Coverage in WSNs for Structural Health Monitoring

X. Liu[#1] , J.N. Cao[#2], W.Z. Song[*3], Steven Lai[#4], J.J.Li[#5], Y.Liu[#6]

[#] *Dept. of Computing, the Hong Kong Polytechnic University*
Hong Kong, China
(csxfliu[1],csjcao[2],csylai[4],csjli[5],csyangliu[6])@comp.polyu.edu.hk

[*] *Dept. of Computer Science, the Georgia State University*
USA
songwz[3]@cs.gsu.edu

Abstract—**WSNs have been regarded as a new sensing paradigm for structural health monitoring (SHM) due to the low cost, easy deployment and high scalability. To extend the system lifetime, sensor nodes deployed on a structure can be partitioned into multiple groups and groups work in a round-robin fashion. To ensure that each group can still monitor the whole structure effectively, coverage in WSNs for SHM needs to be studied. Although coverage problem has been studied extensively in WSNs and many energy-efficient coverage-preserving protocols have been proposed for various monitoring applications, they would fail in SHM because monitoring structural condition uses a totally different scheme from other applications. The sensing region of a sensor in SHM is no longer a circle, a sphere, or even a convex area. In addition, the sensing region of a group of sensors cannot be determined by combining the sensing regions of individual sensor together, which is a very important assumption used in all of the traditional coverage models.**

In this paper, we first study the coverage problem in SHM and define a new coverage model: SHM-coverage, which is directly connected with damage detection capability of sensor nodes. The difference between the SHM-coverage and the existing coverage is clearly described. We then propose two methods to partition the deployed sensor nodes into disjoint subsets where each subset is able to 'SHM-cover' the whole structure. The effectiveness of the proposed approaches is demonstrated through results from both simulation and experiment.

I. Introduction

Wireless sensor networks (WSNs) present a new prototype of computing systems consisting of small and cheap nodes which scattered in the environment in order to monitor the spatial-temporal phenomena that people are interested. Applications of wireless sensor networks include battlefield surveillance, environmental and habitat monitoring, and biological detection [1].

Considering the limited energy supply of wireless sensor nodes, duty-cycling is widely used in WSNs. Deployed sensor nodes are partitioned into multiple groups and groups work in a round-robin order to monitor the phenomena that application users are interested. To ensure that each group is still able to effectively monitor the phenomena occurs in a given area, network coverage, which measures how well an area is monitored by a sensor network, has been extensively studied. Many energy efficient coverage-preserving protocols have been proposed for WSNs [2-9]. These protocols have been applied to many applications including the intruder detection in battlefield surveillance, habitat monitoring, or fire detection et al. However, to the authors' best knowledge, none of them can be applied to a special monitoring application, structural health monitoring (SHM).

The objective of SHM is to monitor the integrity of structures and pinpoint the existence and location of possible damage. WSNs have been regarded as a new sensing paradigm for SHM due to the low cost, easy deployment and high scalability. In a typical WSN-based SHM system, wireless sensor nodes equipped with accelerometers or strain gauges are deployed on the structure to be monitored to measure its responses under ambient or external input forces. The vibration data collected from deployed sensor nodes are then transmitted to a central server, where the structure's modal parameters, which characterize the dynamic behaviour of the structure, are obtained (we delay a more detailed description of modal parameters, including natural frequencies and mode shape matrix, in the Appendix). Change on these modal parameters is the indication of structural damage[10].

From the discussion above, it can be seen that SHM uses a totally different monitoring scheme from other applications. In traditional monitoring applications, the energy emitted by an event of interest is received by the deployed sensor nodes, and then they can directly give 0/1 about the occurrence of event from their own point of view. Generally, an event can be detected by sensor nodes neighboring to the location of the event. However, in SHM, detection of an event (i.e. damage) is not so straightforward. Damage is detected by examining changes in the modal parameters. These parameters are identified using measurement data from multiple sensor nodes and are the global features of the structure. Sensor nodes which are able to detect damage are not restricted to those which are near the damage location. The sensors should be chosen as those whose measurement data, when used together, can identify modal parameters with or above a pre-defined accuracy.

Coverage in SHM then becomes different from other monitoring applications. This difference can be more clearly observed from the perspective of sensing region. In most of the traditional coverage problem[3-9] the sensing region of a sensor is modelled as a circle (in 2D space) or a sphere (in a 3D space) centred at the sensor with radius as its sensing range. Although this assumption is relaxed in some applications, they still require the sensing region to be at least a convex function [2]. Another important assumption used in all of the traditional coverage models is that the sensing region of a sensor set S is the union of the sensing regions of individual sensors in S:

$$A(\{s_1, s_2, \dots s_N\}) = A(s_1) \cup A(s_2) \dots \cup A(s_N) \tag{1}$$

where $A(s_i)(i = 1, \dots N)$ and $A(\{s_1, s_2, \dots s_N\})$ are the sensing regions of sensor s_i and the sensor set $\{s_1, s_2, \dots s_N\}$, respectively. In other words, the full coverage is achieved by accumulating the coverage area of individual sensor.

However, in SHM, the sensing region of a sensor is no longer a circle, a sphere, or even a convex area. If we were to define a sensing region for a sensor or a sensor set in SHM as was in traditional monitoring applications, it should be either the whole structure (if accuracy of modal parameters identified from the sensor/sensor set satisfies the requirement) or 0 (if it fails). More importantly, the aforementioned union operation is also not valid in SHM applications. For example, assume accurate enough modal parameters of a structure can be identified only when the data of a sensor set S are used together. This implies that the sensing region of each sensor node in S is '0', but that of the set S is the whole structure. Therefore, we do not know whether a structure is 'covered' by a sensor set S even the sensing region of each sensor node in S is given. Different models of the sensing region causes previous coverage-preserving protocols no longer applicable in SHM.

In this paper, we are the first to study the coverage problem in SHM and define a new coverage model: SHM-coverage, which is directly connected with damage detection capability of sensor nodes. We first give a criterion to determine whether a given set of sensor nodes can 'SHM-cover' a structure. Based on the criterion, we proposed two approaches, one heuristic and the other based on genetic algorithm (GA), to divide sensor nodes deployed on a structure into disjoint sets. Sensor nodes in each subset are connected and each subset is able to 'SHM-cover' the whole structure. Sensor nodes in each set can work in a round-robin order and each set can monitor the whole structure effectively. The lifetime of the system is therefore significantly extended. Through simulation and experiment, the effectiveness of the proposed approaches is demonstrated.

