
- [97] C. Hou, F. Nie, X. Li, D. Yi, and Y. Wu, "Joint embedding learning and sparse regression: A framework for unsupervised feature selection," *IEEE Transactions on Cybernetics*, vol. 44, no. 6, pp. 793-804, 2014.
- [98] J. H. Jia, N. Yang, C. Zhang, A. Z. Yue, J. Y. Yang, and D. H. Zhu, "Object-oriented feature selection of high spatial resolution images using an improved Relief algorithm," *Mathematical and Computer Modelling*, vol. 58, no. 3-4, pp. 619-626, Aug, 2013.
- [99] G. Wang, J. Ma, and S. L. Yang, "Igf-Bagging: Information Gain Based Feature Selection for Bagging," *International Journal of Innovative Computing Information and Control*, vol. 7, no. 11, pp. 6247-6259, Nov, 2011.
- [100] O. Ozturk, A. Aksac, A. Elsheikh, T. Ozyer, and R. Alhajj, "A Consistency-Based Feature Selection Method Allied with Linear SVMs for HIV-1 Protease Cleavage Site Prediction," *Plos One*, vol. 8, no. 8, Aug 23, 2013.

- [101] H. C. Peng, F. H. Long, and C. Ding, "Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, Aug, 2005.
- [102] Z. Zhao, and H. Liu, "Searching for interacting features in subset selection," *Intelligent Data Analysis*, vol. 13, no. 2, pp. 207-228, 2009.
- [103] Y. Liu, F. Tang, and Z. Zeng, "Feature selection based on dependency margin," *IEEE Transactions on Cybernetics*, vol. 45, no. 6, pp. 1209-1221, 2015.
- [104] J. Wang, L. Wu, J. Kong, Y. Li, and B. Zhang, "Maximum weight and minimum redundancy: A novel framework for feature subset selection," *Pattern Recognition*, vol. 46, no. 6, pp. 1616-1627, 2013.
- [105] M. B. Imani, M. R. Keyvanpour, and R. Azmi, "A Novel Embedded Feature Selection Method: A Comparative Study in the Application of Text Categorization," *Applied Artificial Intelligence*, vol. 27, no. 5, pp. 408-427, May 28, 2013.
- [106] C. H. Yang, L. Y. Chuang, and C. H. Yang, "IG-GA: A Hybrid Filter/Wrapper Method for Feature Selection of Microarray Data," *Journal of Medical and Biological Engineering*, vol. 30, no. 1, pp. 23-28, 2010.
- [107] A. S. Ghareb, A. A. Bakar, and A. R. Hamdan, "Hybrid feature selection based on enhanced genetic algorithm for text categorization," *Expert Systems with Applications*, vol. 49, pp. 31-47, 2016.
- [108] J. Apolloni, G. Leguizamón, and E. Alba, "Two hybrid wrapper-filter feature selection algorithms applied to high-dimensional microarray experiments," *Applied Soft Computing Journal*, vol. 38, pp. 922-932, 2016.
- [109] M. M. Kabir, M. Shahjahan, and K. Murase, "A new hybrid ant colony optimization algorithm for feature selection," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3747-3763, 2012.
- [110] C. P. Lee, and Y. Leu, "A novel hybrid feature selection method for microarray data analysis," *Applied Soft Computing Journal*, vol. 11, no. 1, pp. 208-213, 2011.
- [111] H. H. Hsu, C. W. Hsieh, and M. D. Lu, "Hybrid feature selection by combining filters and wrappers," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8144-8150, 2011.
- [112] Z. Zhu, Y. S. Ong, and M. Dash, "Wrapper-filter feature selection algorithm using a memetic framework," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 1, pp. 70-76, 2007.
- [113] B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimization for feature selection in classification: A multi-objective approach," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1656-1671, 2013.
- [114] K. Nag, and N. R. Pal, "A Multiobjective Genetic Programming-Based Ensemble for Simultaneous Feature Selection and Classification," *IEEE Transactions on Cybernetics*, 2015.

- [115] D. Chakraborty, and N. R. Pal, "Selecting useful groups of features in a connectionist framework," *IEEE Transactions on Neural Networks*, vol. 19, no. 3, pp. 381-396, Mar, 2008.
