

# Dependable Structural Health Monitoring Using Wireless Sensor Networks

Md Zakirul Alam Bhuiyan, *Member, IEEE*, Guojun Wang, *Member, IEEE*, Jie Wu, *Fellow, IEEE*, and Jiannong Cao, *Fellow, IEEE*,

## Abstract

As an alternative to current wired-based networks, wireless sensor networks (WSNs) are becoming an increasingly compelling platform for engineering structural health monitoring (SHM) due to relatively low-cost, easy installation, and so forth. However, there is still an unaddressed challenge: the application-specific dependability in terms of sensor fault detection and tolerance. The dependability is also affected by a reduction on the quality of monitoring when mitigating WSN constrains (e.g., limited energy, narrow bandwidth). We address these by designing a dependable distributed WSN framework for SHM (called DependSHM) and then examining its ability to cope with sensor faults and constraints. We find evidence that faulty sensors can corrupt results of a health event (e.g., damage) in a structural system without being detected. More specifically, we bring attention to an undiscovered yet interesting fact, i.e., the real measured signals introduced by one or more faulty sensors may cause an *undamaged location to be identified as damaged (false positive)* or a *damaged location as undamaged (false negative)* diagnosis. This can be caused by faults in sensor bonding, precision degradation, amplification gain, bias, drift, noise, and so forth. In DependSHM, we present a distributed automated algorithm to detect such types

M. Z. A. Bhuiyan is with the School of Information Science and Engineering, Central South University, Changsha, China, 410083, and the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122. E-mail: zakirulalam@gmail.com.

G. Wang is with the School of Information Science and Engineering, Central South University, Changsha, China, 410083. E-mail: csgjwang@gmail.com (Corresponding Author).

J. Wu is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122. E-mail: jiewu@temple.edu.

J. Cao and X. Liu are with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong. E-mail: {csjcao,csxfliu}@comp.polyu.edu.hk.

of faults, and we offer an online signal reconstruction algorithm to recover from the wrong diagnosis. Through comprehensive simulations and a WSN prototype system implementation, we evaluate the effectiveness of DependSHM.

### Index Terms

Wireless sensor networks, structural health monitoring, dependability, fault detection, fault-tolerance, energy-efficiency.

## I. INTRODUCTION

Wireless sensor networks (WSNs) consist of a number of sensor nodes that can collaborate with each other to perform monitoring tasks. WSNs have been widely deployed on the ground, vehicles, structures, and the like for enabling various applications, e.g., target detection, scientific observation, safety-related, and traffic monitoring [1], [2], [3], [4], [5], [6], [7], [8]. A WSN typically consists of a large number of resource-limited sensor nodes working in a self-organizing and distributed manner. Sensor nodes Applications of WSNs include military sensing, wildlife tracking, traffic surveillance, health care, environment monitoring. Recent work has explored that WSNs can be a compelling platform for engineering structural health monitoring (SHM), due to relatively low-cost, easy installation, and so forth [9], [10], [11], [12], [13]. In a typical SHM system, the interest is in monitoring possible changes (e.g., damage, crack, corrosion) on physical structures (e.g., aerospace vehicles, buildings, bridges, nuclear plants, etc.) and providing an “alert” at an early stage to reduce safety-risk. This prevails throughout the aerospace, civil, structural, or mechanical (ACSM) engineering communities.

Both ACSM and computer science (CS) communities have already addressed numerous challenges/requirements, including data acquisition, compression, aggregation, damage detection, distributed computing. However, there is still an unaddressed challenge: *the application-specific dependability*, which is the ability of a WSN providing application-specific meaningful monitoring results under sensor faults. Particularly, such a system should be able to detect the sensor data faults online and take recovery actions immediately to avoid meaningless monitoring operations. In fact, dependability is highly desired in a WSN-based SHM, as an “alert” about a structural event conveys a serious concern with public safety and economic losses.

On the one hand, SHM algorithms in wired sensor networks used by ACSM are generally

centralized/global-based [11], [14], [15], [16], in which they may not need to seriously consider data collection quality and synchronization errors, etc. This is because they may not often handle data losses or mismatch, as there are no issues like poor wireless connectivity, narrow bandwidth, and energy constraints. The dependability is affected by a reduction on the quality of data when mitigating the constraints. Moreover, once data from the WSN is collected at a centralized base station (BS), it becomes complex to scrutinize all the collected data (including faulty signals).

On the other hand, significant efforts have been made for specific fault types in WSNs [17], [18]. Some prominent schemes, namely, *decision fusion (or 0/1 decision)*, *threshold-based decision*, *heartbeat reception* have been suggested for fault-tolerant phenomenon (such as an event) detection problems [17], [19], [20], [21], [22]. These often use simplified data and few measurements to adequately detect certain faults. However, they are not able to function properly in an SHM system, since SHM algorithms use totally different methods to detect a damage event. For example, the algorithms need raw measured signals rather than the decision fusion, and the analysis of signals (vibration, strain, damping, etc.) that requires a substantial knowledge from ACSM domains (e.g., finite element model updating, Eigen matrix, mode shape properties) [23], [13], [10], [24]. We have evidence from experimental settings that when there is a change in structural health properties (as shown Fig. 1a), 0/1 decision schemes tell sensor 5 is faulty, but they cannot tell what happen (faulty signals or damage event) around sensors 4 and 6. Regarding all these issues above, a question might be posed: is it possible to have a dependable SHM system using WSNs?

The answer is positive. In this paper, we design a dependable and distributed WSN framework for SHM (called DependSHM) that jointly considers ACSM and CS requirements. In DependSHM, we propose an algorithm to detect sensor faults efficiently under the constraints of the WSN. Dependability in WSNs suffers from various types of faults, including, transceiver failure, link errors, security attacks (e.g., collusion), etc [18], [25]. Numerous efforts are being published every day in handling these fault types. Instead, we are interested in some types of sensor faults that are common but difficult to identify: *sensor debonding (when a sensor partially or completely debonds from the host structure)*, *faulty signals*, *faults in offset, bias, precision degradation*, and *the amplification gain factor of signals, noise faults, node missing or failure*.

Most of the sensor data faults fall within these fault models and they directly interrupt a WSN system from detecting damage. Sensors with some of these faults seem to work properly,

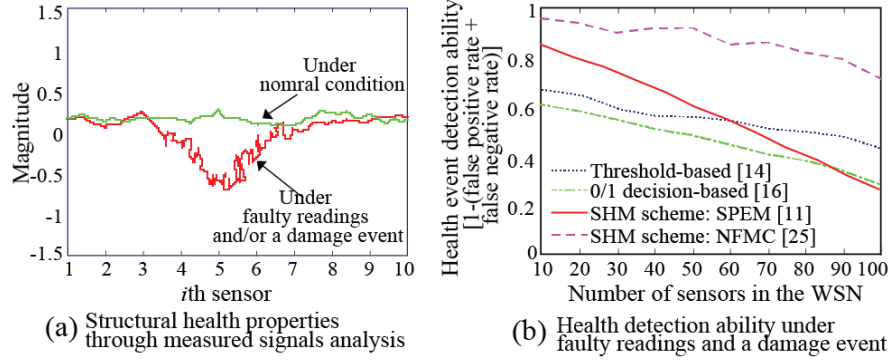


Fig. 1. Investigation of the dependability performance of different schemes in structural health monitoring (SHM).

to communicate to neighbors, to exchange *heartbeats*, but they return incorrect readings or decisions. Under any of the fault occurrences in a practical SHM, we discover a fact that goes to SHM system dependability: *both faulty and non-faulty sensors can generate abnormal signals or decisions* (i.e., remarkable changes in the measured signals). The difficult part is that sensor data, the only available information, will be affected by both structural damage and sensor faults. We further discover an interesting fact that such a possibility can cause an *undamaged location to be identified as damaged (false positive)* or a *damaged location can be given undamaged (false negative)* diagnosis. When we transform these false positive and negative rates into a structural health event detection ability as the performance of system dependability (as shown in Fig. 1b), we find that those decision based and current SHM schemes do not perform well.

We use a new general measurement, *mutual information independence (MII)*, between two signals  $u$  and  $v$  from two different sensors for evaluating results in the absence of the ground truth. We think that mutual statistical information can be used as an *indicator* to decide on a sensor fault detection in conjunction with damage detection. We attempt to reconstruct faulty sensor signals using Kalman filter techniques so that if there is damage, it can be recovered after the reconstruction. This does not require any costly actions, including sensor grouping, faulty sensor avoiding, masking, isolating, or replacing.

Our major contributions in this paper are as follows:

- We study a WSN-based SHM system dependability problem and design DependSHM to address the problem. This task is by no means easy, as it requires multi-domain knowledge and is associated with optimizing WSN resource constraints.

- We propose a non-faulty data collection algorithm, by which we utilize an online faulty sensor detection algorithm based on the function of MII. Although we focus on sensor faulty signals in DependSHM, MII does not rely on a particular fault type.
- In DependSHM, we present a recovery algorithm to reconstruct faulty sensor signals based on the Kalman filter technique. The recovery is directly applicable to any kind of spatially and temporally correlated signals that are caused by numerous sensor faults in a WSN-based SHM system.
- We evaluate DependSHM via simulations using real data sets, adopted from a SHM system deployed on the GNTVT structure [26]. We implement a prototype system developed by the TinyOS [27] running on the Imote2, and verify it on a test structure. The results show that a careful use of recovery from faulty signals in DependSHM is effective and can lead to a dependable WSN-based SHM system.

This paper is organized as follows. Section 2 reviews related Work. Section 3 provides system models and problem formulation. Section 4 presents the DependSHM framework. The faulty sensor detection algorithm is in Section 5. Faulty sensor signal reconstruction is detailed in Section 6. Performance evaluation is outlined in Section 7. Section 8 concludes this paper.

## II. RELATED WORK

**Dependability in WSN-based SHM.** WSNs have been widely suggested and validated in experimentation for SHM system by both the ACSM and CS communities in recent years [10], [11], [13], [14], [15], [24], [9], [28]. Existing schemes already have sufficient contributions to ACSM and CS requirements [2], [9], [10], [11], [12], [13], [14], [15], [16], [24], [29], [30], [31], [32], [32], but they suffer from the dependability problem.

On the one hand, generally data can be corrupted at four stages, namely acquisition, processing and local decisions, wireless transmission, and the final analysis at the BS. Among them, the most important stage is the acquisition stage that can ensure the quality of sensor readings in WSNs at the beginning. The quality is also affected when application-specific requirements are considered, including high-resolution data, raw data, non-faulty data, dependable and real-time decision-making to analyze actual structural health conditions. These additional requirements are traditionally guaranteed by using wired networks. To make WSNs effective alternatives to wired network system instruments, a first step in this direction is SHM system dependability in terms

of detection of faulty sensor readings and a collection of non-faulty readings at the BS, and then dependable monitoring results. In this paper, we take such a step.

**Work from Generic WSN Applications Related to SHM Dependability.** Fault tolerance in WSNs has been studied extensively by researchers in computer science [17], [19], [20], [13], [33], [20], [22]. The application background is largely event/target detection in battlefield surveillance, environment monitoring, etc. The general objective is that when some sensor nodes give faulty readings, how to achieve the correct detection of an event/target over a specific region. Among them, a large part of the schemes on fault detection are off-line and centralized-based. Most schemes rely on various detection methods, including correlation analysis, 0/1 decision, value fusion, decision rules or threshold. Some more details can be found in our earlier work [20], [9], [33], [22].

Although dependability support by fault detection and tolerance problem in SHM looks similar to the problem of making binary “0/1” detection decisions or value fusions [9], [20], [21], it is fundamentally different from them. To check the validity of this assumption and the dependability, we have conducted WSN-based SHM experiments on a physical structure (the settings are described in the later part). For these experiments, faulty signals are injected into sensor nodes to validate the systems capabilities to autonomously detect and tolerate the fault. As shown in Fig. 1, these schemes do not show a satisfactory performance in SHM; specifically, they cannot identify what exactly occurs in a WSN-based SHM. We can see that when there is a change in structural health properties (as shown Fig. 1a, a 0/1 decision scheme [21] tells sensor 5 is faulty, but it cannot tell what happens (faulty signals or damage event) around sensors 4 and 6. There are also changes in signals of sensors 5 and 6. Those existing schemes show here a high rate of *fault positive* and *false negative* rates, resulting in a low system dependability. When we consider WSN-based system dependability as the structural health detection ability, we can see that these schemes show low dependability performance, as shown in Fig. 1b. The methods of detecting faulty sensors by measured signals, removing faulty signals from the measured signals, and then identifying what happens exactly in a structure are different from the methods in those schemes.

Numerous techniques towards the area of fault detection and isolation (FDI) have been proposed, e.g., model-based techniques, knowledge-based techniques, or a combination of both [34]. There are also techniques on fault-tolerant data aggregation that deal with faulty sensor

readings caused by security attacks, such as node compromising, collusion [18], [25]. They use some filtering algorithms (e.g., iterative filtering) at the upper-stream nodes (e.g., cluster head) to remove the faulty signals. Though the algorithms seem to be applicable for our case, we could not justify them. However, it may be difficult to apply such filtering algorithms at a upper-stream node once such a high-resolution big acceleration data from a number of nodes reaches at a upper-stream node and the upper-stream node filters all the raw data. Various constraints in WSNs and application requirements might be an issue in them.

**Work from SHM Applications Related to SHM Dependability.** On the contrary, there are also fault detection schemes from ACSM engineering domains [35], [36], [37], [38]. The concepts in most of them are associated with system failure detection dating back to the 1980s. Here, the failure does not imply to faults in a WSN system, but to faults (e.g., damage) in a physical systems. FDI concepts (described previously) have also been implemented in a number of engineering disciplines, such as ACSM, to improve the availability and reliability of SHM systems [38]. However, these are centralized and computationally-intensive, and usually developed in wired networks.

A noteworthy WSN deployment method for SHM applications called SPEM [16]. SPEM nicely explains the ACSM and CS requirements and is verified on the real structure. It adjusts the quality of sensor locations to better fit WSN requirements; meanwhile, the adjustment satisfy ACSM location-quality requirements. We have verified SPEM under sensor faults in simulations and found that the SHM dependability performance in SPEM drastically decreases from 87% to 28% as the number of sensors in the WSN increases, as shown in Fig. 1b. This is just because of a lack of the dependability support.

To the best of our knowledge, as the first step, we have worked towards the WSN-based SHM application-specific dependability, and have got preliminary results [9], [33]; and this paper is an extension. Our first work is about sensor fault detection algorithm and structural damage event in WSN-based SHM [33] that works on Natural Frequency extraction and Matching in Clusters (NFMC for short), and then tackles faulty sensor readings. However, it has several shortcomings, described in [9]. After improving the shortcomings, we check the dependability performance by NFMC, as shown in Fig. 1b. Its dependability performance falls between 96% and 76%, which is much better than all other schemes. However, such a performance is still not enough to put our confidence in a WSN-based SHM system.