The main contributions of this paper are the following:

1. A new coverage model SHM-coverage is firstly defined which is more suitable for structural health monitoring studies. The difference between the SHM-coverage and the traditional coverage is clearly described.
2. We propose a heuristic method (in both centralized and distributed version) and a GA method to partition the network into subsets where each subset can SHM-cover the structure. Use the obtained subset, system lifetime can be extended and *this extension is not at the expense of damage detection capability*, and
3. The performance of our approach is demonstrated through both simulation and experiment data

The structure of the paper is organized as follows. In section II, we introduce the related works, which mainly focus on the traditional coverage protocols in WSN. In section III, the preliminaries, the definition of SHM-coverage and the problem formulation are provided. In section IV, we present a heuristic approach and a GA approach for energy efficient coverage-preserving scheduling. The simulation results on a suspension bridge and the experimental results on a 12-floor structure are proposed in section V. Section VI concludes the paper.

II. RELATED WORKS

In this section, the previous work of energy-efficient coverage protocols is briefly reviewed. The purpose of the review is to demonstrate that these protocols are not applicable for SHM.

Energy efficient coverage-preserving protocols can be mainly divided as centralized and distributed methods. In [2] and [3], centralized protocols are proposed and energy efficient coverage is transformed to the set cover problem: the algorithm allocates sensor nodes into maximum number of mutually exclusive sets of sensor nodes, where each cover completely covers the area. After dividing the sensor nodes into disjoint cover set, a schedule can be worked out by activating these subsets successively to extend network lifetime.

When designing distributed protocols, the basic idea is mainly as follows: by exchanging information with the active neighbors, a sensor node knows whether or not its sensing region has already been covered by its active neighbors and will then be activated or go to sleep accordingly. In [4], a distributed coverage configuration protocol (CCP) is proposed which can configure a sensor network to any coverage degree. Using CCP, a scheduling mechanism can activate only a small number of sensor nodes to perform coverage to extend the system lifetime. Tian et al. [5] devised a distributed algorithm that ensures complete coverage using the concept of 'sponsored area'. Whenever a sensor node receives a packet from one of its working neighbors, it calculates its sponsored area. If the union of all the sponsored areas of a sensor node covers the coverage disk of the node, the node turns itself off. In [6], an Optimal Geographical Density Control (OGDC) algorithm is proposed. The OGDC algorithm can configure a sensor network with the characteristics of full-coverage, network connectivity, and maximum energy conservation. The energy is conserved by controlling the density of the active nodes.

Besides coverage, a WSN must also provide satisfactory connectivity so that sensed data can be delivered to aggregation points or a sink node. The relationship between coverage and connectivity has been provided in [4] and [6].

Given a convex region A, a set of sensors with the uniform sensing range R_c and the communication range R_s, complete coverage of A implies connectivity if and only if $R_c \geq 2R_s$. Some protocols such as [4] use this assumption without paying extra efforts on connectivity.

However, the protocols mentioned above would fail in SHM because monitoring a structure uses a totally different scheme from other monitoring applications. Whether a sensor or a set of sensor can cover the whole structure is determined by accuracy of the identified modal parameters. The sensing region of a sensor or a sensor set in SHM is therefore no longer a circle or even a convex area. In addition, the sensing region of a sensor set is not the accumulated sensing region of individual sensor, which is a very important assumption in all the coverage models. In addition, since we cannot define sensing region for individual sensor node as before, $R_c \geq 2R_s$ is no longer valid and connectivity must always be considered whenever designing protocols for WSN in SHM.

In this paper, we study the coverage problem used for SHM and define a special coverage: SHM-coverage. SHM-coverage of a WSN is directly associated with its damage detection capability. Consequently, we proposed two approaches to divide the sensor nodes into disjoint sets while each set can effectively monitor the condition of the whole structure. These subsets can then work in a round-robin order and the system lifetime can be significantly increased.

III. PROBLEM FORMULATION

We now formally define the SHM-coverage problem addressed in this paper. We start by giving the definition of SHM-coverage.

SHM-coverage (Definition 1): A structure is said to be SHM-covered by a sensor set S *iff* using S, the modal parameters of the structure can be identified with no less than a pre-defined accuracy.

From the definition, it can be seen that SHM-coverage is tightly connected with the ability to detect structure damage since damage is detected by examining the changes of identified modal parameters. If a sensor set S can SHM-cover a structure, we can say that a certain level of damage, wherever occurs on the structure, can be detected using data from S.

We then need to find a criterion to evaluate the accuracy of identified modal parameters. In this paper, the modal parameters are identified using a classical modal identification method, the Eigen-Realization Algorithm (ERA) [11]. The ERA builds a Hankel matrix from collected vibration data and then implements singular value decomposition (SVD) to extract modal parameters. The accuracy of identified modal parameters using a sensor set S can be evaluated by the condition number (which is the ratio of the largest to the smallest singular value) of the mode shape matrix corresponding to S (denoted as Φ_S). The larger the condition number of Φ_S, the less accurate of identified modal parameters will be. Φ_S can be obtained using the finite element model (FEM) of the structure.

In this paper, an upper bound for the condition number is defined. If the condition number of a sensor set S is below this value, the modal parameters identified from S will be accurate enough to detect damage. In another word, sensor set S will *SHM-cover* the structure. Note that this upper bound is structure-dependent, and should be determined in practice by the measurement noise and the damage level required to be detected. An example of the upper bound will be given in Section V.B.

The previous description is summarized as the second definition of SHM-coverage:

SHM-coverage (Definition 2): For a given structure to be monitored and a user-defined upper bound value α, a sensor set S deployed on the structure is said to be able to SHM-cover the structure *iff* the condition number of Φ_S is equal or smaller than α.

Some observations of the SHM-coverage are listed as follows:

1. For a given sensor set S, we can determined whether S is able to SHM-cover the whole structure using the condition number of its mode shape matrix Φ_S

2. For given sensor set S, the condition number of Φ_S is determined by the two factors: 1) the structure's FEM 2) the number and the locations of sensor nodes in S. Generally speaking, the more number of sensor nodes and the more independent of the rows of Φ_S, the smaller the condition number will be.