- [116] K. C. Lin, and Y. H. Hsieh, "Classification of Medical Datasets Using SVMs with Hybrid Evolutionary Algorithms Based on Endocrine-Based Particle Swarm Optimization and Artificial Bee Colony Algorithms," *Journal of Medical Systems*, vol. 39, no. 10, 2015.
- [117] S. Khan, M. Nazir, N. Riaz, and M. Khan, "Optimized features selection using hybrid PSO-GA for multi-view gender classification," *International Arab Journal of Information Technology*, vol. 12, no. 2, pp. 183-189, 2015.
- [118] S. Gunasundari, and S. Janakiraman, "A hybrid PSO-SFS-SBS algorithm in feature selection for liver cancer data," *Lecture Notes in Electrical Engineering*, Springer Verlag, 2015, pp. 1369-1376.
- [119] V. Bolon-Canedo, N. Sanchez-Marono, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," *Knowledge and Information Systems*, vol. 34, no. 3, pp. 483-519, Mar, 2013.
- [120] M. X. Liu, and D. Q. Zhang, "Pairwise Constraint-Guided Sparse Learning for Feature Selection," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 298-310, Jan, 2016.
- [121] P. Buteneers, K. Caluwaerts, J. Dambre, D. Verstraeten, and B. Schrauwen, "Optimized Parameter Search for Large Datasets of the Regularization Parameter and Feature Selection for Ridge Regression," *Neural Processing Letters*, vol. 38, no. 3, pp. 403-416, Dec, 2013.
- [122] A. C. Haury, P. Gestraud, and J. P. Vert, "The Influence of Feature Selection Methods on Accuracy, Stability and Interpretability of Molecular Signatures," *Plos One*, vol. 6, no. 12, Dec 21, 2011.
- [123] C. De Mol, S. Mosci, M. Traskine, and A. Verri, "A Regularized Method for Selecting Nested Groups of Relevant Genes from Microarray Data," *Journal of Computational Biology*, vol. 16, no. 5, pp. 677-690, May, 2009.
- [124] V. Bolón-Canedo, N. Sánchez-Marono, A. Alonso-Betanzos, J. M. Benítez, and F. Herrera, "A review of microarray datasets and applied feature selection methods," *Information Sciences*, vol. 282, pp. 111-135, 10/20/, 2014.
- [125] T. T. Nguyen, A. W. C. Liew, M. T. Tran, X. C. Pham, and M. P. Nguyen, "A novel genetic algorithm approach for simultaneous feature and classifier selection in multi classifier system," in *2014 IEEE Congress on Evolutionary Computation, CEC 2014*, 2014, pp. 1698-1705.
- [126] Z. Zhu, Y.-S. Ong, and M. Dash, "Markov blanket-embedded genetic algorithm for gene selection," *Pattern Recognition*, vol. 40, no. 11, pp. 3236-3248, 11//, 2007.
- [127] B. Xue, W. Fu, and M. Zhang, "Differential evolution (DE) for multi-objective feature selection in classification." in *16th Genetic and Evolutionary Computation Conference, GECCO 2014*, pp. 83-84.

- [128] D. D. M. Mason, "Encyclopedia of Data Warehousing and Mining, 2nd edition," *Online Information Review*, vol. 33, no. 2, pp. 389-390, 2009.
- [129] L.-Y. Chuang, H.-W. Chang, C.-J. Tu, and C.-H. Yang, "Improved binary PSO for feature selection using gene expression data," *Computational Biology and Chemistry*, vol. 32, no. 1, pp. 29-38, 2008.
- [130] H. B. Nguyen, B. Xue, I. Liu, and M. Zhang, "Filter based backward elimination in wrapper based PSO for feature selection in classification." in *2014 IEEE Congress on Evolutionary Computation, CEC 2014*, pp. 3111-3118.
- [131] G. Wang, H. E. Chu, Y. Zhang, H. Chen, W. Hu, Y. Li, and X. Peng, "Multiple parameter control for ant colony optimization applied to feature selection problem," *Neural Computing and Applications*, 2015.
- [132] X. T. Li, and M. H. Yin, "Multiobjective Binary Biogeography Based Optimization for Feature Selection Using Gene Expression Data," *Ieee Transactions on Nanobioscience*, vol. 12, no. 4, pp. 343-353, Dec, 2013.