As an extension, this paper includes several aspects. (i) We deal with the *problem of SHM system dependability* and design DependSHM for the problem. (ii) We propose a WSN framework to observe the dependability by removing faulty sensor data from structural damage data and by finding a fault *indicator* based on MII. To ensure the dependability, we devise a new method for tolerance to a missing or failed sensor. We then address the challenge of when both *structural damage* and *sensor fault* occur at the same time, and identify what exactly happens in a structure. Generally, it is not easy to make sure that there is a *faulty sensor* but there is *no damage*. Particularly, we attempt to make DependSHM efficient for recovery from the sensor faults that occur for a short duration; thus, masking or isolating the sensor (as required in NFMC) is not needed. (iii) For the faulty sensor, we present a process flow of the Kalman filter and a graphical representation of it for the signal reconstruction. (iv) The motivation for the dependability, clarifications on system models, and SHM-specific terms are given. (v) A detailed performance evaluation and new results are presented.

### III. MODELS AND PROBLEM FORMULATION

In this section, we provide some necessary models and definitions. Then, we formally formulate our problem.

#### A. Network Model

We assume that a set  $P$  of  $m$  wireless sensors is in charge of performing different types of application tasks (e.g., sensing the vibration, strain, and damping signals, pressure, temperature, etc., in the context of SHM) and sending its measurements to neighboring nodes. A reference 2D building model is shown in Fig. 2a, where sensors (white circle) are deployed on it and a remote monitoring center or BS station location (colored circle) is at a remote place. Fig. 2b shows the traditional WSN framework for SHM (which is similar to the framework in [16]), while Fig. 2c shows the proposed distributed, dependable WSN-based SHM framework, DependSHM.

Consider that the set of sensors is deployed on a structure by finding locations from a set of candidate locations of the structure;  $L = \{l_0, l_1, l_2, \dots, l_m\}$ , where sensor  $i$  is placed at location  $l_i$ , and  $l_0$  is a suitable location of the BS. For high-quality monitoring results, we follow engineering-driven sensor deployment method [13], [16], [39], [40]. Regarding DependSHM in Fig. 2c, sensor  $i$  can be allowed to acquire data, analyze it locally (prepare natural frequency



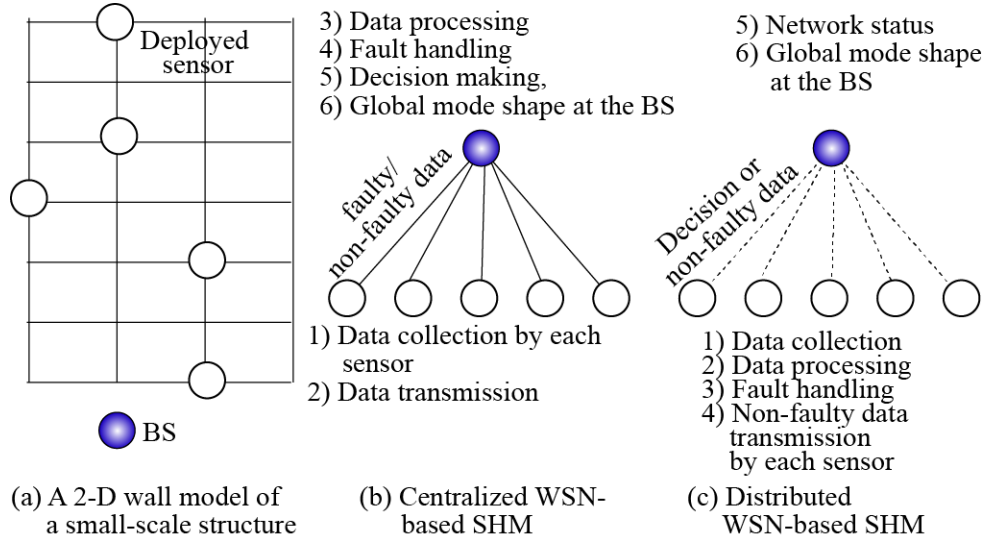


Fig. 2. WSN-based SHM frameworks.

if needed), identify faulty readings, and finally compute mode shapes locally or transmit the non-faulty raw data to the BS (see Appendix B for the natural frequency and mode shape definitions).

Let  $R_{max}$  and  $R_{min}$  be the maximum and minimum communication ranges of a sensor, respectively.  $R_{min}$  is used to maintain local topology, where a number of sensors is allowed to share their signals with their neighbors for damage detection, also used for fault detection. The intention of adopting adjustable communication range is to reduce energy cost for transmission. Note that Imote2 sensor platform supports discrete power levels [41]. Two local topologies can be seen in Fig. 3; each sensor can be overlapped by one or more sensors. When a sensor communicates to the BS directly,  $R_{max}$  is used. Each sensor corresponds to a vertex in a network graph denoted by  $G$ , and two vertices are connected in  $G$  if their corresponding sensors communicate directly. The graph  $G$  is called the communication graph of this WSN.

### B. Sensor Faults

1) *Fault Model*: We focus on the following set of sensor faults that occur in a real WSN-based SHM system:

- Sensor debonding—it is a very common fault in a WSN-based SHM that occurs when a wireless sensor slightly or completely debonds/detaches from the host structure. This affect

is seen in terms of accurate vibration capturing from the structural response.

- Faulty signals—these are caused by precision degradation, breakage, etc., especially in vibration signal capturing. For example, a sensor reports a constant value for a large number of successive samples, where the constant value is either very high or very low compared to the “normal” or “reference” value.
- There are faults in offset, bias, and the amplification gain factor of signals. For example, the offset fault is due to calibration errors in sensor signals, which differs from the normal value by a constant amount, but the sensor readings still exhibit normal patterns.
- Noise faults are caused by longer duration noisy readings that affect a number of successive samples.
- There is also node missing or failure.

Sensors with these faults seem to work properly (except for the last type), to communicate to neighbors and exchange messages, but they return incorrect values or decisions. We tackle these faults in DependSHM.

2) *Fault Detection Model:* We assume that sensor  $i$  exchanges its signals with its neighbors in each sampling instant  $t$  in  $T_d$ , where  $T_d$  is the monitoring round, i.e., the time is divided into discrete sampling periods. In each period,  $i$  broadcasts its current readings to one-hop neighboring nodes using  $R_{min}$ . Besides the readings,  $i$  may be enabled to make a decision locally or recover the mode shape and forwarded it to the BS. The signal  $y_i^t$  measured by a faulty sensor at  $t$  is subject to noise effect  $\sigma$ . Then, let  $y$  be the measured output reading that would be transmitted to the neighboring nodes, which is given as follows:

$$y = y_i^t + \sigma \quad (1)$$

The measurement noise,  $\sigma$ , for non-faulty sensors may be a small random noise in practice. However, it also can greatly affect damage event detection. Consider that a subset  $N$  of sensors is non-faulty at time  $t$ . In SHM, when all of the sensors are non-faulty, it is easily possible to estimate the mode shape from the signals. However, if a sensor is faulty, it is possible to produce predicted signals for the mode shape by using neighbors’ signals, correlation statistics, and the extent of  $\sigma$ . Suppose that damage may occur at time  $t$ , anywhere in the structure. A subset  $D \subset P$  of sensors around the damage area is possibly able to detect the damage. Some of the sensors from  $D$  may provide faulty signals. Thus, to detect faulty sensors, the sensors in

$D$  split into two further subsets:

$N$  = sensors assumed to be non-faulty

$F$  = sensors assumed to be faulty

Note that these two sets are disjoint so that

$$N \cap F = \{\} \text{ and } N \cup F = D, \text{ where } D \subset P$$

Generally, sensors anywhere in the WSN can be faulty/failed. However, we put an emphasis on continuous monitoring and on those sensors whose signal have changed significantly (due to a fault/damage).

**Definition 1** [MII: Mutual Information Independence]. *A function denoted by  $\omega()$  is defined by the quantify of how much the measurement correlation between sensor nodes in  $N$  and sensor nodes in  $F$  deviate from the correlation model.*

We state the MII function as an indirect vibration signal measurement. Assume that a prior correlation model  $C$  of  $x_P^t$  presents [42].  $C$  can be given as a reference set by all immediately-stored data in the sensor local memory after WSN system initialization. The MII function between two signals of sensors  $i$  and  $j$  at time  $t$  in  $D$ , is given as follows.

$$\omega(y_i, y_j, C) \tag{2}$$

Consider that the sensors in  $N$  and  $F$  capture vibration signals and broadcast their measured signal sets  $y_N$  and  $y_F$ , respectively. Thus, MII between the two sets of signals of  $N$  and  $F$  is given as follows:

$$\omega(y_N, y_F, C) \tag{3}$$

Given  $R$  consecutive signals, the MII function estimates the correlation between  $y_N$  and  $y_F$  at time  $t$  as:

$$\Delta(N, F) = \sum_{t=1}^R \omega(y_N, y_N, C) - \sum_{t=1}^R \omega(y_F, y_N, C) \tag{4}$$

$\Delta(N, F)$  is achieved by reducing the deviation between non-faulty sensors in  $N$ , and by maximizing the deviation between non-faulty sensors in  $N$  and faulty sensors in  $F$ .  $D$  can be controlled by the system user considering  $R_{min}$ , neighborhood size, or network density.

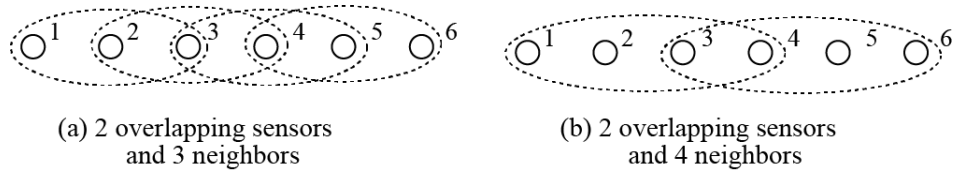


Fig. 3. Topologies with different numbers of overlapping sensors in different neighborhoods.

### C. Energy Cost Model ( $cost(e_i)$ )

One important objective is to minimize the energy cost of the network. Let  $cost(e_i)$  denote the total energy cost of sensor  $i$ , including measurement, computation, transmission, and overhead. Consider a shortest path routing model [13], [16]; there is a path from sensor  $i$  to neighboring sensor node or the BS  $j$ :  $q = z_0, z_1 \dots z_k$ . Sensor  $i$  propagates the data to them. We can find the  $i$ th hop sensor on each path and calculate the amount of traffic that passes along on the paths within each round of monitoring data collection ( $T_d, d=1,2,\dots, n$ ). Then, the  $cost(e_i)$  can be decomposed into the following four parts:

$$cost(e_i) = e_T + e_{comp} + e_{samp} + e_{oh} \quad (5)$$

Here, i)  $e_T$  is the maximum energy cost for data transmission over a link between a transmitter and a receiver. ii)  $e_{comp}$  is the energy cost for processing data locally, e.g., computing equation (6). iii)  $e_{samp}$  is the energy required for a sampling cost of  $M$  data points [33], [43]. iv)  $e_{oh}$  is the additional overhead for some causes, e.g., fault detection, signal reconstruction, copying data to a local buffer, and network latency. See Appendix A for an extended version of the energy cost model.

### D. Problem Statement

**Given:** a set  $P$  of  $m$  sensors and a BS, which are deployed over a physical structure and are involved in monitoring structural health that is reported to the BS.

**Find:** a subset  $D (\subset P)$  of sensors that involves in detecting structural damage event such that the sensors in  $N(\subseteq D)$  are non-faulty and the sensors in  $F = D - N$  are faulty, subjected to the following constraints:

- Data delivery:  $\forall q = z_0 \dots z_k$  used for data delivery, where  $\overline{q[j-1]q[j]} \leq R_{max}, j = 1 \dots k$ ;
- Connectivity:  $\forall i = 1 \dots m, \exists q = z_0 \dots z_k, q[0] = l_i, q[k] = l_0$ ;

- Structural modeling.

**Objectives:** minimize  $\Delta(N, F)$  and minimize  $cost(e_i)$ , and maximize dependable mode shape computation.

#### IV. DEPENDSHM FRAMEWORK

In this section, we present DependSHM, the dependable WSN-based SHM framework. It is divided into three stages: distributed framework for structural health event detection (e.g., damage, crack, corrosion), faulty sensor detection, and faulty reading reconstruction.

##### A. Basics of Structural Event Detection Algorithm

The central focus of SHM is the detection and localization of events (considering damage) within various types of structures. Generally speaking, SHM techniques rely on measuring structural responses to ambient vibrations or forced excitation. Ambient vibrations can be caused by earthquakes, wind, passing vehicles, or forced vibrations, or can be delivered by hydraulic or piezoelectric shakers. It can also be caused by a damage occurrence.

A variety of sensors, e.g., accelerometers, strain gauges, or displacement can be used to measure structural responses. SHM techniques infer the existence and location of damage by detecting differences in local or global structural responses before/after a damage occurs. The responses are usually comprised of frequencies in the tens of Hz, and can be sensed using relatively inexpensive low-noise MEMS-based accelerometers. Increasingly, the ACSM communities are becoming interested in active sensing techniques [10], which measure structural responses to forced excitations. In order to identify the damage, two necessary structural characteristics are important: mode shape ( $\Phi$ ) and natural frequency ( $f$ ).

##### B. Mode Shape Computation at Each Sensor

Each type of structure (aircraft, building, bridge, etc.) has a tendency to vibrate with much larger amplitude at some frequencies than others.  $f$  and  $\Phi$  rely on structural material properties, geometry, and assembly of its constituent members. We use *state space model*, which is widely accepted by ACSM communities for capturing structural dynamics to compute  $\Phi$  [23], [10]. We mention the process we use for  $\Phi$  computation, which is the same process is used for

designing faulty signal reconstruction (after fault detection). The state space matrices for a finite-dimensional linear structural dynamic system can be succinctly obtained by the linear differential equation:

$$M\ddot{x} + Kx + \sigma = F(x, t) \quad (6)$$

Here, function  $F(x, t)$  is the response of the structure over a period of interest at certain sensor locations, where  $x$  is the structural response at time instant  $t$ .  $M$  and  $K$  are the matrices of mass and stiffness coefficients of the various elements of the structure, respectively<sup>3</sup>.  $\sigma$  is the signal to noise. In (6), damping is neglected for an advantage in detection (see Appendix C for details).