3. The condition number corresponding to a sensor set S cannot be determined by accumulating the condition number of each sensor node in S. Therefore, whether a sensor set S can SHM-cover a structure cannot be determined by considering sensors in S individually, the relationship of sensors in S must also be considered. In other words, the sensing region of S cannot be determined by combining that of each sensor node in S together.

Now, we consider designing the energy efficient protocols for WSN-based SHM. The basic idea is however very simple: sensor nodes deployed on a structure is divided into as disjoint subsets and each subset is able to SHM-cover the whole structure. We also require that sensor nodes in each subset are connected. After this division is finished, the obtained subsets can be scheduled to be active successively. To maximize the lifetime of the entire network, the problem then becomes how to maximize the number of disjoint subsets that can SHM-cover the structure.

The following notations are used to formulate the problem.

V: The set of all the sensors

m: All the sensors can be divided into m subsets and each subset is able to SHM-cover the structure.

$G(V_i, E_i)$: The graph for i^{th} subset. The set V_i consists of all the nodes in the subset and an edge exists from node u to v if they can communicate with each other

CN_i: The condition number of the V_i.

α: The user-specified upper bound of condition number. Then the problem is formulated as:

Objective: Max m

Subject to:

1. $\cup_{1\le i\le m} V_i \subseteq V$
2. Sensor nodes in V_i is connected according to $G(V_i, E_i)$
3. $V_i \cap V_j = \emptyset, 1 \le i, j \le m, i \ne j$
4. $CN_i \le \alpha, 1 \le i \le m$

IV. Proposed Approaches

In this section, we will describe two approaches, one heuristic and the other based on the genetic algorithm.

A. The Heuristic Method for Energy Efficient Scheduling in WSN-based SHM

In this section, we present a heuristic method to solve the problem defined in the previous section. The main idea of the heuristic method is to iteratively construct subsets V_i by choosing sensor which is connected to the current sensor set and its participation can greatly improve the damage detection ability of the current sensor subset. In another word, from all the nodes connected to the current V_i, the one is chosen whose participation can minimize the decrease of the condition number of the subset. In this way, we can potentially increase the number of generated subsets and ensure sensors in each subset are connected.

How to choose the first sensor node in each V_i is very important. It should be noted that since each subset is constructed in a greedy manner, it is possible that there remain some sensors in S which are not be able to SHM-cover the structure because of the following reasons: (1) they are not connected, or (2) although they are connected, the condition number is larger than α (3) or both. A large number of remaining sensor nodes inevitable decreases the number of subsets than can be potentially constructed. Therefore, the basic idea of choosing the first sensor node in each subset is to avoid the remaining of un-connected sensor nodes after the whole process is finished. A simple example is illustrated in Fig. 1 which shows the topology of nine sensor nodes. Assume we start from node 5 and the obtained subset is $V_1 = \{5,2,4,6,8\}$. This is obviously not a good choice because no more subsets can be further constructed since the remaining sensor nodes 1,3,7,9 are not connected. This can be partially attributed to the fact that we constructing from a sensor node with the maximum degree (the degree of node 5 is 8). When node 5 is removed after V_1 is constructed, the number of links of the remaining sensor nodes deceases significantly.

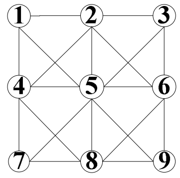

Figure 1 Example of how to choose the first sensor node

To address this problem, sensor node with the minimum degree is selected when start constructing each subset. In case more than one sensor nodes have the same degree, the importance of these sensor nodes will be ranked using the effective independence method (EFI) [12] and the one with the highest value is selected. This criterion tries to avoid the remaining of important sensor nodes. It should be noted that the EFI method must be carried out each time when a new V_i is to be constructed. This is because the order of the importance may be changed as certain sensors are deleted.

A more formal and complete description of the designed algorithm is shown in Algorithm 1. The first node in V_i is chosen using function $getMinDegree(S)$ to find out the node with minimum degree and $getMaxEFI(\{x\})$ is used in case of a tie. The condition number of the V_i is then calculated using $getCN(V_i)$. If the condition number is still larger than the pre-define upper bound α, more sensors need to be added into V_i. Any sensors in S which are directly connected with V_i are selected as candidate sensors using $getNeighbors(V_{i,}S)$. From the obtained sensors, one whose participation can minimize the condition number of V_i is chosen and added into V_i (using function $getMinCN(\{x\}, V_i)$). In a greedy manner, sensors are added to V_i. When V_i has no neighbors, sensor nodes in V_i are removed from S. If at some iteration, V_i can provide SHM-coverage (i.e. $getCN(V_i) \le \alpha$), the process for constructing V_i is finished and a new subset V_{i+1} will be constructed in the same manner. The heuristic method stops when no more subset can be constructed which is able to SHM-cover monitored structure.

Since each subset is constructed in a greedy manner, it is possible that there remain some sensors in S that do not belong to any constructed subsets. These sensor nodes are denoted as set Re in Algorithm 1. To make the best use of these sensor nodes, they are put back to existing subsets to improve their damage detection capability. From all of the subsets, the one which has the maximum condition number has the priority to first choose the remaining nodes (this subset is chosen using function $getMaxCN(V)$). Assume this subset is denoted as V_{Max}, any nodes in Re which are directly connected with V_{Max} are candidates and from which, one whose participation can minimize the decrease of the condition number of V_{Max} is chosen and added into V_{Max}. The condition numbers of all the subsets are then re-evaluated and the process is re-iterated until no remaining sensor nodes can be utilized. Using this procedure, a balanced improvement among all subsets can be obtained using the remaining sensor nodes.

We use a simple example to illustrate the proposed algorithm in a step-by-step manner. A total of 16 sensor nodes, labelled from A to P, are deployed on a plate structure with fixed bottom edge (see Fig. 2(a)). Fig. 2 (b) represents the connectivity topology of these sensor nodes. Assume the modal parameters of the first three modes are needed to detect possible damage. Fig. 3 illustrates the theoretical first three mode shapes of the plate. These mode shapes are to be used when calculating the condition number of different subsets. We also assume that condition number of each subset should be smaller than 2 so that the identified modal parameters can be accurate enough to detect a certain level of damage.