- [133] G. Dudek, "An Artificial Immune System for Classification with Local Feature Selection," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 6, pp. 847-860, Dec, 2012.
- [134] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, and C. A. C. Coello, "A survey of multiobjective evolutionary algorithms for data mining: Part i," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 1, pp. 4-19, 2014.
- [135] K. J. Kim, and S. B. Cho, "An evolutionary algorithm approach to optimal ensemble classifiers for DNA microarray data analysis," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 3, pp. 377-388, Jun, 2008.
- [136] J. Lu, T. Zhao, and Y. Zhang, "Feature selection based-on genetic algorithm for image annotation," *Knowledge-Based Systems*, vol. 21, no. 8, pp. 887-891, 2008.
- [137] R. Li, J. Lu, Y. Zhang, and T. Zhao, "Dynamic Adaboost learning with feature selection based on parallel genetic algorithm for image annotation," *Knowledge-Based Systems*, vol. 23, no. 3, pp. 195-201, 2010.
- [138] D. Y. Harvey, and M. D. Todd, "Automated Feature Design for Numeric Sequence Classification by Genetic Programming," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 4, pp. 474-489, 2015.
- [139] M. Sheikhan, and N. Mohammadi, "Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection," *Neural Computing & Applications*, vol. 21, no. 8, pp. 1961-1970, Nov, 2012.
- [140] M. Sheikhan, and N. Mohammadi, "Time series prediction using PSO-optimized neural network and hybrid feature selection algorithm for IEEE load data," *Neural Computing & Applications*, vol. 23, no. 3-4, pp. 1185-1194, Sep, 2013.
- [141] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems Magazine*, vol. 22, no. 3,

- pp. 52-67, Jun, 2002.
- [142] S. D. Muller, J. Marchetto, S. Airaghi, and P. Koumoutsakos, "Optimization based on bacterial chemotaxis," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 16-29, Feb, 2002.
- [143] J. H. Cho, and D. H. Kim, "Intelligent feature selection by bacterial foraging algorithm and information theory," in *3rd International Conference on Advanced Communication and Networking, ACN 2011*, Brno, 2011, pp. 238-244.
- [144] D. Y. Liang, and W. K. Zheng, "An intelligent feature selection method based on the Bacterial Foraging Algorithm," *Intelligent Structure and Vibration Control, Pts 1 and 2*, vol. 50-51, pp. 304-308, 2011.
- [145] P. Kora, and S. R. Kalva, "Hybrid Bacterial Foraging and Particle Swarm Optimization for detecting Bundle Branch Block," *SpringerPlus*, vol. 4, no. 1, 2015.
- [146] Z. W. Gao, C. Cecati, and S. X. Ding, "A Survey of Fault Diagnosis and Fault-Tolerant Techniques-Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757-3767, Jun, 2015.
- [147] R. Doraiswami, C. P. Diduch, and J. Tang, "A New Diagnostic Model for Identifying Parametric Faults," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 533-544, May, 2010.
- [148] Z. Yu, F. Wang, and L. Liu, "An adaptive fault-tolerant robust controller with fault diagnosis for submarine's vertical movement," *International Journal of Control, Automation and Systems*, vol. 13, no. 6, pp. 1337-1345, 2015.
- [149] A. Akhenak, E. Duviella, L. Bako, and S. Lecoeuche, "Online fault diagnosis using recursive subspace identification: Application to a dam-gallery open channel system," *Control Engineering Practice*, vol. 21, no. 6, pp. 797-806, Jun, 2013.
- [150] M. Döhler, and L. Mevel, "Subspace-based fault detection robust to changes in the noise covariances," *Automatica*, vol. 49, no. 9, pp. 2734-2743, 2013.
- [151] M. S. Mahmoud, and A. U. D. Qureshi, "Model identification and analysis of small-power wind turbines," *International Journal of Modelling, Identification and Control*, vol. 17, no. 1, pp. 19-31, 2012.
- [152] Z. Ge, M. Zhang, and Z. Song, "Nonlinear process monitoring based on linear subspace and Bayesian inference," *Journal of Process Control*, vol. 20, no. 5, pp. 676-688, 2010.
- [153] D. Henry, A. Zolghadri, M. Monsion, and F. Cazaurang, "Fault diagnosis in induction machines using the generalized structured singular value," *Control Engineering Practice*, vol. 10, no. 6, pp. 587-598, 2002.