In traditional SHM algorithms, the state space model is computed in a centralized/global fashion. We argue that it could be quite costly for the resource-limited WSN. Considering SHM as a big data application, to make use of the WSN for SHM, we mitigate this problem by allowing each sensor to work only with local structural responses rather than the global. For this purpose, We modify the model, considering the implementation of the model for each sensor location. To reduce the system order, a transformation of the state space into mode coordinates is necessary. This transformation is derived by determining a diagonal matrix, which contains a certain number of Eigen frequencies covering the natural frequency ranges of interest. By applying the mode transformation, based on the mass normalization  $\{\phi_i\}$ , the refined  $\Phi$  is given by:

$$x = \Phi h \quad (7)$$

where  $h$  is the mode participation factor. We have,

$$\ddot{h}_i + \delta_i h_i = H_i \quad (8)$$

(7) implies the refined response of the structure that is a sum of the responses in each mode. We can express it more explicitly as follows:

$$x = \sum_{i=1}^n h_i \phi_i \quad (9)$$

where  $\delta_i$  is the  $i$ th eigenvalue, and  $H_i = \phi_i^T f$  is the  $i$ th mode of responses under force or ambient excitation input. The summation is given over all of the  $n$  modes of the structure that

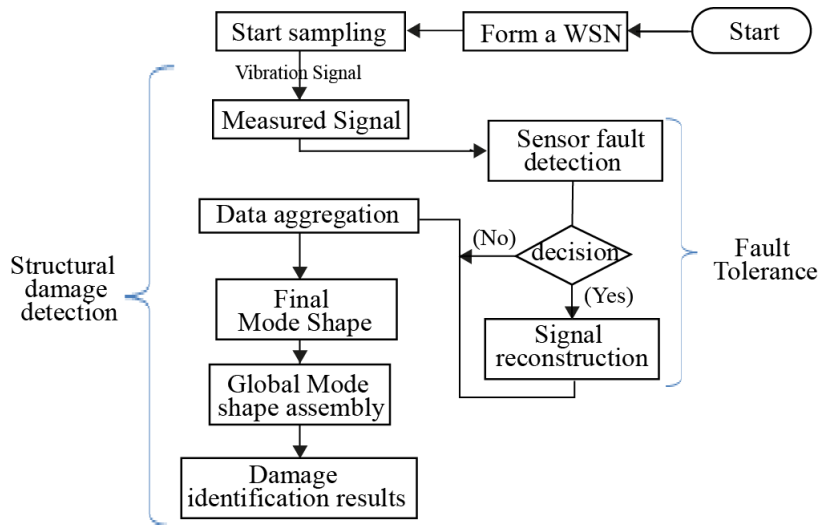


Fig. 4. Providing the dependability in the WSN Framework: the step by step process of sensor fault tolerance and damage detection.

are measured by the individual sensor. Typically, only the lower modes are important because the force excitation is concentrated in these modes. Each sensor extracts its local  $\Phi$  that can be used for both faulty signals and damage event detection. The method of extraction, including the difference between  $f$  and  $\Phi$  can be found in Appendix B.

*Is it Possible to Compute  $\Phi$  under Sensor Faults:* In practice,  $\Phi$  is greatly affected by a faulty sensor signals (see Appendix B for more detail), especially when a sensor is placed at a optimal location [13], [16]. If a signal is detected as faulty, the measured signal is reconstructed directly for the actual mode shape, by collecting the signals from the sensor location or reference signals. We use neighboring sensor nodes' signals for detecting a faulty sensor and reconstruct its signal. In this work, not all the sensors' signals using (6) will be measured. The measured output of a sensor  $i$  at time  $t$ ,  $y_i^t$ , can be obtained by:

$$y_i^t = \mathbf{Q}x \quad (10)$$

where  $\mathbf{Q}$  is the measurement matrix.

### C. General Overview of DependSHM

Fig. 4 summarizes the whole WSN framework for SHM. Once the WSN starts operating, in each monitoring round ( $T_d$ ), a number of signals is measured at each sensor. Based on the

measured signals, each sensor identifies faulty sensors by using MII. If any faulty sensor is detected, the neighboring sensors reconstruct the signals, while a faulty sensor itself can do the same task if it still works. It is highly possible that a sensor exhibits faulty behavior temporarily in many cases. However, if a sensor fails or it is missing from its location, we still guarantee the signal reconstruction for the sensor. We do not assume to isolate or deploy a new sensor as it is a costly task doing so.

In addition, for the high-quality SHM, we must provide monitoring information of each sensor location, since ACSM communities mostly deploy sensors at optimal locations [13], [16]. Each sensor locally computes the final  $\Phi$  and transmits to the BS directly, which is a relatively small amount of data. The BS assembles all the received final  $\Phi$ s and identifies the structural damage. There is a high possibility that the BS does not receive any of  $\Phi$  caused by data packet-loss or others. If sensor forwarding fails, the BS still has the final results received through the neighboring sensor nodes. In addition, we allow each sensor to keep the final results in the local memory until a sensor receives an acknowledgment from the BS. Each set of raw  $\Phi$  is of a number of KBs while each final result or  $\Phi$  computed by a sensor is a number of bytes. In this framework, each final  $\Phi$  received from the neighbors is not processed.

## V. FAULTY SENSOR DETECTION

This section describes non-faulty sensor data collection and faulty sensor detection algorithms, according to the model described earlier.

### A. Data Collection and Faulty Sensor Indication

We assume that sensors are likely to generate abnormal signals. The signals are measured by the vibration, which may be incorrect compared to the neighbors, previous signals, or reference signals. We first show data collection at every sensor in the WSN. A subset of sensors, say  $D$  of sensors that are in a sensor's  $R_{min}$ , share their data with each other, and participate in faulty sensor detection.

Algorithm 1 simply presents the data collection method based on the neighborhood. While theoretically this procedure involves multi-hop communication, consider the fact that for SHM application, the radio communication range of current sensor nodes exceeds the area in which sensors gather signals. We limit sensors to communicate within the one-hop neighboring nodes.



---

**Algorithm 1: Non-faulty Data Collection for Damage Detection**


---

**Input:** a neighborhood with a bounded degree,  
 $t \leftarrow$  transmission at a time slot

**Step1:** **for**  $i = 1$  to  $m$  **do:**  
node  $i$  acquires data and buffers it

**Step2:**  $i$  transmits its data to its neighboring node  $j$

**for**  $t = 1$  to  $n$  **do:**  
**for** each node  $i$  **do:**  
**If**  $i$  has data not forwarded to  $j$  **then**  
transmits a new data to its  $j$  in  $t$

**Step3:** call Algorithm 2

**Step4:** data aggregation (extracting frequency sets, compute  $\Phi^k$ )  
compute final  $\Phi$

**Step5:** transmit the final  $\Phi$  toward the BS

**return** health status of every sensor location

---

We think that multi-hop communication is not mature enough. The nodes that are one-hop away from the BS will directly send the data; otherwise, the data is sent through one or more intermediate nodes. In Step1 of Algorithm 1, every sensor acquires signals captured from vibration responses of the structure, and buffer them temporarily. Then, it transmits and receives the measured signals. The sensors check if there are any faulty sensors, i.e., a sensor with faulty signals via Step2.

Step3 executes Algorithm 2. When a remarkable change appears in a sensor's signals, there is a possibility that a sensor is faulty. The MII is used to detect faults. Let us consider the statistical dependency between the two sensors' signals quantified by MII.  $\omega$  measures the information about one sensor that is shared by another sensor in the set of signals in  $D$ . It is seen that  $\omega$  changes as soon as a sensor fault occurs, because the faulty signal is not present in the reference or other sensor signals.

We use a joint Gaussian distribution based correlation model. Multivariate Gaussian distribution has been used to accurately model the correlation of many types of signals in literature [22]. Each signal is broadcast to sensors in  $D$ , where  $i$ th sensor signal  $y_i^t \in y_D^t$  and  $j$ th sensor signal  $y_j^t \in y_D^t$ ,  $i, j \in D$  and  $D \subset P$ . For simplicity,  $y_i^t$  as  $u$  and  $y_j^t$  as  $v$  are denoted hereafter.

Hence, it would be worth considering how to find joint probability density between two signals.

The statistical dependency/independency between the two Gaussian distributed time signals  $u$  and  $v$  can be expressed in the form of the joint probability density  $p(u, v)$  of signals:

$$p(u, v) = \frac{1}{2\pi\tau_u\tau_v\sqrt{1-\rho_{uv}}} e^{-\frac{1}{2(1-\rho_{uv}^2)} \left[ \left( \frac{u-\mu_u}{\tau_u} \right)^2 - 2\rho_{uv} \frac{(u-\mu_u)(v-\mu_v)}{\tau_u\tau_v} + \left( \frac{v-\mu_v}{\tau_v} \right)^2 \right]} \quad (11)$$

where  $\mu_u$ ,  $\mu_v$ ,  $\tau_u$ , and  $\tau_v$  are the means and the standard deviations of the signals  $u$  and  $v$ , respectively.  $\rho_{uv}$  is the correlation coefficient between the two signals. It is given by:

$$\rho_{uv} = \frac{E \{(u - \mu_x)(v - \mu_y)\}}{\tau_u\tau_v} \quad (12)$$

The correlation coefficient can also sometimes be used to determine if two signals are statistically independent. On one hand, if  $|\rho_{uv}| = 1$ , there is a strong correlation between the two signals. On the other hand, if  $|\rho_{uv}| = 0$ , the two signals are not correlated. The correlation can be interpreted as a weak form of statistical dependency. In [20], it is shown that two random variables, which are not correlated, can even so be statistically dependent. This is why we take the statistical dependency or independency. The product of the marginal densities  $\rho_u$  and  $\rho_v$  of the signals  $u$  and  $v$ , respectively, is given by:

$$p(u, v) = p(u)p(v) \quad (13)$$

If the expression in (11) is equal to the product of the marginal densities in (13), the signals are completely independent. One possibility to quantify the statistical dependency between two signals is to calculate the MII of them, as follows:

$$\omega(u, v, C) = \int \int p(u, v) \log \frac{p(u, v)}{p(u)p(v)} du dv \quad (14)$$

The base of the logarithm determines the units in which information is measured. (14) shows that if  $u$  and  $v$  are independent,  $\omega$  becomes zero. A forward approach is to divide the range of  $u$  and  $v$  into finite bins and to count the number of sampled pairs of  $h_o = (u_o, v_o), o = 1, 2, \dots, n$ , falling into these finite bins. This count allows for approximately determining the probabilities, replacing (15) by the finite sum:

$$\omega_{bin}(u, v, C) = \sum_{a,b} p_{uv}(a, b) \log \frac{p_{u,v}(a, b)}{p_u(a)p_v(b)} \quad (15)$$

where  $p_u(a) \approx n_u(a)/n$  and  $p_u(b) \approx n_u(b)/n$  are the probabilities based on the number of points  $n_u(a)$  and  $n_v(b)$  falling into the  $a$ th bin of  $u$  and the  $b$ th bin of  $v$ , respectively. The joint

probability is  $p_{uv}(a, b) \approx n(a, b)/n$  based on the number  $n(a, b)$  of points falling into box nos.  $a, b$ . MII is non-negative and symmetric:

$$\omega(u, v, C) = \omega(v, u, C) \geq 0 \quad (16)$$

The MII for all possible combinations of sensor outputs  $y_r$  and  $y_s$  (except  $r = s, i = 1, 2, \dots, r, j = 1, 2, \dots, s$ ) is computed, which leads to an  $\omega$ -matrix for all combinations of  $r$  and  $s$ . The basic idea is that the MII changes when a sensor fault  $f_r$  is present. Suppose that it is in the  $r$ th channel or index:

$$\tilde{y}_r = y_r + f_r \quad (17)$$

This fault appears only in the  $r$ th channel. Thus, we should expect that all combinations with index  $r$  should show a reduction of  $\omega$ . This allows us to localize the faulty sensor. One or more faulty sensors can be simultaneously detected in the same way. One possibility to visualize the faulty sensor is to use the relative change as a sensor fault indicator  $\lambda_{y_r}^\omega$ :

$$\lambda_{y_r}^\omega = \frac{|\omega_{y_r} - \omega_{ref}|}{\omega_{y_r}} \quad (18)$$

where  $y_r$  is an actual data set and the lower index *ref* is one reference data set. The method based on MII is able to detect sensor faults in different combinations of them.

### B. Algorithm 2: Faulty Sensor Detection

Under centralized detection, the BS handles the damage and faulty sensor detection process. In each decision cycle, the BS makes a decision about the faulty sensors, solely based on the  $k$  most recent signals received from each sensor. The BS computes the MII for each signal, and chooses the signal with the maximal independence for fault detection. This detection is not suitable for resource-constrained WSNs. For example, if each sensor needs to send all its signals to the BS (where each sequence of signals or raw natural frequencies can be from  $X0\text{kb}$  to  $X000\text{kb}$ ,  $X = 1, 2, \dots$ ), the centralized WSN may not be able to operate for a given period of time. In a large-scale WSN deployment, the situation becomes serious. After capturing data at high frequency in SHM, sensors should reduce data before transmission.

In contrast, the faulty sensor detection (see Algorithm 2) can execute in a distributed manner where each sensor makes a decision on the collected signals locally, as described earlier. In the algorithm, if the local decision on a sensor's signals,  $\lambda_{y_r}^\omega > 0.5$ , the sensor is faulty. This

---

**Algorithm 2: Distributed Faulty Sensor Detection**


---

Decision:  $(\lambda_{n_r}^\omega \leq 0.5$ : non-faulty),  $(\lambda_{n_r}^\omega > 0.5$ : faulty)

**for** each sensor  $i \in N$  where  $N \subseteq D$  **do**:

//set initial decision as a uniform distribution

$(\lambda_{n_r}^\omega)_i \leftarrow 0$  // each sensor  $i$  is non-faulty

**loop**

**for** each sensor  $i$  **do**:

$(\lambda_{n_r}^\omega)_{j(\text{neighbor})} \leftarrow$  receive neighbor  $j$ th decision

samples  $\leftarrow N$  samples from  $(\lambda_{n_r}^\omega)_{j(\text{neighbor})}$

**for** each sample  $h$  **do**: //  $h$  is an index

$R_h \leftarrow$  non-faulty sensors

$F_h \leftarrow$  faulty sensors

$f_h \leftarrow (R_h, F_h)$  //Eq. 3

**for** each sensor  $i$  **do**:

**if**  $\lambda_{n_r}^\omega > 0.5$  **then**

$i$  marks/reports itself as a faulty sensor node

**if**  $i$  does not transmit the decision **then**

$j$  marks/reports about  $i$  as a faulty sensor node

call Algorithm 3 //  $i$ th sensor signal reconstruction

**end**

---

means that MII is high on the sensor's measured signals. The distributed method only requires neighbors to be synchronized. In addition, the detection is almost immediate and online, since a sensor does not need to wait for the signals from sensor nodes at more than one hop away. Moreover, the detected faulty signal set is not forwarded toward the BS; thus, the communication cost is relatively low. The energy cost becomes lower.