Algorithm 1. The Heuristic Method
Input: Sensor set S, threshold α
Output: Set of subsets V

```
i=1, V = ∅, Re = ∅
while S is not empty do
  {x} = getMinDegree(S)
  if |{x}| > 1, V_i = getMaxEFI({x}),
  else V_i = x
  end if
  while getCN(V_i) > α, do
     {x} = getNeighbors(V_i, S)
     if {x} == ∅
       Re = Re ∪ V_i, remove V_i from S
       if S == ∅, break; end if
     else
       x = getMinCN({x}, V_i)
       V_i = x ∪ V_i, Remove x from S
     end if
  end while
  V = {V, V_i}, i=i+1;
end while
W = V, V = ∅
while Re ≠ ∅ ,do
  V_Max = getMaxCN(W);
  {x} = getNeighbors(V_Max, Re)
  if {x} == ∅
     V = {V, V_Max}
     remove V_Max from W,
     if W == ∅, break; end if
  else
     x = getMinCN({x}, V_Max);
     V_Max = x ∪ V_Max,  V = {V, V_Max}
     Remove x from Re
  end if
end while
Output V
```

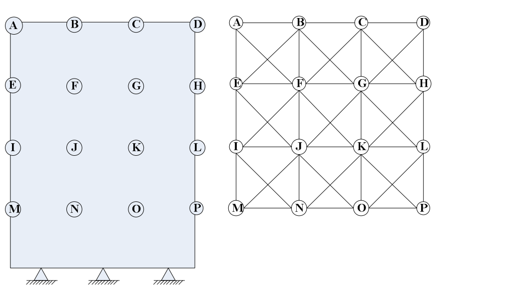

Figure 2 Plate structure and sensor nodes (a) plate structure and location of deployed sensor nodes (b) topology of sensor nodes

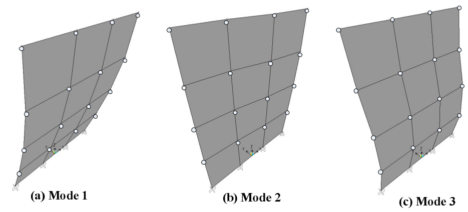

Figure 3 The first three theoretical mode shapes of the plate (a) mode 1 (b) mode 2 (c) mode 3

The heuristic method starts by finding out the node with the minimum degree. From Fig. 4(a), it can be seen that nodes A,D,M and P have the same minimum degree 3. Among these four nodes, D is the most important one according to the EFI criterion and therefore the subset construction starts from D (The selected nodes are marked with black). Node D has three neighbors, namely, C,G and H (neighbours to the current subset are marked with gray). From these three sensor nodes, node G is selected since when combined with D, subset {D,G} has the minimum condition number than other possible subsets. By calculating the condition number of current subset $V_1 = \{D, G\}$, more sensor nodes need to be added. The neighbors of V_1, which are B,C,F,H,J,K,L, become candidates to be included into V_1 in the next step (see Fig. 4(c)). From these candidates, node B is chosen in the similar way as was in the previous step (see Fig. 4(c)). This process continues until the condition number of V_1 is smaller than the threshold 2. Fig. 4(e) illustrates the constructed subset V_1={B,D,G,J}. Nodes in V_1 are then removed from the available sensor pool S.

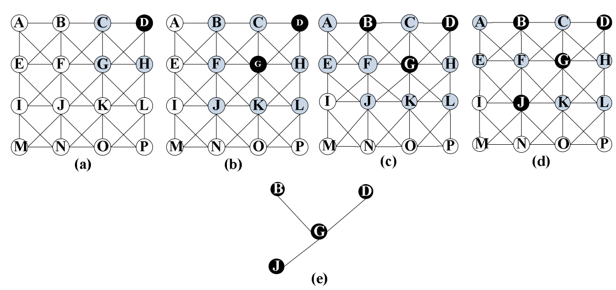

Figure 4: The Process of Constructing the First Subset V_1 (a) Stage 1 (b) Stage 2 (c) Stage 3 (d) Stage 4 (e) Nodes contained in V1

Likewise, the construction of the second subset V_2 and the third subset V_3 is illustrated in Fig. 5 and Fig. 6, respectively. Fig. 7 summarizes the three subsets obtained and the corresponding condition numbers.

There remain three unselected nodes, E,M and P, after V_3 is constructed. No subsets can be further constructed using E,M,P and they are put back to the existing subsets. From Fig. 7 (d), V_3 has the largest condition number and therefore has the priority to first select the remaining sensor nodes. E is chosen and added into V_3. The three subsets are then re-evaluated and the procedure re-iterates. Fig. 8 shows the final three subsets and the corresponding condition numbers. Compared with Fig. 7, the condition numbers of all the subsets are further decreased.

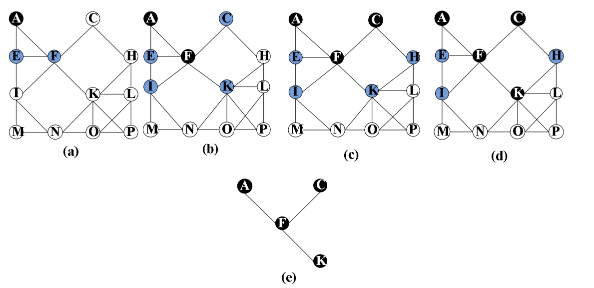

Figure 5 The process of constructing subset V_2 (a) Stage 1 (b) Stage 2 (c) Stage 3 (d) Stage 4 (e) nodes contained in V_2

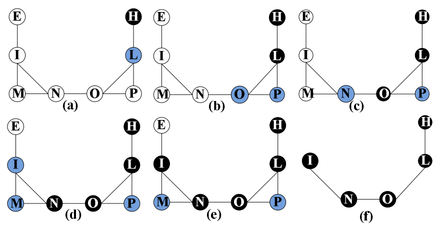

Figure 6 The process of constructing subset V_3 (a) Stage 1 (b) Stage 2 (c) Stage 3 (d) Stage 4 (e) Stage 5 (f) nodes contained in V_3

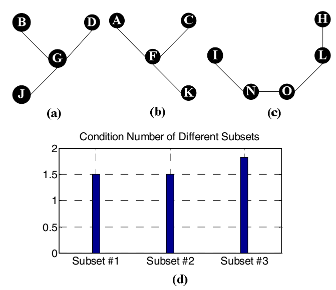

Figure 7 Summary of the three subsets constructed and the corresponding condition numbers

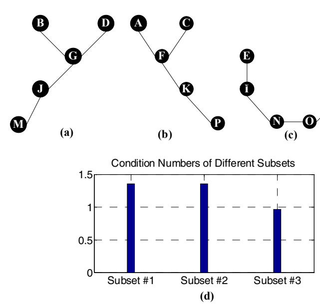

Figure 8 Summary of the three subsets constructed and the corresponding condition numbers after E,M,P are added into subsets

Based on the centralized method, we introduce a distributed algorithm for the problem where the nodes use only local information when constructing the subsets. Different from the centralized method, in the distributed method, each node will execute the algorithm and then send messages to other nodes so that they can also start to execute the algorithm.