- [154] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME--Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.
- [155] Z. Wang, and H. Shang, "Kalman filter based fault detection for two-

- dimensional systems,” *Journal of Process Control*, vol. 28, pp. 83-94, 2015.
- [156] G. Heredia, and A. Ollero, “Detection of sensor faults in small helicopter UAVs using Observer/Kalman filter identification,” *Mathematical Problems in Engineering*, vol. 2011, 2011.
- [157] S. Jayaram, “A new fast converging Kalman filter for sensor fault detection and isolation,” *Sensor Review*, vol. 30, no. 3, pp. 219-224, 2010.
- [158] A. Mirzaee, and K. Salahshoor, “Fault diagnosis and accommodation of nonlinear systems based on multiple-model adaptive unscented Kalman filter and switched MPC and H-infinity loop-shaping controller,” *Journal of Process Control*, vol. 22, no. 3, pp. 626-634, Mar, 2012.
- [159] H. X. Le, and S. Matunaga, “A residual based adaptive unscented Kalman filter for fault recovery in attitude determination system of microsatellites,” *Acta Astronautica*, vol. 105, no. 1, pp. 30-39, Dec, 2014.
- [160] I. Maalej, D. Ben Halima Abid, and C. Rekik, “Fuzzy Kalman observer for fault detection in nonlinear discrete systems based on type-2 fuzzy logic,” *International Journal of Automation and Control*, vol. 9, no. 4, pp. 261-282, 2015.
- [161] N. Chatti, B. Ould-Bouamama, A. L. Gehin, and R. Merzouki, “Signed Bond Graph for multiple faults diagnosis,” *Engineering Applications of Artificial Intelligence*, vol. 36, pp. 134-147, 2014.
- [162] M. Yu, and D. W. Wang, “Model-Based Health Monitoring for a Vehicle Steering System With Multiple Faults of Unknown Types,” *IEEE Transactions on Industrial Electronics*, vol. 61, no. 7, pp. 3574-3586, Jul, 2014.
- [163] G. G. Yen, and L. W. Ho, “Online multiple-model-based fault diagnosis and accommodation,” *IEEE Transactions on Industrial Electronics*, vol. 50, no. 2, pp. 296-312, Apr, 2003.
- [164] W. Chen, and M. Saif, “Fuzzy nonlinear unknown input observer design with fault diagnosis applications,” *JVC/Journal of Vibration and Control*, vol. 16, no. 3, pp. 377-401, 2010.
- [165] W. Chen, and M. Saif, “Design of a TS based fuzzy nonlinear unknown input observer with fault diagnosis applications,” in *Proceedings of the American Control Conference*, 2007, pp. 2545-2550.
- [166] Z. Gao, C. Cecati, and S. X. Ding, “A survey of fault diagnosis and fault-tolerant techniques-part II: Fault diagnosis with knowledge-based and hybrid/active approaches,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3768-3774, 2015.
- [167] A. Garcia-Perez, R. D. J. Romero-Troncoso, E. Cabal-Yepez, and R. A. Osornio-Rios, “The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors,” *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 2002-2010, 2011.
- [168] M. Grbovic, W. C. Li, N. A. Subrahmanya, A. K. Usadi, and S. Vucetic, “Cold Start Approach for Data-Driven Fault Detection,” *IEEE Transactions*

- on *Industrial Informatics*, vol. 9, no. 4, pp. 2264-2273, Nov, 2013.
- [169] J. Strackeljau, S. Goreczka, and D. Behr, "A fault detection concept for single class problems," in *9th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2012, CM 2012 and MFPT 2012*, London, 2012, pp. 675-682.
- [170] M. P. Michaelides, and C. G. Panayiotou, "SNAP: Fault Tolerant Event Location Estimation in Sensor Networks Using Binary Data," *IEEE Transactions on Computers*, vol. 58, no. 9, pp. 1185-1197, Sep, 2009.
- [171] M. P. Michaelides, C. Laoudias, and C. G. Panayiotou, "Fault Tolerant Detection and Tracking of Multiple Sources in WSNs using Binary Data," *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 Held Jointly with the 2009 28th Chinese Control Conference (Cdc/Ccc 2009)*, pp. 3769-3774, 2009.