MII does not rely on particular fault types. The algorithm 2 based on MII is able to detect different kinds of faults (as modeled before). However, it may fail to detect a node missing or failing. We provide Appendix E for handling the node failure or node missing.

## VI. FAULTY SIGNAL RECONSTRUCTION

In this section, we propose a Kalman Filter technique (KF) for faulty wireless sensor's signal reconstruction. The KF has received extensive attention to describe the recursive solutions of predicting state variables for linear systems [44]. We consider it, as it can generate the best estimation if the optimal filter is linear among all the linear observers, because it minimizes the

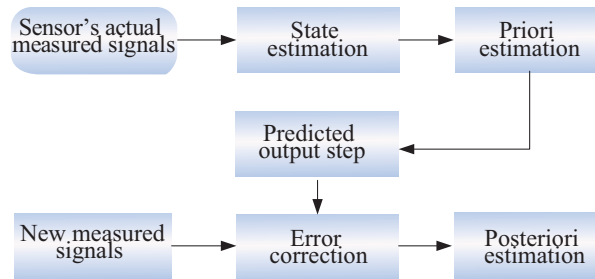


Fig. 5. Sensor signal process flow under the KF.

error covariance. In ACMS engineering domains, the KF has been studied for on-line damage detection [29]. Here, we utilize it for faulty sensor signal detection.

#### A. Kalman Filter in State Space Representation

The description of KF is made with the help of the state-space representation of the structural system, as described in Section 4.2. One radical concept of the KF is that the state estimation is recursively corrected by the actual physical system outputs. Then, using (6), the equation of motion for time discrete and time invariant cases are given as follows:

$$\begin{aligned} l_t &= M_{t-1}z_{t-1} + K_{t-1}u_t + \sigma_{t-1} \\ m_t &= M_t z_t + K_t u_t + \sigma_t \end{aligned} \quad (19)$$

$u_t$  is the excitation at a specific frequency at time  $t$ .  $M_t$  and  $K_t$  are transition matrices. The signals  $\sigma_{t-1}$  and  $\sigma_t$  represent the measurement noises, respectively (refer to Fig. D in Appendix D for the state-space equation-based KF). When measuring the responses of a dynamical structural system by sensors, the actual signals produced by the sensors are contaminated by noise due to internal manufacturing defects, physical interference, or external environmental effects. According to features of the KF, we assume that every measurement from the wireless sensors contains noise; thus, if the noise measurement is zero, the KF collapses. Setting the mean of noise as zero is a common practice:  $E[\sigma_t] = 0$ . Noises are assumed to be independent of each other, and are normally distributed with covariance matrices,  $c_v = [\sigma\sigma^T]$ .

The underlying KF information is that KF is a recursive algorithm consisting of a loop (see Fig. D ) which is passed through for each time instant  $t$ . The estimation of the system state for  $t$  is determined from the weighted average of the *actual measured value* at time instant  $t$  and the *prediction of the system states* for this time instant. The weight factors of this average are determined from estimated uncertainties in each loop, which are also connected to the predicted

system state and to the new measured value. The lower the uncertainty, the higher is the weight factor; i.e., Kalman gain ( $K_t$ ). The uncertainty is calculated with the help of covariance matrices [45].

We shortly describe the faulty signal reconstruction process. At first a *priori* state estimate  $P_k$  for the state vector  $l_t^{pr}$  of the system are estimated [45]:

$$\begin{aligned} l_t^{pr} &= M_{t-1}z_{t-1} + K_{t-1}u_t \\ P_k &= M_{t-1}z_{t-1}M_{t-1}^T + c_v \end{aligned} \quad (20)$$

After getting the measured value  $l_t^{pr}$ , a *posteriori* state can be estimated in the correction step, see (21). For the *posteriori* estimation, the difference between the measured and estimated signals are weighted by the Kalman gain factor  $K_k$ .

$$l_t^{post} = l_t^{prio} + K_t[m_t - M_t l_t^{prio} - K_t u_t] = l_t^{prio} + K_k[m_t - m_t^{prio}] \quad (21)$$

$$K_k = P_k M_t^T [M_t P_k M_t^T + c_v]^{-1} \quad (22)$$

The *priori* estimated error covariance  $P_k$  in the prediction step is used to update the Kalman gain factor in (21) whereas  $P_k$  itself is updated by the a *posteriori* estimated error covariance. The KF is used here based on the Matlab implementation, which delivers the optimum Kalman gain ( $K_k$ ) together with the steady state error covariance matrix ( $P_k$ ).

For the system state estimation using the KF, the process and the measurement noise covariance are determined. Since the system input is assumed to be unknown and the model uncertainty is high, the process noise covariates are set to high values. The values for the measurement noise covariance matrix will be determined a *priori* by the first estimation of the measurement error  $m_t - l_t^{post}$ , as given in (21). The overall process flow of the KF is illustrated in Fig. 5.

### B. Sensor Signal Reconstruction Algorithm

If there is a sensor detected as faulty using Algorithm 2, the sensor signal reconstruction algorithm (see Algorithm 3) is used by the neighboring sensor nodes. The basic idea of the signal reconstruction is as follows. When a sensor signal does not correspond to the modeled system (monitoring its location) and it was erroneously assumed that this signal has a low measurements noise covariance, then the signals from the other sensors cannot be correctly reconstructed, and the difference between the measured and estimated signals will be high. If the value of the covariance for the faulty sensor signals is set as high, then the KF will reconstruct

---

**Algorithm 3: Sensor Signal Reconstruction**


---

- Step 1: Detect the faulty sensor by changes of indicator  $\lambda_{nm}^{MI}$  (Algorithm 2)
- Step 2: Recreate the state space model of the structure w.r.t. the sensor location (including faulty or nonexistent sensor) on the structure [see  $M$  and  $K$  matrices in Eq. (6)]
- Step 3: Determine the process and measurements noise covariance matrices //(see Eq. (20))
- Step 4: Set the covariance value(s) of the assumed faulty sensor(s) to be high
- Step 4: Solve the state space equation of motion through KF with the estimated state vector  $l_t^{Post}$  (in signal correction step). The KF is driven by the input  $u_t = 0$  and the measured signals  $y_t^i$
- 

all the signals, including the incorrect signals, with the help of the other signals and the model. In this manner, it is possible to reconstruct more than one signal simultaneously. The number of signals that can be reconstructed rely on the number of neighboring nodes in case of the distributed system, all of the nodes in the network in case of the centralized system, and on the quality of the model. For a better understanding of Algorithm 3, the procedure is broken into several steps.

By means of Algorithm 3, when just one sensor node does not work properly, it is possible to identify and reconstruct it only with the help of KF. For this purpose, Steps 2 to 5 are applied. Here, Steps 3 to 5 have to be calculated several times. The number of loops over these steps corresponds to the number of neighboring nodes in the case of the distributed WSN or all of the nodes in the case of the centralized WSN. In each loop, the measurement variance of one sensor is set to be high. In addition, we attempt to detect a missing sensor and construct its signals that can be found in Appendix E.

## VII. PERFORMANCE EVALUATION

### A. Simulation Studies

1) *Methodology*: We conduct comprehensive simulations using MATLAB to evaluate DependSHM that includes the faulty sensor detection methods and signal reconstruction algorithm. We use

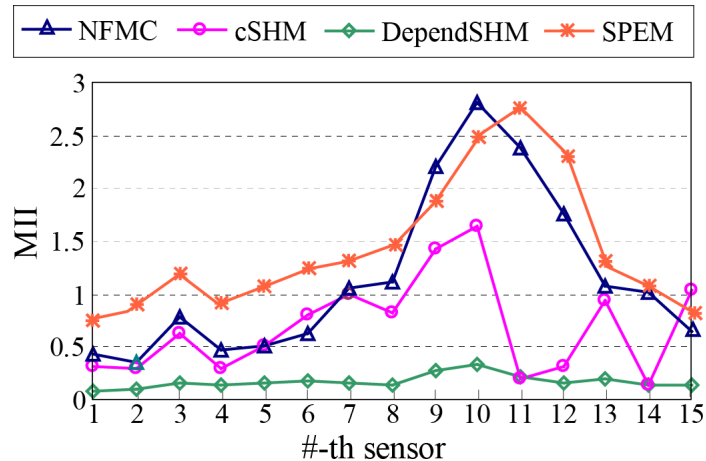


Fig. 6. Performance of different fault detection methods: achieved MII under sensor faults.

real data sets collected by the SHM system employed on the high-rise GNTVT [26] and a SHM toolsuite [46]. We use the data sets for the 100-sensor case in our simulations. We perform the WSN deployment via our WSN-based deployment scheme suggested in [13], which is supported by the ACSM engineering deployment methods [16]. The simulation environment is a  $450 \times 50$  sensing field regarding structural environment, e.g., bridge, building, aircraft.

The background data is simulated as vibration influenced by the 100 sensor locations in the field. A random Gaussian noise is added to all the data. The mean of the noises is zero, and the standard deviation is 10% of the real signals. From the data sets, a set of data is used as *reference data* to train the joint distribution, and another set of similar data is used for testing. A random Gaussian noise is added to all the data. The mean of the noises is zero, and the standard deviation is 10% of the real signals. From the data sets, a set of data is used as *reference data* to train the joint distribution, and another set of similar data is used for testing. The noise is present in both the data sets. Thus, the trained correlation model reflects the noises. In the distributed detection method, each sensor makes a decision based on signals received from neighboring nodes within  $R_{min}$ . After a sensor receives a decision, it recomputes its MII and chooses to change its decision accordingly. The energy cost and routing models described in Section 3 are used for evaluation.

For comparison, we implement other three schemes, including SPEM [16] and NFMC [33]. We compare their performance with DependSHM. We consider two schemes for observing



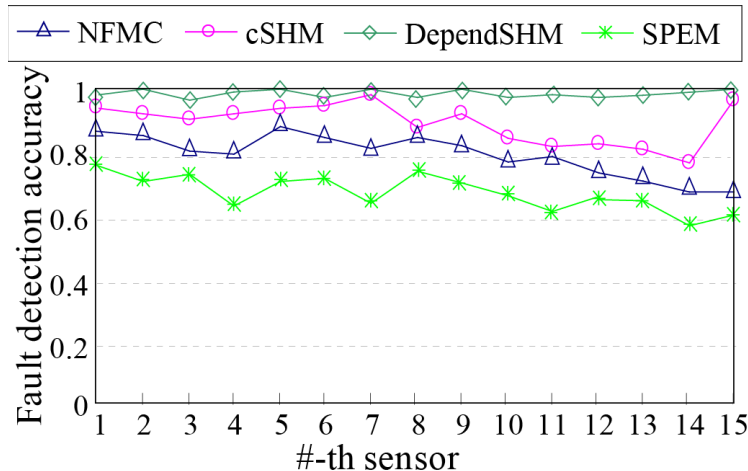


Fig. 7. Performance of different fault detection methods: the fault detection accuracy.

the performance of our fault detection and tolerance methods: i) distributed fault detection under localized data processing (DependSHM); ii) centralized fault detection under localized data processing (cSHM). SPEM is a WSN deployment method for SHM that nicely explains CS requirements and is verified on the GNTVT. It adjusts the quality of sensor locations to better fit WSN requirements. It is a centralized data processing method. We intend to verify its performance under fault detection and tolerance support and compare with DependSHM. NFMC is our preliminary fault detection and tolerance WSN-based SHM scheme, which is based on the natural frequency extraction and matching.

Using simulation results, we compare DependSHM with them in several aspects under the fault injections: i) fault detection accuracy; ii) dependability in terms of detection ability and mode shape ( $\Phi$ ) recovery, etc., and iii) energy cost of the WSN. Here, the *detection ability* is the rate that is calculated by the percentage of successful faulty sensor detection to the percentage of the amount of the sensor fault injection. This includes both the false positive and false negative occurrences. Here, false positive cases are recorded as an *undamaged location* of the structure is identified as a *damaged location*, and false negative cases are recorded as a *damaged location* of the structure is identified as an *undamaged location*.

2) *Results*: In the first set of simulations, we implement all three schemes under the sensor fault injection (through modifying a number of sensors' signals randomly in the data sets). A fraction of the sensor nodes is randomly selected and the modified faulty signals are fed

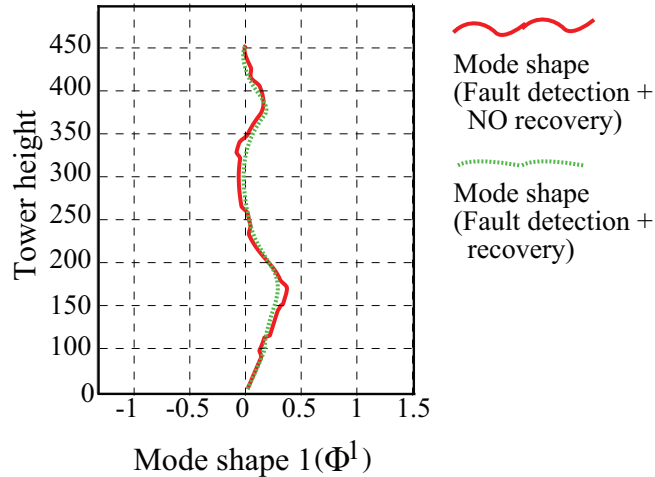


Fig. 8. Observation on the actual mode shape curvature vs. mode shape curvature in cases of faulty sensors and recovery (signal reconstruction).

into their acquisition modules. We vary the number of faulty sensors from 15% to 25%. Each sensor node broadcasts its readings towards the neighboring sensor nodes. Each of the faulty readings is replaced by a random number independently drawn from a uniform distribution in the deployment field (0, 450). Fig. 6 shows MII achieved by the four schemes. Out of them, DependSHM achieves the smallest value, followed by cSHM method. SPEM method performs poorly, since it requires centralized data processing and shows a significant amount of data packet loss during transmission. Due to heavy data losses, its performance on the MII is low. Nevertheless, NFMC still outperforms SPEM in many sensor fault detection cases.

Fig. 7 depicts the fault *detection accuracy*, which is computed as  $\text{accuracy} = (\text{true positive} + \text{true negative})/\text{all}$ . The detection accuracy in DependSHM is about 98%, which outperforms others. In SPEM, the detection accuracy is poorer (less than 80%) than that of others, while it is from 75% to 85% in NFMC. One major cause is that peak natural frequency signals used in NFMC and DependSHM achieve higher MII. However, we experience that the fault detection accuracy rate becomes lower in NFMC and SPEM than in DependSHM and cSHM, as the number of faulty sensor nodes in the WSN increases.