Since the distributed method will execute in a node-by-node fashion, the key difference between the centralized and the distributed approach lies in three functions: $getMinDegree()$, $getNeighbors()$, and $getMaxEFI()$. In the distributed method, all the three functions will return the values which correspond to the current executing node's neighborhood instead of the whole network.

Before the algorithm, all nodes will first send HELLO messages to obtain their neighbors' information. Then the execution starts from a seed node which is the sink. Based on the nodes' neighborhood information, nodes that have the smallest number of neighbors will have a higher priority to be added to the subsets first. The priority is achieved by assigning different tokens to the different nodes.

Initially, the sink node will automatically be assigned to token 0. After the execution of the algorithm, the sink node will assign tokens to its neighbors according to their degree and contribution to decrease the condition number. Similar to the centralized algorithm, the nodes with lower degree and higher contribution to decrease the condition number will be assigned tokens with smaller ID. The nodes won't start the algorithm until it is its own turn. Based on the timeslots, the distributed algorithm will have similar effect as the centralized algorithm.

B. The GA Method for Energy Efficient Scheduling in WSN-based SHM

Genetic algorithms (GA) are optimization algorithms which evolve solutions in a manner analogous to the Darwinian principle of natural selection [13]. In this paper, we use GA to maximize the number of subsets while guarantee that sensor nodes in each subset are connected and can SHM-cover the structure. The first hurdle of using GA is working out how to best encode the possible solutions as genes. In the current problem, we encode the sensor nodes as like '0120310231'. In the encoding, each position corresponds to a sensor node and its value is the subset number to which it belongs. For example, in the encoding above, the first sensor node belongs to the subset 0, and the second and the third sensors belong to subset 1 and 2, respectively. In the encoding above, a total of ten sensors are divided into 4 subsets, and sensors included into each subset are sensors {#1,#4,#7} for subset 0,{#2,#6,#10} for subset 1, {#3,#8} for subset 2 and {#5, #9} for subset 3. For example, if the total number of available sensor nodes is N, and they are divided into k subsets, then '$a_1a_2…a_N$' would be an appropriate encoding, where a_i (i=1,…,N) is an integer chosen from candidate group {0,1,2,..k-1}. This encoding can be used when the number of disjoint cover sets to be divided is known *a-priori*.

However, since we are trying to find out the maximum number of subsets that can be obtained, this number is not known. To solve this problem and still use the encoding above, the following procedures are adopted. First we give an initial value of the subset number k. k can be as small as 2 if no priori guess is used. GA is then used to find out the 'optimal division' under the current subset number k. The 'optimal division' minimizes the maximum condition number of the generated subsets. After obtaining the optimal division and corresponding subsets, the subsets are evaluated. If the subset with the maximum condition number is smaller than the upper bound α, k=k+1. GA is then carried out based on this new number and the whole process reiterates until the condition number constraint fails. It should be noted that it may take a significant time if the initial k is a small value while the number of potential subsets is a large. One approach is to use some initial guess on k. For example, we can carry out the heuristic approach proposed in the previous section first and use the obtained subset number as the initial guess of k. Another approach is to dynamically determine the increase of k in each iteration according to the distance between the maximum condition number of the obtained subsets and the condition number upper bound.

Having decided on a representation, the next step is to generate, at random, an initial population of possible solutions. The number of genes in a population depends on several factors, including the size of each individual gene, which itself depends on the size of the solution space.

Having generated a population of random genes, it is necessary to decide which of them are fittest in the sense of producing the best solutions to the problem. To do this, a fitness function is required to provide a measure of the suitability of the solution. The fitness function $\mathcal{F}$ of a set of subsets $\mathrm{V} = \{\mathrm{V}_1, \mathrm{V}_2, \ldots, \mathrm{V}_\mathrm{m}\}$ is represented as:

$$\mathcal{F}(\mathrm{V}) = \text{ Inf (if any } \mathrm{V}_\mathrm{i} \text{ is not connected)} \qquad (2)$$
$$= \max\big(getCN(\mathrm{V}_1), \ldots getCN(\mathrm{V}_\mathrm{m})\big) \text{ (else)}$$

where Inf represents 'infinite'. $getCN(\mathrm{V}_\mathrm{i})$ is the function to calculate the condition number of subset V_i.

The fitness function in Eq. 2 considers both connectivity and SHM-coverage. The fitter genes, which having the smaller fitness function value, will be used for mating to create the next generation of genes which will hopefully provide better solutions to the problem. Once sufficient genes have been selected for mating, they are paired up at random and their genes combined to produce two new genes. The most common method of combination used is called crossover. Here, one point crossover is used in which a position along the genes is chosen at random and the substrings from each gene after the chosen point are switched. A gene having the maximum fitness value among a population is called elite and carried through unchanged to the next generation.

V. Validation of the Proposed Methods

A. Simulation Results

To test the effectiveness of the proposed methods, a simulated suspension bridge is generated by SAP2000 [14] (see Fig. 9(a)). Some important dimensions of the bridge are also marked in the figure.

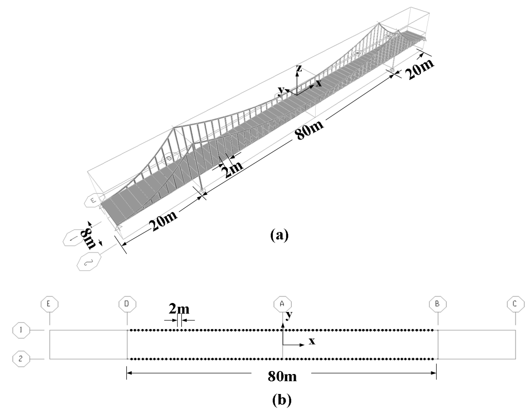

Figure 9 The simulated suspension bridge (a) The dimensions of the suspension bridge (3D), (b) the sensor locations (X-Y plane)

A total of 78 wireless sensor nodes are used to monitor the vibration at the transverse direction (z direction in Fig.9(a)) of the deck of the bridge. These sensors are evenly spaced at the outer side and inner side in the middle span of the deck with distance of 2m (see Fig. 9(b)). We also assume the communication range of each node is 10m.