- [172] M. P. Michaelides, and C. G. Panayiotou, "Fault Tolerant Maximum Likelihood Event Localization in Sensor Networks Using Binary Data," *IEEE Signal Processing Letters*, vol. 16, no. 5, pp. 406-409, May, 2009.
- [173] X. Xu, X. Gao, J. Wan, and N. Xiong, "Trust index based fault tolerant multiple event localization algorithm for WSNs," *Sensors*, vol. 11, no. 7, pp. 6555-6574, 2011.
- [174] K. Sudha, and V. Thulasi Bai, "A hybrid fault liberal control model for nonlinear system processes," *International Journal of Applied Engineering Research*, vol. 9, no. 24, pp. 29443-29465, 2014.
- [175] H. Gao, C. Xu, and C. Li, "Quantum-inspired cultural bacterial foraging algorithm for direction finding of impulse noise," *International Journal of Innovative Computing and Applications*, vol. 6, no. 1, pp. 44-54, 2014.
- [176] V. Franklin, and K. Paramasivam, "Bacterial foraging optimization supported utility based call admission control framework for 3GPP LTE networks," *Life Science Journal*, vol. 10, no. 2, pp. 394-400, 2013.
- [177] K. M. Bakwad, S. S. Pattnaik, B. S. Sohi, S. Devi, B. K. Panigrahi, and S. V. R. S. Gollapudi, "Bacterial foraging optimization technique cascaded with adaptive filter to enhance peak signal to noise ratio from single image," *IETE Journal of Research*, vol. 55, no. 4, pp. 173-179, 2009.
- [178] B. Biswal, H. S. Behera, R. Bisoi, and P. K. Dash, "Classification of power quality data using decision tree and chemotactic differential evolution based fuzzy clustering," *Swarm and Evolutionary Computation*, vol. 4, pp. 12-24, Jun, 2012.
- [179] R. Panda, and M. K. Naik, "Fusion of infrared and visual images using bacterial foraging strategy," *WSEAS Transactions on Signal Processing*, vol. 8, no. 4, pp. 145-156, 2012.
- [180] B. Niu, and H. Wang, "Bacterial colony optimization," *Discrete Dynamics in Nature and Society*, vol. 2012, 2012.
- [181] R. N. Khushaba, A. Al-Ani, and A. Al-Jumaily, "Feature subset selection using differential evolution and a statistical repair mechanism," *Expert*

- Systems with Applications*, vol. 38, no. 9, pp. 11515-11526, Sep, 2011.
- [182] B. Niu, Y. Fan, H. Wang, L. Li, and X. Wang, "Novel bacterial foraging optimization with time-varying chemotaxis step," *International Journal of Artificial Intelligence*, vol. 7, no. 11 A, pp. 257-273, 2011.
- [183] H. Wang, X. J. Jing, and B. Niu, "A weighted bacterial colony optimization for feature selection," *Intelligent Computing in Bioinformatics*, vol. 8590, pp. 379-389, 2014.
- [184] Z. Liu, J. Qu, M. J. Zuo, and H.-b. Xu, "Fault level diagnosis for planetary gearboxes using hybrid kernel feature selection and kernel Fisher discriminant analysis," *The International Journal of Advanced Manufacturing Technology*, vol. 67, no. 5-8, pp. 1217-1230, 2012.
- [185] J. Kennedy, and R. Eberhart, "Particle swarm optimization," in *IEEE International Conference on Neural Networks - Conference Proceedings*, 1995, pp. 1942-1948.
- [186] "The Fundamentals-AI, Data Clustering," <http://www.bindichen.co.uk/post/AI/fuzzy-c-means.html>.
- [187] H. Wang, and X. J. Jing, "Vibration Signal Based Fault Diagnosis in Complex Structures: A Beam-Like-Structure Approach," *Structural Health Monitoring*, 2017.
- [188] K. Dragan, M. Dziendzikowski, S. Klimaszewski, S. Klysz, and A. Kurnyta, "Energy correlated damage indices in fatigue crack extent quantification," *10th International Conference on Damage Assessment of Structures, DAMAS 2013*, 2013, pp. 1186-1193.
- [189] M. P. Michaelides, C. Laoudias, and C. G. Panayiotou, "Fault Tolerant Localization and Tracking of Multiple Sources in WSNs Using Binary Data," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1213-1227, Jun, 2014.