**Dependability Verification.** We next observe the structural health condition using WSNs in simulations. We estimate mode shape ( $\Phi$ ) curvature under sensor faults and the signal reconstruction of the faulty sensors. We recover the first  $\Phi$  (see Fig. 8) of the simulated structure with 100

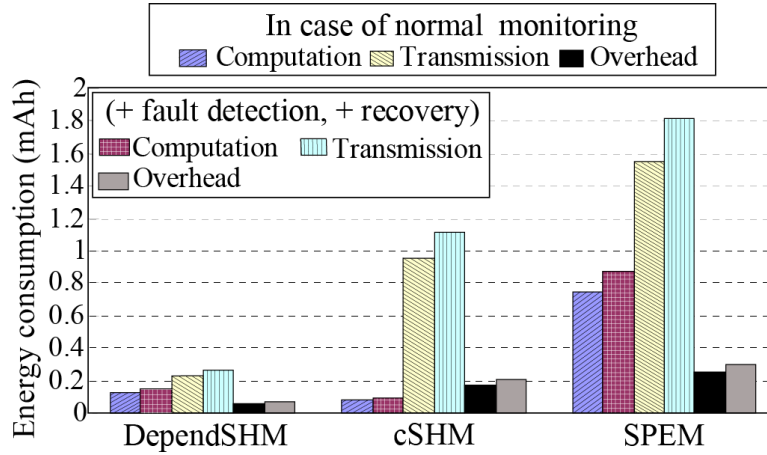


Fig. 9. Energy cost in different measures analyzed by five rounds of monitoring.

locations, which cover up to 450 meters of the structure.  $\Phi$  is extracted, based on sensor collected signals in DependSHM. We can see the impact on the health status, in which the actual mode shape is distorted under the sensor faults, which is successfully recovered by the corresponding sensor signals' reconstruction. This implies that, if there is no appropriate faulty signal detection and tolerance methods, having successful monitoring operations will be difficult to achieve. Thus, a WSN-based SHM system without having such methods will not be dependable. More results and analysis of the *performance of WSN-based SHM system dependability* in different schemes can be found in Appendix F.

**Energy Cost.** We next observe the energy cost in the first five rounds of monitoring seen in Fig. 9 for DependSHM, cSHM, and SPEM schemes. We consider two cases: the amount of energy cost in the case of normal monitoring operation when there is no fault injection in the WSN and the WSN needs to provide health monitoring; the amount of energy cost in cases of monitoring operations when there are the fault injection and recovery from the sensor faults through the signal reconstruction. We calculate the energy cost for computation, transmission, and overhead under both localized and centralized data processing. We did not consider the energy cost for measurement, as we consider the same amount of energy cost for the measurement in all the schemes. We can see that the amount of energy cost for communication in SPEM and cSHM is very large compared to DependSHM. The amount of energy cost in cSHM is seen to be around 60% more than that of DependSHM, while it is 90% more than that of DependSHM.

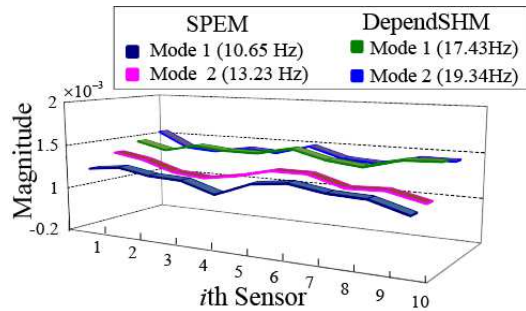


Fig. 10. Global mode shape ( $\Phi$ ) curvatures based on each sensor individual frequencies in SPEM and DependSHM.

## B. WSN Prototype System Implementation

1) *Methodology and Wireless Sensor Platform:* We validate our scheme by implementing a proof-of-concept system using the TinyOS on Imote2 platforms [47]. Our main objective is to verify i) the dependability and ii) the energy-efficiency of the system. We target the accuracy or successful  $\Phi$  identification, because it can provide us with the answer, whether or not a WSN-based SHM system is dependable in terms of various sensor faults. We employ 10 integrated Imote2s called SHM motes on a test structure (refer to Fig. G1 and Appendix G1 for more detail); an additional Imote2 is located 15 meters away as the BS mote, and a PC as a command center for the BS mote and data visualization. The test structure has 10 floors; at each floor, a mote is deployed to monitor the structure's horizontal accelerations. In the experiment,  $R_{min}$  is adjusted by the diameter of the structure, which is adjusted by estimating the height of the test structure and each floor.

*Fault Injection.* To produce a sizable vibration response of the test structure, we collected the original data by vertically exciting the test structure using a magnetic shaker. We inject the sensor faults in two cases: i) debonding fault between the 5th sensor and the structure; ii) precision degradation fault during acceleration signal capturing by the 10th sensor. The sensors are expected to work properly but exhibit faulty acceleration measurements or decisions.

2) *Experiment Results:* In the first set of experiments, we compute mode shapes ( $\Phi$ ) in the base-line structural system, when there are no damage events and no sensor faults. These are computed by sensor initial identified natural frequency (presented in detail in Appendix G2). Fig. 10 demonstrates two mode shapes of the structure, captured by using the identified frequencies in both SPEM and DependSHM schemes. The results compare the exact mode shapes obtained

by centralized WSN where the motes transmit their measured signals to the BS. On the other hand, in DependSHM, the final mode shapes provided by the each mote are combined at the BS. The errors between two processes are analyzed. The accuracy of mode shapes identification in SPEM is at least 13% lower than the accuracy in DependSHM. It is found that DependSHM has around 16% better accuracy than that of SPEM under topology 2 (as shown in Fig. 3). The global mode shape computed at the BS assembles all of the sensors' final results. Note that in the mode shape assembling, mode shapes from different motes correspond to the slight difference in the set of natural frequencies.

Fig. 11a shows experimental fault detection results. Remarkable changes in signals of the 5th and 10th sensors and some of their neighboring nodes are detected. The MII changes in both of the sensor fault cases can be seen in Fig. 11b. Some of the neighboring nodes, e.g., 4th, 6th, 9th, and so on have also provided an extent of change in their MII. This is because their signals have also been partially affected by the fault injection. The corrupted/faulty signals of a sensor (e.g., the 5th) are reconstructed (details performance analysis on the signal reconstruction can be found in Appendix G3).

The energy cost analysis of the experimental WSN is provided in Appendices G4. We find that, in the case of faulty sensor detection and signal reconstruction, DependSHM consumes a small amount of energy in computation with a slight overhead, which is 5% to 8% of the total energy cost in each round. Meanwhile, it saves a significant amount of energy for communication (which is at least three times when compared to its counterparts).

**Dependability verification: Identification of what exactly happened in the structure.** When computing the mode shapes, we should notice that the signals of the faulty sensors contribute to the global mode shape computation. Thus, the mode shape values corresponding to the failed sensors are changed drastically. If there is a missing sensor, the mode shape result is affected. Considering faulty signals, missing signals (in the case of sensor missing), or irregular signals (may be due to the damage), the mode shapes will be affected. But we need to know exactly what happened in a WSN-based SHM system so that we can realize whether WSN-based SHM is dependable or not.

Now, we identify exactly what happened in the structure. Recall that there is a possibility of both sensor fault and damage occurrence at the same location. If there is a change in the signals with a single sensor only, the sensor may be faulty. If the change is present with multiple

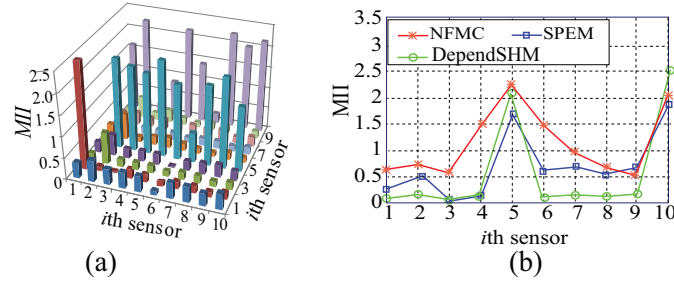


Fig. 11. (a) Distributed sensor fault detection (5th sensor and 10th sensor are faulty); (b) MII achieved by the two detection methods under the sensor faults.

sensors, there is possibly damage. However, if there is damage, it cannot be identified before the faulty sensor detection. As the fault injected, the 5th sensor should be faulty at a different time. As shown in Fig. 12a, during the SHM operation, at  $T_d=5$ , the 5th sensor is detected as faulty, and at  $T_d=20$ , the 10th sensor is detected as faulty. The changes in the mode shape are computed at those time intervals.

We inject structural physical damage through removing the plates on the 5th and 10th floors, since sensors located at these floors are faulty. They are not able to provide appropriate damaged information. We can see in Fig. 12b that the neighbors (4th, 6th, and 9th) are able to detect an extent of changes (i.e., the presence of a damage) in the structure, where the slightly affected mode shapes clearly appeared.

Under the same experiment and excitation setting, we further conduct experiments in which a faulty sensor signal is reconstructed by using our algorithm. The mode shape's curvature is recovered significantly at the 10th sensor location, as shown in Fig. 12c. This means that there is possibly a damage, since the mode shape is still slightly affected. However, the changes in mode shape at the 5th sensor still remains and is slightly recovered at the neighbors. It provides the correctness of DependSHM and the dependability in WSN-based monitoring. If there was no recovery solution, the damage would not be identified and the changes at the sensor near the faulty sensor would not be discovered, which is the exact opposite of the 5th sensor cases.

Further proof of the damage detection can be seen in Fig. 12d. Here, we replace the plate on the same floors. When the sensors wakeup and start monitoring, that time  $T_d=42$ . In Fig. 12d, no remarkable change appears at the 10th sensor location. Distorted mode shape information at the neighbor locations is completely recovered. We can say that no MII appeared. In contrast, at the

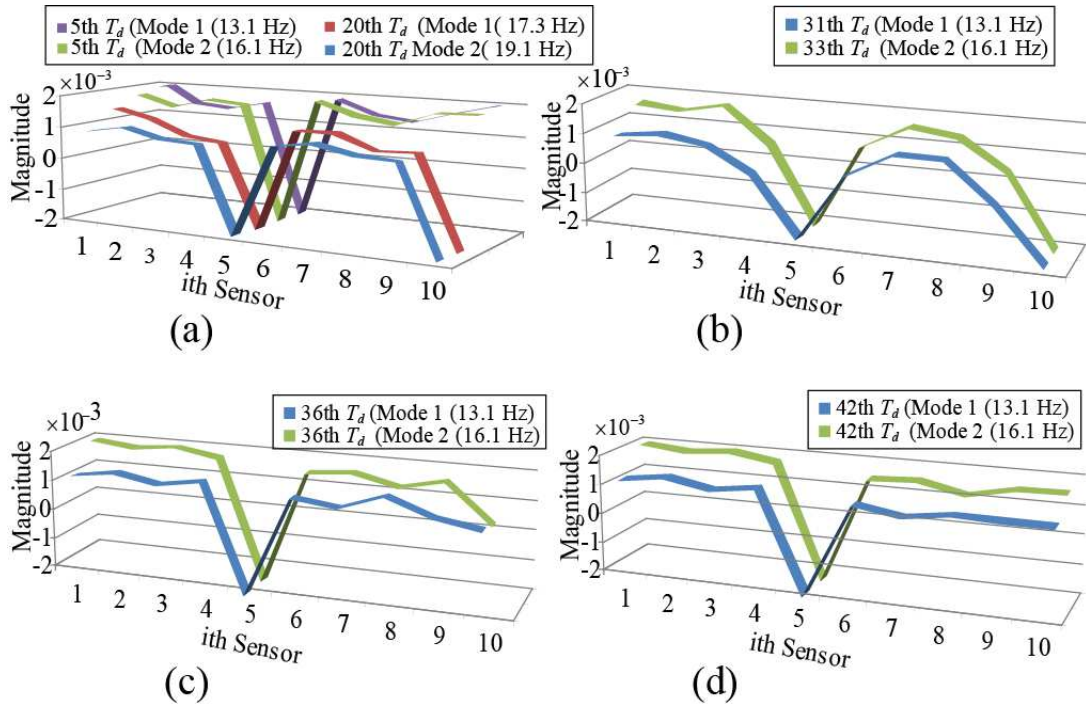


Fig. 12. The mode shape's curvature under sensor faults and recovery from the faults.

5th location, mode shape remains unrecovered but there is no remarkable mode shape curvatures at the neighbor sensors' location. It proves that the 5th sensor is surely faulty, while there was damage at the 10th sensor at this period of monitoring (but which is not clearly detected as the prior damage).

Through an analysis, the quality of the faulty sensor signal reconstruction is about 92% compared to the base-line results under the fault-free condition (as shown in Fig. 10, and TABLE G1 in Appendix G). The inference can be drawn from the above analysis that, in the presence of sensor faults, a damage can be successfully detected in DependSHM.

## VIII. CONCLUSION

In this paper, we proposed a dependable WSN-based SHM scheme, DependSHM, by making the best use of resource-constrained WSNs for SHM and incorporating requirements of both engineering and computer science domains. DependSHM includes two complementary algorithms for sensor fault detection and faulty sensor's signal reconstruction. It is able to provide the quality of SHM in the presence of sensor faults automatically, which does not need any

network maintenance for the fault detection and recovery, and does not consume significant WSN resources for the recovery. In the future, we plan to study decentralized computing architectures in WSNs, which can be integrated by the computing system issues and structural engineering system techniques in conjunction. Such an architecture is highly expected to reduce data traffic for data-intensive SHM and energy cost in WSNs.

#### ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China under Grant Nos. 61272151, 61472451, 61402543, 61202468 and in part by ISTCP grant 2013DFB10070, in part by China postdoctoral research fund, in part by NSF under grants ECCS 1231461, ECCS 1128209, CNS 1138963, CNS 1065444, and CCF 1028167, and in part by HKRGC under GRF grant PolyU5106/11E and HK PolyU Niche Area Fund RGC under grant N\_PolyU519/12.