The theoretical first 5 mode shapes of the structure are illustrated in Fig. 10. These theoretical mode shapes will be used to calculating the condition number of different subsets in the proposed methods.

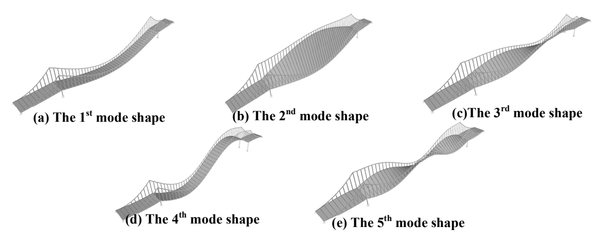

Figure 10 The first five theoretical mode shapes of the bridge

The deployed 78 sensor nodes are to be divided into disjoint subsets. We require that condition number of each subset should be smaller than 2 so that the identified modal parameters can be accurate enough to detect a certain level of damage.

Fig. 11 and Fig. 12 illustrate the results using the centralized and distributed heuristic methods, respectively. Sensor configuration in each subset and its corresponding condition number are illustrated. It can be seen that both of the two methods generate 6 subsets. Sensor nodes in each subsets are connected and each subset can provide required modal identification accuracy (condition number <2). However, the subsets generated using the centralized method is slightly better than the distributed counterpart.

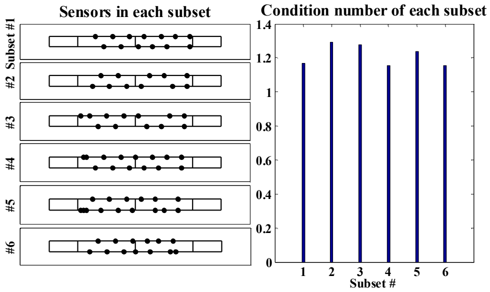

Figure 11 Results from the heuristic method (centralized) (a) Sensors in each subset (b) Condition number of each subset

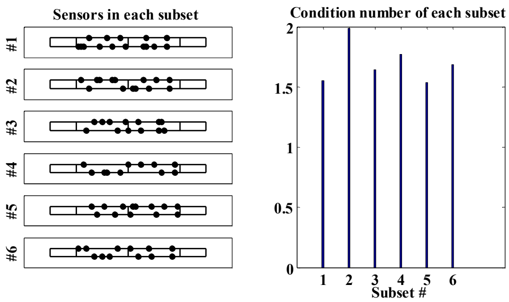

Figure 12 Results from the heuristic method (distributed)

Fig.13 shows the results of the GA method proposed in this paper. It can be seen that compared with the heuristic methods, the GA method generated more subsets. However, it should be noted that the GA method needs an initial guess of subset number and cannot directly give the maximum number of subset that can generate as was in the heuristic method. Therefore, the GA method can take longer time than the heuristic methods. In addition, when using the GA methods, we found that it is highly advisable not to generate the chromosome from random, but to use a various chromosomes that can guarantee the connectivity first. This technique can significantly decrease the computational time of the GA method.

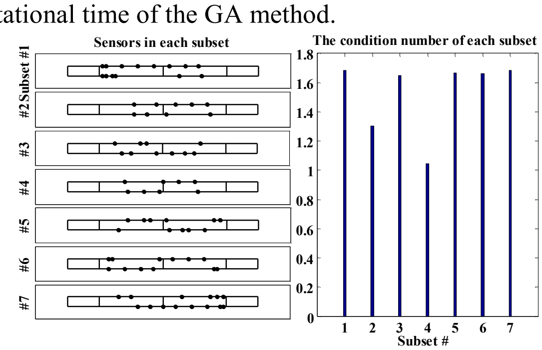

Figure 13 Results from the GA Method (a) Sensors in each subset (b) Condition number of each subset

For comparison, Fig. 14 shows the results of a clustering which quite arbitrarily divides the available 78 sensors into 6 clusters and only ensures the connectivity of each subset without considering the condition number. It can be seen that the condition numbers of these subsets can exceed far from the required threshold. In this division, the maximum condition number is higher than 1000 (in subset #1 and #6) and even the smallest condition number (subset #3) is well above the threshold 2. It will be shown later that that when these subsets are used to monitor the structure, the modal parameters identified from these subsets would not be accurate enough to detect damage.

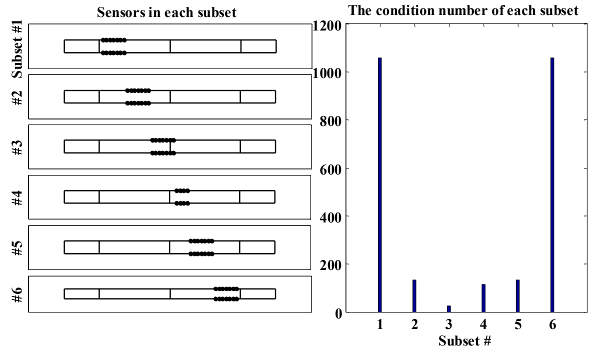

Figure 14 Results from the traditional clustering (a) sensor configuration in subsets (b) condition numbers of different subsets

To complete simulation part, we wish to further demonstrate how the condition number will affect the accuracy of identified modal parameters. Impulse responses of the simulated suspension bridge are generated at these 78 sensor nodes. The response time series were sampled at 200Hz, providing approximately 28 points per period for the highest frequency mode in the simulation. Noise was added to the sensor data at each sample as a zero-mean Gaussian sequence with variance σ^2. In this simulation, σ^2 is chosen such that the ratio of the σ to the root-mean-square sensor output averaged over all the 78 sensors is 15%. As was described, the ERA was used to extract the mode shapes and natural frequencies from simulated response. The block Hankel matrix in ERA has 50 block rows and 250 columns. In this work, the correct order of the system was always assumed to be known such that the results of each identification analysis for each subset sensor configuration could be compared directly.