- [190] D. C. Kammer, "Sensor Placement for on-Orbit Modal Identification and Correlation of Large Space Structures," *Journal of Guidance Control and Dynamics*, vol. 14, no. 2, pp. 251-259, Mar-Apr, 1991.
- [191] S. Jin, K. G. Yu, X. M. Lai, and Y. H. Liu, "Sensor placement strategy for fixture variation diagnosis of compliant sheet metal assembly process," *Assembly Automation*, vol. 29, no. 4, pp. 358-363, 2009.
- [192] S. Jin, Y. H. Liu, and Z. Q. Lin, "A Bayesian network approach for fixture fault diagnosis in launch of the assembly process," *International Journal of Production Research*, vol. 50, no. 23, pp. 6655-6666, 2012.
- [193] "Traveling Salesman Problem – Genetic Algorithm," <http://blogs.mathworks.com/pick/2011/10/14/traveling-salesman-problem-genetic-algorithm/>.
- [194] S. Minakuchi, I. Yamauchi, N. Takeda, and Y. Hirose, "Detecting an arrested crack in a foam-core sandwich structure using an optical fiber sensor embedded in a crack arrester," *Advanced Composite Materials*, vol. 20, no. 5, pp. 419-433, 2011.

- [195] M. Gan, C. Wang, and C. Zhu, "Multiple-domain manifold for feature extraction in machinery fault diagnosis," *Measurement: Journal of the International Measurement Confederation*, vol. 75, pp. 76-91, 2015.
- [196] V. T. Tran, and B. S. Yang, "An intelligent condition-based maintenance platform for rotating machinery," *Expert Systems with Applications*, vol. 39, no. 3, pp. 2977-2988, 2012.
- [197] X. Qi, Z. Yuan, and X. Han, "Diagnosis of misalignment faults by tachless order tracking analysis and RBF networks," *Neurocomputing*, vol. 169, pp. 439-448, 2015.
- [198] R. K. Singleton, E. G. Strangas, and S. Aviyente, "Extended kalman filtering for remaining-useful-life estimation of bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1781-1790, 2015.
- [199] S. Bahrapour, B. Moshiri, and K. Salahshoor, "Weighted and constrained possibilistic C-means clustering for online fault detection and isolation," *Applied Intelligence*, vol. 35, no. 2, pp. 269-284, Oct, 2011.
- [200] J. P. Anita, and P. T. Vanathi, "Multiple fault diagnosis and test power reduction using genetic algorithms," *International Conference on Eco-Friendly Computing and Communication Systems, ICECCS 2012*, 2012, pp. 84-92.
- [201] R. Doraiswami, and L. Cheded, "Fault Diagnosis of a Sensor Network: A Distributed Filtering Approach," *Journal of Dynamic Systems Measurement and Control-Transactions of the Asme*, vol. 135, no. 5, Sep, 2013.
- [202] S. Kolluri, and M. Bhushan, "Audit of sensor networks for efficient fault diagnosis," *Journal of Process Control*, vol. 23, no. 6, pp. 881-893, Jul, 2013.
- [203] J. F. C. Markmiller, and F. K. Chang, "Sensor Network Optimization for a Passive Sensing Impact Detection Technique," *Structural Health Monitoring-an International Journal*, vol. 9, no. 1, pp. 25-39, Jan, 2010.
- [204] O. A. Vanli, C. Zhang, A. Nguyen, and B. Wang, "A minimax sensor placement approach for damage detection in composite structures," *Journal of Intelligent Material Systems and Structures*, vol. 23, no. 8, pp. 919-932, May, 2012.
- [205] C. J. Li, and H. Lee, "Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics," *Mechanical Systems and Signal Processing*, vol. 19, no. 4, pp. 836-846, Jul, 2005.
- [206] X. Jing, and Q. Li, "A nonlinear decomposition and regulation method for nonlinearity characterization," *Nonlinear Dynamics*, vol. 83, no. 3, pp. 1355-1377, 2016.
- [207] X. Jing, and Z. Lang, "Frequency Domain Analysis and Design of Nonlinear Systems based on Volterra Series Expansion: A Parametric Characteristic Approach," Monograph by Springer, 2015.