#### REFERENCES

- [1] J. Gao, J. Li, Z. Cai, and H. Gao, "Composite event coverage in wireless sensor networks with heterogeneous sensors," in *Proc. of IEEE INFOCOM'15*, pp. 1–9.
- [2] M. Z. A. Bhuiyan, G. Wang, J. Cao, and J. Wu, "Energy and bandwidth-efficient wireless sensor networks for monitoring high-frequency events," in *Proc. of IEEE SECON'13*, pp. 194–202.
- [3] M. Z. A. Bhuiyan, G. Wang, J. Wu, T. Wang, and X. Liu, "Resource-efficient vibration data collection in cyber-physical systems," in *the 15th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2015)*, 2015, pp. 1–8.
- [4] G. Wang, M. Z. A. Bhuiyan, J. Cao, and J. Wu, "Detecting movements of a target using face tracking in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 4, pp. 939–949, 2014.
- [5] M. Z. A. Bhuiyan, G. Wang, and A. V. Vasilakos, "Local area prediction-based mobile target tracking in wireless sensor networks," *IEEE Transactions on Computers*, 2014, <http://doi.ieeecomputersociety.org/10.1109/TC.2014.2346209>.
- [6] M. Z. A. Bhuiyan, G. Wang, J. Cao, and J. Wu, "Local monitoring and maintenance for operational wireless sensor networks," in *Proceedings of the 11th IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA'13)*, 2013, pp. 837–844.
- [7] G. Wang, M. Z. A. Bhuiyan, and Z. Li, "Two-level cooperative and energy-efficient tracking algorithm in wireless sensor networks," *Wiley's Concurrency and Computation: Practice & Experience*, vol. 22, no. 4, pp. 518–537, 2010.
- [8] C. Li, M. Z. A. Bhuiyan, and Y. Liu, "Impact of pursuer speed on the performance of distributed surveillance sensor network," in *Proceedings of the 2010 Second International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC '10)*, 2010, pp. 121–124.
- [9] M. Z. A. Bhuiyan, J. Cao, G. Wang, and X. Liu, "Energy-efficient and fault-tolerant structural health monitoring in wireless sensor networks," in *Proc. IEEE SRDS'12*, 2012, pp. 310–310.
- [10] X. Liu, J. Cao, W. Song, P. Guo, and Z. He, "Distributed sensing for high-quality structural health monitoring using WSNs," *IEEE Transactions on Parallel and Distributed Systems*, doi:10.1109/TPDS.2014.2312911, 2015.



- [11] G. Hackmann, W. Guo, G. Yan, Z. Sun, C. Lu, and S. Dyke, "Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 63–72, 2014.
- [12] A. Jindal and M. Liu, "Networked computing in wireless sensor networks for structural health monitoring," *IEEE/ACM Transactions on Networking*, vol. 20, no. 4, pp. 1203–1216, 2012.
- [13] M. Z. A. Bhuiyan, G. Wang, J. Cao, and J. Wu, "Sensor placement with multiple objectives for structural health monitoring," *ACM Transactions on Sensor Networks*, vol. 10, no. 4, pp. 1–45, 2014.
- [14] X. Weng, H.-H. Zhu, J. Chen, D. Liang, B. Shi, and C.-C. Zhang, "Experimental investigation of pavement behavior after embankment widening using a fiber optic sensor network," *Structural Health Monitoring*, vol. 14, no. 1, pp. 46–56, 2015.
- [15] R. Melnik and B. Sowinski, "The selection procedure of diagnostic indicator of suspension fault modes for the rail vehicles monitoring system," in *Proc. of the EWSHM*, 2014, pp. 1–9.
- [16] B. Li, D. Wang, F. Wang, and Y. Q. Ni, "High quality sensor placement for SHM systems: Refocusing on application demands," in *Proc. of IEEE INFOCOM*, 2010, pp. 1–9.
- [17] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized fault-tolerant event boundary detection in sensor networks," in *Proc. of the IEEE INFOCOM*, 2005, pp. 1–9.
- [18] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 1, pp. 98–110, 2015.
- [19] Z. Yuan, H. Xue, Y. Cao, and X. Chang, "Exploiting optimal threshold for decision fusion in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2014, pp. 1–7, 2014.
- [20] T. Clouqueur, K. K. Saluja, and P. Ramanathan, "Fault tolerance in collaborative sensor networks for target detection," *ACM Transactions on Computers*, vol. 53, no. 3, pp. 320–333, 2004.
- [21] T. Wang, L. Chang, D. Duh, and J. Wu, "Fault-tolerant decision fusion via collaborative sensor fault detection in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 756–768, 2008.
- [22] P. Schaffer and I. Vajda, "CORA: Correlation-based resilient aggregation in sensor networks," *Ad Hoc Network*, vol. 7, no. 6, pp. 1035–1050, 2009.
- [23] D. N. Mitiku and M. E. Getachew, "Data processing algorithms in wireless sensor networks for structural health monitoring," Master Thesis, Swedish Institute of Computer Science, Royal Institute of Technology, January, 2012.
- [24] S.-H. Sima, B. F. Spencer, H. Jo, and J. F. Carbonell-Mrquez, "Decentralized random decrement technique for efficient data aggregation and system identification in wireless smart sensor networks," *Probabilistic Engineering Mechanics*, vol. 26, no. 1, pp. 81–91, 2011.
- [25] S. G. Vadlamudi and P. P. Chakrabarti, "Robustness analysis of embedded control systems with respect to signal perturbations: Finding minimal counterexamples using fault injection," *IEEE Transactions on Dependable and Secure Computing*, vol. 11, no. 1, pp. 45–58, 2014.
- [26] Structural health monitoring for guangzhou new tv tower (gntvt) using sensor networks. [Online]. Available: <http://www.cse.polyu.edu.hk/benchmark/>
- [27] TinyOS documentation. [Online]. Available: [http://docs.tinyos.net/tinywiki/index.php/T2\\_on\\_I mote2](http://docs.tinyos.net/tinywiki/index.php/T2_on_I mote2)
- [28] M. Z. A. Bhuiyan, G. Wang, J. Cao, J. Wu, and X. Liu, "Localized decision making in wireless sensor networks for online structural health monitoring," Central South University, Changsha, China, Tech. Rep. TR-SISE-02-Decision.pdf, 2013.
- [29] L. Zhoua, S. Wu, and J. N. Yang, "Experimental study of an adaptive extended kalman filter for structural damage identification," *Journal of Infrastructure Systems*, vol. 14, no. 1, pp. 42–51, 2014.

- [30] M. Z. A. Bhuiyan and G. Wang, "Reliable shortest paths in wireless sensor networks: Refocusing on link failure scenarios from applications," in *Proceedings of the 20th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2014)*, 2014, pp. 1–9.
- [31] M. Z. A. Bhuiyan, J. Cao, and G. Wang, "Backup sensor placement with guaranteed fault tolerance for structural health monitoring," in *Proceeding of the Sixth Beijing-Hong Kong International Doctoral Forum (IDF'11)*, 2011, pp. 1–4.
- [32] M. Z. A. Bhuiyan, G. Wang, J. Wu, X. Xiaofei, , and X. Liu, "Application-oriented sensornetwork architecture for dependable structural health monitoring," in *the 21st IEEE Pacific Rim International Symposium on Dependable Computing (IEEE PRDC 2015)*, 2015, pp. 1–8.
- [33] X. Liu, J. Cao, M. Z. A. Bhuiyan, S. Lai, H. Wu, and G. Wang, "Fault tolerant WSN-based structural health monitoring," in *Proc. of IEEE/IFIP DSN'11*, 2011, pp. 37–48.
- [34] R. Isermann, *Fault-Diagnosis Applications: Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems*. Heidelberg, Germany: Springer, 2011.
- [35] Y. Bao, H. Li, X. Sun, Y. Yu, , and J. Ou, "Compressive samplingbased data loss recovery for wireless sensor networks used in civil structural health monitoring," *Structural Health Monitoring*, vol. 12, no. 1, pp. 78–95, 2013.
- [36] X. Cai and F. Wu, "Multiobjective fault detection and isolation for flexible air-breathing hypersonic vehicle," *Journal of Systems Engineering and Electronics*, vol. 22, no. 1, pp. 52–62, 2011.
- [37] A. Gandomi, M. Sahab, A. Rahaei, and M. S. Gorji, "Development in mode shape-based structural fault identification technique," *World Applied Sciences Journal*, vol. 5, no. 1, pp. 29–38, 2008.
- [38] A. Goodloe and L. Pike, "Monitoring distributed real-time systems: A survey and future directions," National Aeronautics and Space Administration (NASA), Hampton, VA, USA, Tech. Rep. NASA/CR2010-216724, 2010.
- [39] M. Z. A. Bhuiyan, G. Wang, J. Cao, and J. Wu, "Deploying wireless sensor networks with fault tolerance for structural health monitoring," *IEEE Transactions on Computers*, vol. 64, no. 2, pp. 382–395, 2015.
- [40] M. Z. A. Bhuiyan, J. Cao, and G. Wang, "Deploying wireless sensor networks with fault tolerance for structural health monitoring," in *Proceedings of the IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS'12)*, 2012, pp. 194–202.
- [41] B. F. Spencer, "Imote2 for structural health monitoring: User's guide," University of Illinois at Urbana-Champaign, The Newmark Structural Engineering Laboratory, USA, Tech. Rep., 2011.
- [42] T. Nagayama, B. F. Spencer, G. Agha, and K. Mechitov, "Model-based data aggregation for structural monitoring employing smart sensors," in *Proc. of INSS*, 2006, pp. 1–8.
- [43] X. Liu, J. Cao, S. Lai, C. Yang, H. Wu, and Y. Xu, "Energy efficient clustering for WSN-based structural health monitoring," in *Proc. of IEEE INFOCOM*, 2011, pp. 2768–2776.
- [44] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME-Journal of Fluids Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [45] M. I. Ribeiro, "Kalman and extended kalman filters: Concept, derivation and properties," Institute for Systems and Robotics, Portugal, Tech. Rep. 62, 2004.
- [46] ISHMP Toolsuite. [Online]. Available: <http://shm.cs.uiuc.edu/>
- [47] Imote2 Hardware Reference Manual, Crossbow Technology Inc. Sep. 2007, PN: 7430-0409-01.
- [48] S. Olariu and I. Stojmenovic, "Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting," in *Proc of IEEE INFOCOM*, 2006, pp. 1–19.

- [49] V. Gutnik and A. Chandrakasan, “Embedded power supply for low-power DSP,” *IEEE Transactions on VLSI System*, vol. 12, no. 4, pp. 425–434, 1997.
- [50] A. Pandey, M. Biswas, , and M. Samman, “Damage detection from changes in curvature mode shapes,” *Journal of Sound and Vibration*, vol. 145, no. 2, p. 1991, 321-332.
- [51] F. Perez-Cruz, “Kullback-leibler divergence estimation of continuous distributions,” in *Proc. of IEEE ISIT*, 2008, pp. 1666–1670.

## APPENDIX A

### EXTENDED ENERGY COST MODEL ( $\text{COST}(e_i)$ )

One of our important objectives is to minimize the energy cost of the network regarding various aspects, including the sensor fault detection and recovery, damage event detection. Let  $\text{cost}(e_i)$  denote the total energy cost of sensor  $i$ , including measurement, computation, transmission, and overhead. The sensor  $i$  has discrete power level that it can adjust it in ranges from  $R_{\min}$  to  $R_{\max}$ . In the beginning, sensor  $i$  adopts its minimum power level and then  $i$  may dynamically increase it.

We describe here how energy consumed in transmitting a packet. The maximum energy cost of  $i$  depends on the routing protocol used by the data collection application.

Consider a shortest path routing model [13], [16]; there is a path from sensor  $i$  to neighboring sensor node or the BS  $j$ :  $q = z_0, z_1 \dots z_k$ . Sensor  $i$  propagates the data to them. We can find the  $i$ th hop sensor on each path and calculate the amount of traffic that passes along on the paths within each round of monitoring data collection ( $T_d$ ,  $d=1,2,\dots, n$ ). Then, the  $\text{cost}(e_i)$  can be decomposed into the following four parts:

$$\text{cost}(e_i) = e_T + e_{\text{comp}} + e_{\text{samp}} + e_{\text{oh}} \quad (\text{A23})$$

We describe these terms in the following:

- $e_T$  is the total energy cost for data transmission in a round of data transmission over a link between a transmitter and a receiver, where sensor  $i$  uses its power level from a minimum to a maximum, but not beyond the maximum power. We use a standard energy cost model for calculating the packet transmission cost [48].
- $e_{\text{comp}}$  is the energy cost for processing data locally, e.g., computing equation (6) in Section 4.2. If a sensor is allowed to transmit the raw vibration data to the BS directly,  $e_{\text{comp}}$  would be very low. The cost is mainly due to the onboard processor, such as a micro-controller, DSP chip, or FPGA [49]. These devices consume energy proportional to the number of processing cycles, as well as the maximum processor frequency  $f$ , switching capacitance  $\mu$ , and hardware specific constants  $k$  and  $\beta$ , respectively [49]. We focus on the number of cycles taken for tasks, e.g., equation (6) and the amount of samples taken. The number of cycles required to perform a task on the amount of samples (denoted by  $w$ ) are estimated according to the computational complexity  $O(w)$ , which describes how many basic operations, i.e., averages, additions, multiplications, etc., must be

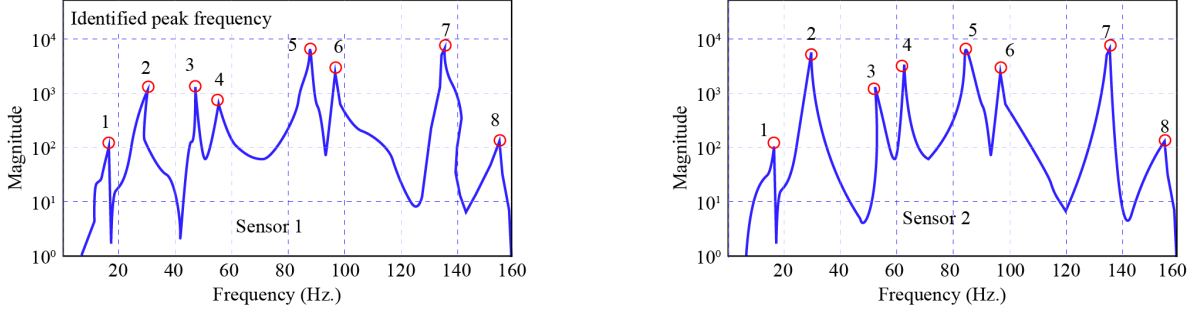


Fig. B1. Based on acquired vibration signal characteristics, measured natural frequencies captured by sensor 1 and sensor 2 under manual input excitation on the structure, respectively. This shows the structural system oscillation (moving back and forth) between its original state and its displaced state, captured by the two in their vicinity.

performed in executing the task. Given these parameters, the computational energy to complete a task can be calculated according to:

$$e_{comp} = O(w) \cdot \mu\left(\frac{f}{k} + \beta\right) \quad (\text{A24})$$

- $e_{samp}$  is the energy required for a sampling cost of  $M$  data points; when sensors capture vibration signals, assuming a maximum 50% overlapping,  $M = (n_a/2 + 1/2) \cdot c_r$ , where  $n_a$  and  $c_r$  are the number of averages mainly for the purpose of noise reduction, that practically ranges from 10 to 20 and cross-correlation factor, respectively [33], [43]. We assume that  $n_a$  and  $c_r$  are set by fixed values on a sensor.
- $e_{oh}$  is any additional overhead for some causes, e.g., fault detection and signal reconstruction, copying data to a local buffer, and network latency.