The error of the identified modes (including mode shapes and natural frequencies) is calculated as follows. For i^{th} mode (i=1..5), the natural frequency error Err_f(i) is defined as:

$$\mathrm{Err\_f(i)} = \frac{\left|f_i^I - f_i^T\right|}{f_i^T} \tag{3}$$

where f_i^I is the identified i^{th} natural frequency using the ERA and the f_i^T is the true one.

The i^{th} mode shape error Err_s(i), is defined as:

$$\mathrm{Err\_s(i)} = 1 - \frac{\left|\{\boldsymbol{\Psi}_i^I\}'\{\boldsymbol{\Psi}_i^T\}\right|^2}{\left|\{\boldsymbol{\Psi}_i^I\}'\{\boldsymbol{\Psi}_i^I\}\right| \cdot \left|\{\boldsymbol{\Psi}_i^T\}'\{\boldsymbol{\Psi}_i^T\}\right|} \tag{4}$$

where $\boldsymbol{\Psi}_i^I$ and $\boldsymbol{\Psi}_i^T$ are the identified and the true i^{th} mode shape vectors, respectively.

Fig. 15 illustrates the identified mode shape error and natural frequency error using the subsets obtained from the centralized heuristic method (see Fig. 11). The mode shape error and natural frequency error of all the subsets in all of the five modes are below 4e-3 and 5e-3, respectively. It can be seen that even with the relatively high noise-to-signal ratio(15% in this case), the modal parameters of the bridge (i.e. mode shapes and natural frequencies) are very accurately identified using data from each subset.

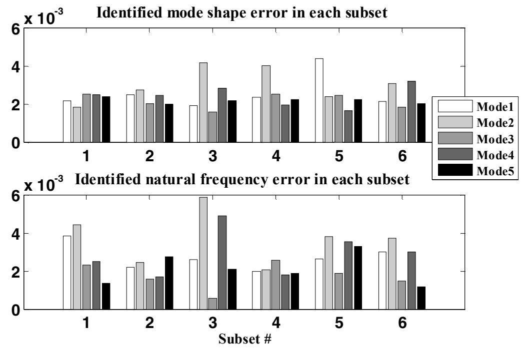

Figure 15 The identified mode shape error and natural frequency error using the subsets in Fig. 11

For comparison, the identified mode shape error and natural frequency error using the subsets illustrated in Fig. 14 are presented in Fig. 16. It can be seen that using subset with large condition number will not be able to identify these modal parameters correctly. Correspondingly, structural damage will not be able to be detected by examining the changes of identified modal parameters.

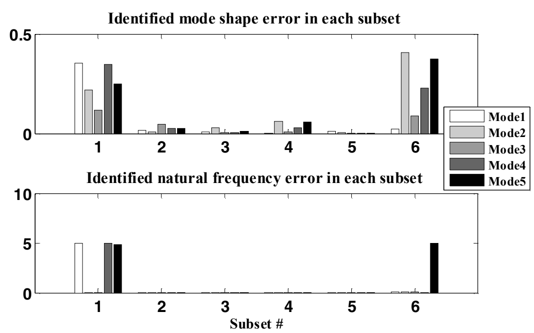

Figure 16 The identified mode shape error and natural frequency error using the subsets in Fig. 14.

B. Experiment

The effectiveness of the proposed approaches is tested through a real implementation. To address the generally high requirements of SHM application in terms of high sampling frequency, complex SHM algorithms and tight synchronized sensing, we designed a particular type of wireless sensor nodes called SHM mote. A SHM mote includes an Imote2, a sensor board, and radio-triggered wakeup & synchronization unit(see Fig.17 (a)). The testing structure has 12 floors, and the SHM motes are deployed on different floors to monitor the structure's horizontal accelerations under the excitation from a hammer (see Fig. 17(b)). Although in general condition, the deployed SHM motes can form a complete network, we decrease the transmission power and use the topology of the network shown in Fig. 17(c). We use a gateway node which is connected to a computer to collect data wirelessly. The SHM motes run modified TinyOS and are configured to sample the accelerometers at frequency of 1024Hz. At the same time, vibration data are also recorded by a wired system for the reference. Modal parameters will be identified using the data sampled from the wired-system and from which, the first three mode shapes and natural frequencies are obtained. These reference modal parameters have two purposes. The mode shapes will be used to calculate the condition number when partitioning the wireless sensor network. The three natural frequencies will be regarded as the references to test the accuracy of identified natural frequencies from the obtained subsets.

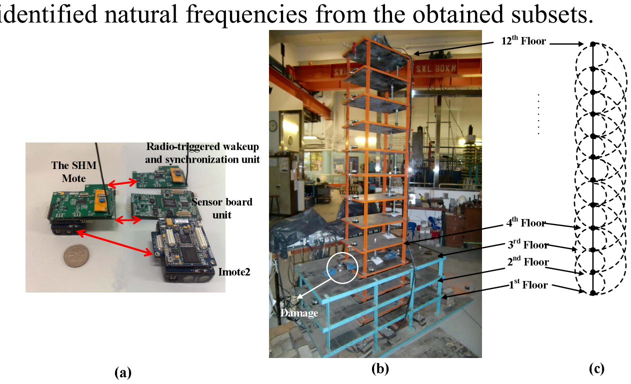

Figure 17 the SHM mote and test structure (a) the SHM mote (b) test structure (c) network topology

The deployed 12 SHM motes are to be partitioned into subsets and we require that the condition number of each subset should be smaller than 5 to be able to detect damage on this structure (the justification of choosing this value will be given later). Fig. 18 illustrates the results and it shows that all of the proposed methods partitioned the 12 nodes into 3 subsets. From the calculated condition numbers, the GA method outperforms the heuristic methods as expected. Fig. 19 illustrates the identified natural frequency errors of each subset. It shows that using data from each subset, the first three natural frequencies can be identified accurately. In other words, sensor nodes in these three subsets can work in a round-robin order to monitor the condition of the structure effectively, and the system lifetime is extended by three times.

At last, we will justify why value 5 is chosen as the threshold of condition number in this experiment. Damage in this test is generated by releasing a support ring on the third floor (see Fig. 17(b)). Table I lists the first three natural frequencies before and after damage, identified using data from the wired system. It can be seen that this damage generates more than 6 percent change on these natural frequencies. Also from Fig. 19, it can be seen that using threshold 5 makes identified natural frequencies error for all the subsets to be under 2%, the damage is therefore can be effectively detected by all of the subsets.