## APPENDIX B

### METHOD OF EXTRACTING LOCAL MODE SHAPE

In Section 4.2, we have described the state-space model for structural mode shape computation ( $\Phi$ ) at individual sensor. Here, we show a method to local  $\Phi$  extraction and explain benefits of utilizing the extracted  $\Phi$  over  $f$  towards sensor fault detection and tolerance.

**Definition B1** [Natural Frequency]. *Every structure has a tendency to vibrate with much larger amplitude at some frequencies than others. Each such frequency is called a natural frequency denoted by  $f$ .  $f$  is an internal vibration signal characteristic of structure, and is different for different structures (such as from building to bridge, from indoor to outdoor). In other words, it is defined as the number of times that a structural system oscillates (moves back and forth) between its original state and its displaced state when assuming there is no outside interference.*

**Definition B2** [Mode shape]. *When subjected to external forces, the response of a structure is conceptually similar to the response of a vibrating string or structural components such as a metal plate. Upon excitation, the*

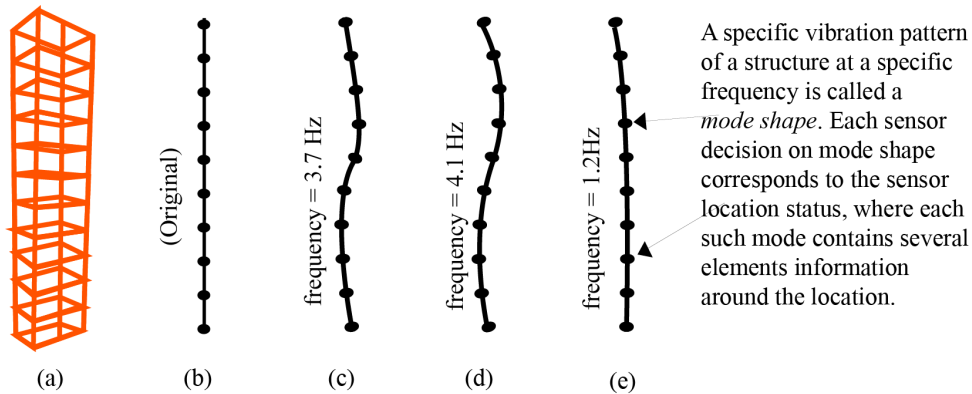


Fig. B2. (a) The finite element model (FEM) of our designed physical infrastructure; (b) its original mode shape; (c)-(e) its three mode shapes: mode 1, mode 2, and mode 3. *FEM* is a computer based numerical model often used for calculating the behavior and strength of structural mechanics, such as vibration and displacement.

*vibrations are a combination of several harmonics (or at a specific frequency of vibrations), known as modes. Each mode deforms the structure into a particular spatio-temporal pattern known as a mode shape, denoted by  $\Phi$ .*

#### A. Local Mode Shape by Each Sensor

As the network modeled in Section 3.1,  $m$  sensors are available for deployment on a structure, and they extract a total of  $p$  mode shapes from the measurement of  $m$  sensors. The corresponding natural frequencies and mode shapes are represented, respectively, as follows:

$$\mathbf{f} = [f^1, f^2, \dots, f^p] \quad (\text{B1})$$

$$\Phi = [\Phi^1 \quad \Phi^2 \quad \dots \quad \Phi^m] = \begin{bmatrix} \phi_1^1 & \phi_2^1 & \dots & \phi_p^1 \\ \phi_1^2 & \phi_2^2 & \dots & \phi_p^2 \\ \dots & \dots & \dots & \dots \\ \phi_1^m & \phi_2^m & \dots & \phi_p^m \end{bmatrix} \quad (\text{B2})$$

where  $f^k$  ( $k = 1, \dots, p$ ) is the  $k^{\text{th}}$  natural frequency,  $\Phi^k$  is the mode shape corresponding to  $f^k$ .  $\phi_i^k$  ( $i = 1, \dots, m$ ) is the value of  $\Phi^k$  at the  $i^{\text{th}}$  sensor. For example, Fig. B1 and Fig. B2 illustrate the first two sensors' natural frequencies and corresponding mode shapes of a physical structure, respectively, which are extracted from measurements of 10 deployed sensors in our prototype system. In the experiment, vertical accelerations at all the given sensors are obtained, and 10% noise is added to all measurements. Under the artificial input excitation, the

measured accelerations (the peak frequency pointed by 1, 2, . . .) at sensors 1 and sensor 2, respectively, refer to Fig. B1 and Fig. B2 (which are obtained by using network topology in Fig. 3).

### B. $\Phi$ over $f$ in Sensor Fault Detection

The difference between  $f$  and  $\Phi$  can be observed by comparing (C1) with (C2) and Fig. C1 with Fig. C2. According to the ACSM theory,  $f$  is not suitable characteristic for damage event detection due to several reasons:

- (i)  $f$  is not a sensitive indicator to damage event, where only severe damage event causes noticeable change on the set of  $f$ ;
- (ii) Due to the global property,  $f$  does not contain any spatial information, and thus localizing damage event is difficult, while damage event detection using  $f$  is computationally inefficient;
- (iii) High frequency modes are more susceptible to additional noise than low frequency modes; iv)  $f$  is susceptible to additional noise [43]; To improve the usability of the  $f$  to detect damage event of small magnitude, high-frequency modes, which are associated with local responses, may be used. However, we argue that adopting  $f$  is not suitable for WSNs considering WSNs' resource limitation;
- (iv) Importantly, a large set of  $f$  is required to be sent to the BS (e.g., SPEM [16], NFMC [33]); damage event detection is greatly affected if a portion of it is lost during transmission.
- (v)  $\Phi$  is directly linked to topology of the structure and  $\Phi$  highly features the dynamics of the structure.

On the other hand, it can be seen from (C2) that  $\Phi$  has elements corresponding to each sensor, thus containing spatial information.  $\Phi$  and its derivatives have been proven to be very sensitive features to detect damage event. It takes into account out-of-frequency-bandwidth modes of the structure, and is also applicable to a complex linear structure. This is why we target on  $\Phi$  computation and observe the impact of sensor faults on  $\Phi$ . However, theoretically,  $\Phi$  is a global parameter of a structure which means that, using sensor deployed on different locations of a structure, the same set of  $\Phi$  may not be obtained. To mitigate this problem, we allow each sensor estimate  $\Phi$  taking measurements about its vicinity (i.e., local structural response).

In this paper, we utilize the mode shape curvature method proposed by civil engineering to identify significant change (i.e., damage event) in the mode shape [50]. The mode shape curvature has high sensitivity to damage event.

## APPENDIX C

### THE REASON OF NEGLECTING DAMPING

In Section 4.2, we have described the space-space model for the structural response measurements by sensors. In (6), we have considered the matrices of mass and stiffness coefficients of the various elements of the structure, but we have neglected the damping.

Damping is neglected in the model (6), considering individual sensor measurement estimation. In any case, (a) faults in the sensors will only be identified if they cause changes in the response of a greater magnitude than the errors in the estimated mode shape, and (b) the modes with low damping, having approximately real modes, will be strongly excited. Thus, undamped mode shapes can be accurately estimated by (6) at each sensor.

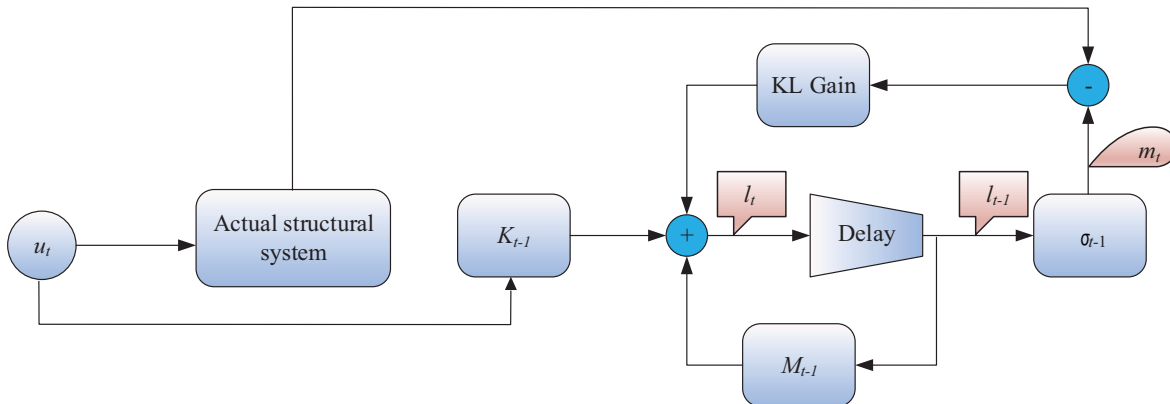


Fig. D. Graphical representation of the state-space equation based Kalman filter that is used in sensor's faulty signal reconstruction.

## APPENDIX D

### THE STATE-SPACE-EQUATION BASED KF FOR SIGNAL ANALYSIS

In Section 6, we proposed the Kalman filter (KF) technique for signal reconstruction. In this Appendix, a graphical representation of the state-space equation based KF is presented in Fig. D. This equation is made with the help of the state-space representation of the structural system, as described in Section 6.1. In Fig. D,  $u_t$  is the structural excitation at a specific frequency at time  $t$ .  $M_t$  and  $K_t$  are transition matrices. The signals  $\sigma_{t-1}$  and  $\sigma_t$  represent the measurement noises, respectively. When measuring the responses of a dynamical structural system by wireless sensors, the actual signals produced by the sensors are contaminated by noises due to internal manufacturing defects, physical interference, or external environmental effects. According to features of the KF, we assume that every measurement from the wireless sensors contains noises; thus, if the noise measurement is zero, the KL collapses. Setting the mean of noise as zero is a common practice:  $E[\sigma_t] = 0$ . Noises are assumed to be independent of each other, and are normally distributed with covariance matrices,  $c_v = [\sigma\sigma^T]$ . The underlying KF information is that KF is a recursive algorithm consisting of a loop, which is passed through for each time instant  $t$ .

## APPENDIX E

### MISSING SENSOR DETECTION METHOD

In Section 5, we presented Algorithm 2 for faulty sensor detection. However, if a sensor is missing or is out of service during the monitoring operation, if the sensor cannot be reached because of communication constraint or failure, or there is an unknown reason, the algorithm cannot guarantee detection of such a sensor node. In order to detect these sensors, we apply a method of Kullback-Leibler divergence (KL) [51] between the measured and

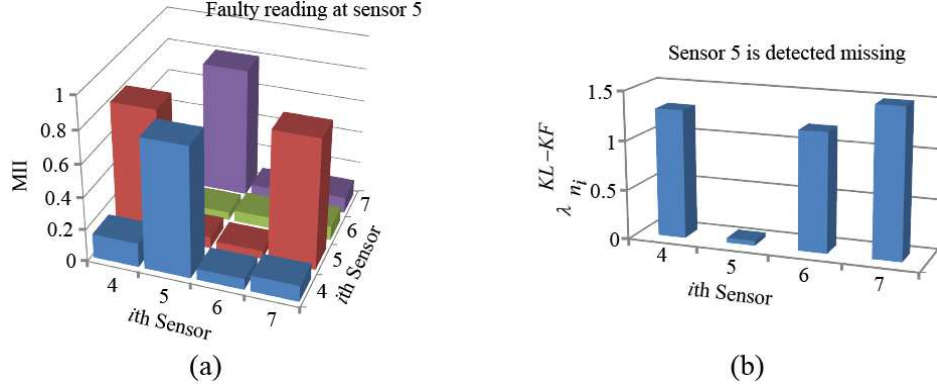


Fig. E. Missing or failed node detection in a WSN-based system: (a) an example of MII change in the sensor signals; (b) the detection result under the KL-KF.

estimated sensor signals and update Kalman filter (KF) with KL, which can be used as a fault indicator  $\lambda_{n_i}^{KL-KF}$  KL-KF (Kullback-Leibler-Kalman-Filter) for such sensors. Note that for a faulty sensor signal reconstruction, we will use the Kalman filter (KL) technique.

The symmetrized form of the KL between the probability distributions of one measured signal ( $y_i$ ) at time  $t$  and with the KF estimated signal  $\hat{y}_i$  is as follows:

$$KL = \frac{1}{2} \sum_i [p_{y_i} - p_{\hat{y}_i}] \log_2 \frac{p_{y_i}}{p_{\hat{y}_i}} \quad (E1)$$

where  $p_{y_i}$  is the probabilities based on the number of points falling into the  $i$ th bin. When the KL between two probability distributions is zero, the signals are identically distributed. The fault indicator is defined as:

$$\lambda_p^{KL-KF} = \frac{1}{p_{\max} - 1} \sum_{p=1}^{p_{\max}} KL_p, 1 < p < p_{\max} \quad (E2)$$

If the faulty sensor  $p$  is not used for the estimation of the  $p_{\max}$  sensor signals, then the KL distance between the measured and estimated signals will be minimal; otherwise, the distance will be higher. This is shown under the network topology in Fig. 3a. It can be seen in Fig. D that without sensor 5, the best estimation is possible, which clearly indicates the sensor is faulty. This method based on KF is able to detect a missing or failed sensor. Also, pure bias faults with the MII method are enhanced further by this KL-KF method. Thus, using KF-KL with the help of Algorithm 2, it can be guaranteed to detect the fault types that produce faulty readings.

We illustrate the justification of sensor fault identification method based on MII (i.e.,  $\omega$ ) through Algorithm 2. In our real experiment, under the manual random excitation and 5th sensor removal, the 5th sensor is detected as faulty. As shown in Fig. E(a), the relative change in MII indicates the sensor 5 as faulty. Actually, the sensor was removed from the location, however, the sensor is detected as missing by  $\lambda_p^{KL-KF}$ , as shown in Fig. E(b). To guarantee a certain redundancy of information in each sensor data set, the initial frequency should be available for



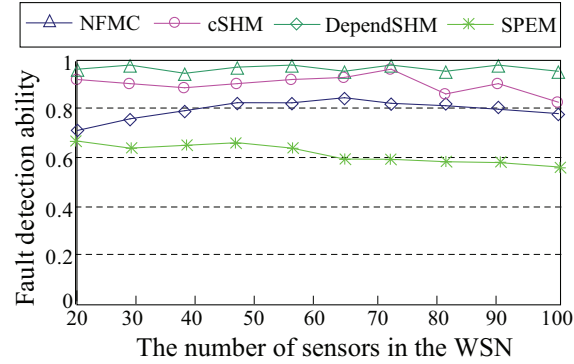


Fig. F1. Dependability verification: fault detection ability of different schemes.

identifying the faulty vibration signal. Therefore, if one of the neighboring nodes is missing, the KL-KF divergence between the measured and estimated sensor signals can be used as a sensor fault indicator.