TABLE I
COMPARE THE NATURAL FREQUENCIES BEFORE/AFTER DAMAGE

	1st	2nd	3rd
Natural frequencies(Healthy)	64.7	161.2	224.5
Natural frequencies(Damaged)	60.2	150.3	200.1
Natural frequency error (calculated by Eq. (3))	6.96%	6.76%	10.87%

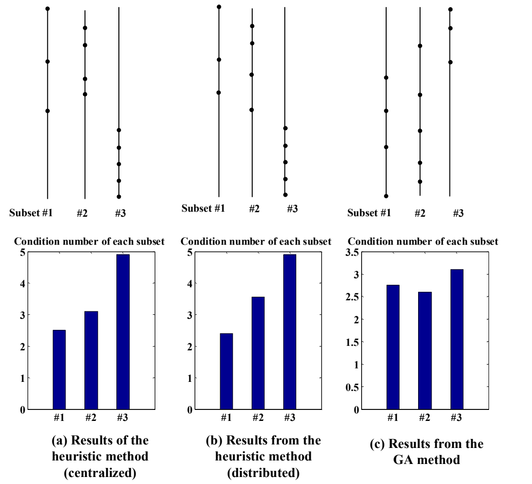

Figure 18 Results from the (a) the heuristic method(centralized) (b) the heuristic method(distributed) (c) the GA method

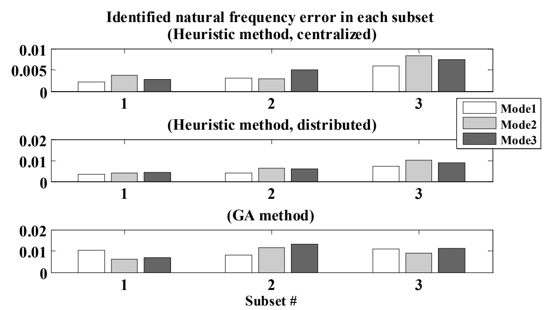

Figure 19 The identified natural frequency error using subsets from different methods

VI. Conclusion

Although coverage problem in wireless sensor networks has been studied extensively and many energy-efficient coverage-preserving protocols have been proposed for various monitoring applications, these protocols would fail in a particular monitoring area: structural health monitoring because monitoring a structure uses a totally different scheme.

To the author's best knowledge, we are the first to study the coverage problem in SHM. We defined a novel coverage: SHM-coverage, which is directly connected with the function of WSN in SHM: damage detection. We gave the criterion to determine whether a given sensors set can 'SHM-cover' a structure and then proposed two methods to divide the deployed sensor nodes into disjoint sets while each set can 'SHM-cover' the whole structure. Through simulation and experimental results, the effectiveness of the proposed approaches is demonstrated and they show great promise for WSN-based SHM.

References

[1] Mainwaring, A., et al. *Wireless sensor networks for habitat monitoring*. 2002: ACM.

[2] Slijepcevic, S. and M. Potkonjak. Power efficient organization of wireless sensor networks. 2002: IEEE.

[3] Gao, S., C. Vu, and Y. Li, Sensor scheduling for k-coverage in wireless sensor networks. Mobile Ad-hoc and Sensor Networks, 2006: p. 268-280.

[4] Wang, X., et al. Integrated coverage and connectivity configuration in wireless sensor networks. 2003: ACM.

[5] D. Tian, N.G. A coverage-preserved node scheduling scheme for large wireless sensor networks. in Proceedings of First International Workshop on Wireless Sensor Networks and Applications (WSNA'02). 2002.

[6] Zhang, H. and J. Hou, Maintaining sensing coverage and connectivity in large sensor networks. Urbana. 51: p. 61801.

[7] Gupta, H., et al., Connected sensor cover: self-organization of sensor networks for efficient query execution. IEEE/ACM Transactions on Networking (TON), 2006. 14(1): p. 55-67.

[8] Chen, B., et al., Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. Wireless Networks, 2002. 8(5): p. 481-494.

[9] Carle, J. and D. Simplot-Ryl, Energy-efficient area monitoring for sensor networks. Computer, 2005. 37(2): p. 40-46.

[10] Farrar, C. and S. Doebling. An overview of modal-based damage identification methods. 1997: Citeseer.

[11] Juang, J. and R. Pappa, Eigensystem realization algorithm for modal parameter identification and model reduction. Journal of Guidance, Control, and Dynamics, 1985. 8(5): p. 620-627.

[12] Penny, J., M. Friswell, and S. Garvey, Automatic choice of measurement locations for dynamic testing. AIAA journal, 1994. 32(2): p. 407-414.

[13] Goldberg, D., Genetic algorithms in search, optimization, and machine learning. 1989: Addison-wesley.

[14] SAP 2000 HELP GUIDE,[online]: www.csiberkeley.com, Computers and Structures Corp.

Acknowledgement

This research is financially supported by Hong Kong ITC under the project ITS/392/09, and by the Hong Kong Polytechnic University under the project 1-BB6C, and by Hong Kong RGC under the GRF project 5102/08E

Appendix

Each mechanical structure has a number of specific vibration patterns which are called '**modes'** by civil engineering. Each mode has a specific **natural frequency** at which it vibrates. In addition, it has a characteristic '**mode shape**' which defines the mode spatially over the entire structure at this natural frequency. Both natural frequency and mode shape are **modal parameters.**

The modal parameters of a structure can be identified by the measurements of deployed sensors on this structure. For example, if we deploy a total of m sensor nodes on a structure and extract a total of p modes from the measurement of these sensors, the mode shapes of these p modes can be written in a m-by-p matrix $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi} = \left[\boldsymbol{\Psi_1}, \boldsymbol{\Psi_2}, \dots \boldsymbol{\Psi_p}\right] = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1p} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2p} \\ . & . & . & . \\ \phi_{m1} & \phi_{m2} & \cdots & \phi_{mp} \end{bmatrix}$$

where $\boldsymbol{\Psi}_k = [\phi_{1k}, \phi_{2k}, \dots \phi_{mk}]'$ is the mode shape corresponding to the k^{th} mode of the structure. ϕ_{ik} (i = 1, 2, …,m) is the k^{th} mode shape value defined at the i^{th} sensor. Each $\boldsymbol{\Psi}_k$ has a corresponding natural frequency value f_k. Examples of structural mode shapes can be found in Fig.3 and Fig. 10.