## APPENDIX F

### MORE RESULTS OF WSN-BASED SHM SYSTEM DEPENDABILITY

In Section 7.1.2, we have partly performed an analysis of the system dependability. In the analysis, we have used a combination of true positive and true negative results in the sensor fault detection accuracy estimation. In this Appendix, we continue the analysis of the performance of the system dependability. We particularly consider the dependability of WSN-based SHM schemes as the ability of fault detection and the ability of structural health event (damage) detection of the schemes.

At first, we discuss the detection ability of different WSN-based schemes. Fig. F1 demonstrates the fault detection ability of DependSHM and other schemes. We can see that the detection ability of DependSHM is much better than that of cSHM, NFMC, and SPEM. NFMC shows higher detection errors than DependSHM, even higher than cSHM. Looking into details of causes, we summarize the following observations under the random fault injection:

- (i) The same pick frequencies cannot be achieved in many neighborhoods or clusters in NFMC;
- (ii) One or more clusters are disconnected from the network, as one or more faulty sensors are isolated based on the natural frequency comparison (although it shows the good ability rate of fault detection in some clusters);
- (iii) The scheme is limited to the frequency matching based fault detection;
- (iv) NFMC fails to detect other types of faults;
- (v) The fault detection ability of SPEM is very low, due to non-faulty reading losses that results in a increased amount of faulty readings;
- (vi) When attempting to recover from the faults, both SPEM and NFMC schemes require a significant amount of energy cost.

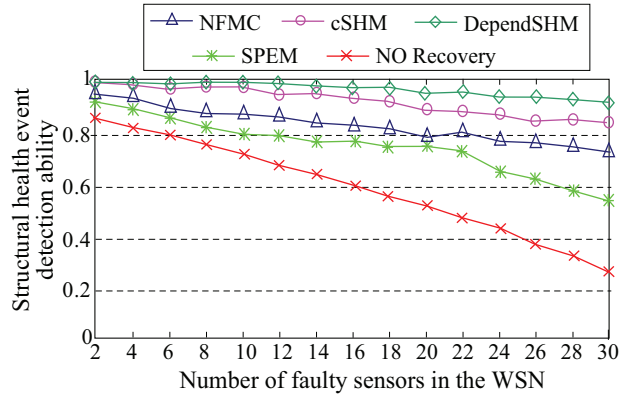


Fig. F2. Dependability verification: structural health event detection ability of different schemes.

We next examine the system dependability in terms of the structural health event detection ability of a system. This can provide us an implication that how much a system can cope with sensor faults and what is the significance of addressing dependability issue in a system. We gather all the false positive and false negative cases appeared in the WSN-based SHM (achieved from a total of 50 simulation runs), and we get an average. Then, we calculate the structural health event detection ability rate as  $1 - (\text{false positive rates} + \text{false negative rates})$ . The results is depicted in Fig. F2. We also take into account the structural health monitoring under *NO* recovery (a preliminary analysis has been done based on these results, as illustrated in Fig. 8). Here, we intend to find evidence that what exactly happens when there is *NO* dependability option (fault detection and recovery) provided.

In Fig. F2, we can see the results, which shows that the structural event detection ability of DependSHM is between 93% and 97.2%, which greatly outperforms others. In SPEM, the detection ability under recovery from sensor faults tolerance algorithm is inferior (between 75% and 92%) among all of the schemes, while it is between 74% and 95% in NFMC and 87% to 95% in cSHM. There can be various reasons that SPEM provides poor detection rate, including i) centralized decision making on the fault detection and tolerance (data losses on the fly is a factor), ii) application-specific sensor deployment, iii) natural frequency matching problem, and so on. In NFMC, the peak natural frequency signals used in the sensor fault detection and recovery, by which the actual mode shape curvature slightly distorted. This lead to a lower MII that results in a lower structural event detection ability. As it can be seen in Fig. F2, the structural event detection ability becomes lower in NFMC and SPEM than in DependSHM and cSHM, as the number of faulty sensor nodes in the WSN increases.

From the results in Fig. F2, the structural event detection ability rate is around 65% in a system with *NO* recovery from sensor faults. It may make us surprised that the monitoring operations in a WSN-based SHM can be often meaningless if there is no dependability option provided. From a deep observation, we have found evidence that faulty sensors can corrupt results of a health event in a structural system without being detected. We have seen that measured signals introduced by some faulty sensors often identify its location as damaged (actually it is undamaged

location). We also have found that some faulty sensor identify its location as undamaged (actually the location is damaged). There are a large number of such wrong diagnoses (false positive and false negative) that lead to a reduced structural event detection ability.

## APPENDIX G

### MORE DETAILS OF THE WSN PROTOTYPE SYSTEM IMPLEMENTATION

#### A. *Extended Detail of the Experimental Setup*

We validate DependSHM by implementing a proof-of-concept system using the TinyOS on Imote2 platforms [47]. Our main objective is to verify i) the dependability and ii) the energy-efficiency of the system. We target the accuracy or successful  $\Phi$  identification, because it can provide us with the answer, whether or not a WSN-based SHM system is dependable in terms of various sensor faults.

The Imote2 (IPR2400) is an advanced wireless sensor platform (off-the-shelf), offering sufficient processing capability and communication resources to locally and continuously monitor vibration characteristics under intensive conditions. Its main board combines a low power PXA271 XScale processor with an 802.15.4 radio (CC2420) and an antenna using 2.4 GHz. The major limitation with it is the energy.

We employ 10 integrated Imote2s called SHM motes on a test structure, as shown in Fig. G1; an additional Imote2 is located 15 meters away as the BS mote, and a PC as a command center for the BS mote and data visualization. The test structure has 10 floors; at each floor, a mote is deployed to monitor the structure's horizontal accelerations. Each mote runs a program (implemented in the nesC language) to process the acceleration data acquired from on-board accelerometers. The BS receives the data packets from the sensors through wireless communication, and relays the data to the PC over a USB cable. The PC commands and sets parameters for the network through BS. Java and Matlab are used to calculate and visualize the whole structural health condition. In the experiment,  $R_{min}$  is adjusted by the diameter of the structure, which is adjusted by estimating the height of the test structure and each floor. Imote2's discrete levels of range are set to use  $R_{min}$  and  $R_{max}$ .

#### B. *Sensor Identified Natural Frequencies*

In Section 7.2.2, we have given experimental mode shapes, estimated based on natural frequencies. In the first set of experiments, we compute the natural frequencies, as shown in TABLE G1. These frequencies are used in creating mode shapes ( $\Phi$ ) in the base-line structural system, when there are no damage events and no sensor faults. Note that such a base-line mode shape should be not fixed but should be dynamic, i.e., a WSN-based SHM system can be enabled to adapt or update its base-line mode shape, taking into account dynamic environments and environmental noise factors. We find that the MII in different frequencies identified at different sensors is low (the result has been shown in Fig. 11).

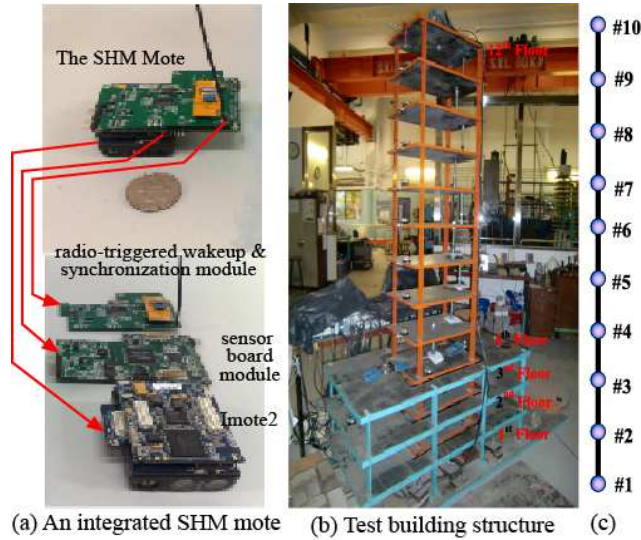


Fig. G1. (a) The SHM mote integrated by Imote2; (b) twelve-story test structure and the placement of 10 SHM motes on it; (c) their deployment.

TABLE G1

IDENTIFIED NATURAL FREQUENCIES BY THE FIRST FIVE SENSORS OF THE EXPERIMENTAL WSN IN SPEM AND DEPENDSHM.

Mode	Frequencies (Hz)									
	Centralized processing (SPEM)					Localized processing (DependSHM)				
	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
1	13.211	12.213	14.131	15.123	13.435	14.134	13.141	15.312	16.856	14.355
2	17.341	14.798	15.112	16.234	15.141	17.741	15.141	17.214	17.852	16.641
3	20.834	21.334	19.134	21.434	19.746	21.341	22.932	21.341	22.344	21.341

### C. Signal Reconstruction at a Faulty Sensor

In Section 7.2.2, we have also provided the sensor fault detection results, where we have found that sensor 5 is faulty. In order to support the results, we hereby can observe the faulty signal reconstruction of the 5th sensor, as shown in Fig. G2. The drift in the measured signal (red line) is corrected by the estimated signal (green lines). We observe the mutual independence under the fault injection at the 5th sensor. In DependSHM, when sensor nodes process data locally, the small value in the MII is achieved, ranging from 2% to 4%, and they are not considered faulty. The MII provides the best value, when there is a remarkable change in the sensor measured signals, i.e., the 5th sensor and 10th sensor are faulty. This reveals that there can be a better accuracy of fault detection in DependSHM in practice, compared to others.

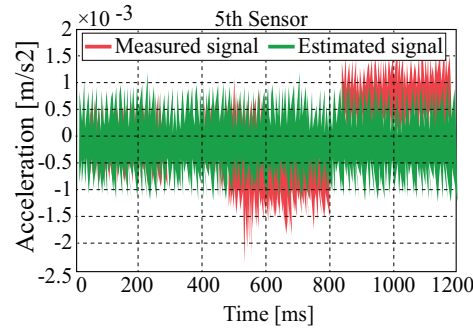


Fig. G2. Signal reconstruction of the 5th sensor (that is detected faulty).

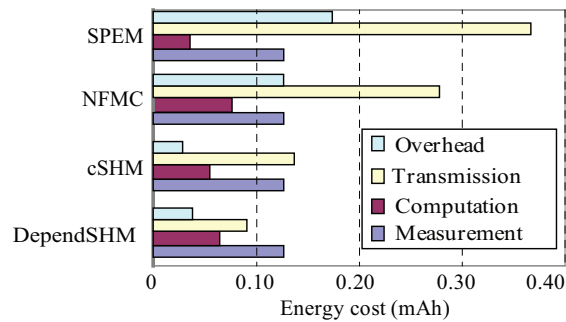


Fig. G3. The performance on the energy cost of the WSN in different schemes.

#### D. Energy Cost ( $cost(e_i)$ )

Due to space limitation, we have not presented the performance of energy cost of the WSN in Section 7.2.2, which we present in this Appendix.

We allow all of the sensors to sleep after each monitoring period to perform power management. The TinyOS 2.0 drivers for the Imote2 supports putting all of the hardware to sleep when it is switched off. This is obvious for a WSN-based SHM system, since a WSN does not always need to run actively in case of specific structural event monitoring. For example, in case of aerospace vehicle monitoring, when it is not flying, the WSN may not need monitoring operations. In another case, the WSN can be scheduled to run periodically or a part of the sensors can be scheduled to wake up periodically and check health event status.  $cost(e_i)$  is calculated by the energy cost for computation, transmission, measurement, and overhead, where the overhead statistics with current cost data is combined. The data sheet can be found in [47].

Fig. G3 shows the energy cost of a round of monitoring,  $T_{d=1}$ . The DependSHM method significantly decreases the energy cost compared to SPEM, from 0.197 mAh to 0.072 mAh. The reason is that the major energy is consumed by the raw signal transmissions to the BS. The actual computation cost in DependSHM is 0.0072 mAh to execute

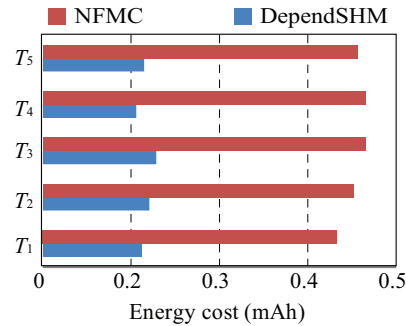


Fig. G4. The performance on the energy cost of the WSN in the first five round of monitoring in DependSHM and NFMFC.

the basic equations and fault detection and signal reconstitution. However, it fully depends on the number of cycles that a sensor CPU requires. It also varies from sensor to sensor based on the tasks a needs to do. A sensor does not need computation for signal reconstruction if there is not fault. In such a case, a sensor can save an average of 0.0027 mAh. More importantly, in DependSys, the computation saves the Imote2 an average of 0.165 mAh during transmission, since it reduces the time that the CC2420 radio is active. The overhead is caused by end-to-end transmission delay and writing/reading data to/from Imote2's memory, since we depend on local processing. In both SPEM and NFMFC methods, transmitting a large amount of raw data in each  $T_d$  (i.e., transmission of natural frequency sets and frequent retransmissions caused by packet losses) increases  $cost(e_i)$ . However, NFMFC achieves slightly lower energy cost for transmission than SPEM.

Further performance analysis of  $cost(e_i)$  in five rounds of monitoring ( $T_d, d = 1, \dots, 5$ ) can be seen in Fig. G4. This shows the actual amount of energy cost required in DependSHM. We can see that DependSHM outperforms NFMFC significantly because of the above causes, cluster maintenance, and network maintenance (e.g., faulty sensor isolation), particularly the set of mode shapes transmitted from the cluster-head to the BS. This is because the final mode shapes of each cluster is transmitted by each cluster-head, while SPEM requires transmission of all natural frequency sets. In our distributed solution, there is no frequent retransmission and the final mode shapes transmitted by each sensor are without sensor fault information. In the case of faulty sensor detection and signal reconstruction, the system consumes a small amount of energy in computation with a slight overhead, which is 5% to 8% of the total energy cost in each round.

In a concluding remark about the results we have found and presented in this paper, our proposed dependable, distributed SHM solution outperforms centralized solution almost in all aspects, including, energy cost of the WSN and offering monitoring system dependability.

