

Efficient Sensor Fault Detection Using Combinatorial Group Testing

Chun Lo, Mingyan Liu

Electrical Engineering & Computer Science
University of Michigan, Ann Arbor
{chunlo, mingyan}@umich.edu

Jerome P. Lynch

Civil and Environmental Engineering
University of Michigan, Ann Arbor
jerlynch@umich.edu

Anna C. Gilbert

Mathematics
University of Michigan, Ann Arbor
annacg@umich.edu

Abstract—This paper introduces a novel use of concepts from combinatorial group testing and Kalman filtering in detecting faulty sensors in a network when faults are relatively rare. By assigning sensors to specific groups and performing Kalman filter-based fault detection over these groups, we can obtain a small binary detection outcome, which can be decoded to reveal the fault state of all sensors in the network. Compared to existing methods, our algorithm achieves similar or better detection accuracy with fewer tests and thus lower computational complexity. We perform extensive numerical analysis using a set of real vibration data collected from the New Carquinez Bridge in California using an 18-sensor network mounted on the bridge.

I. INTRODUCTION

Wireless sensor networks (WSNs) have been successfully used in many applications such as structural health monitoring [1], environmental monitoring [2] and vehicle tracking [3]. With the increasing use of small, low power and low cost sensors, it has also become increasingly critical to ensure the accuracy and integrity of the measured data as low cost sensors are often error prone while the environment in which they are deployed may be harsh. Timely detection of malfunctioning sensors in a system allows it to correct affected sensor readings and arrange for replacement, both of which can prevent further deterioration of the network, and thus should be an essential function to a WSN.

Over the past decade, malfunctioning sensor detection has been studied extensively in many different application contexts. Malfunctioning can be classified into two levels. The first is sensor failure, whereby sensors become irresponsive or ceases to provide data, see e.g., [4]–[6]. The second level of malfunction is sensor faulting, whereby sensors continue to report measurements but the data are intermittently or permanently corrupted. Sensor fault detection is generally more difficult than sensor failure detection because it is typically harder to judge the accuracy of data than it is to determine the absence of data.

Sensor fault detection methods can be further classified into model-based and model-free methods. Model-based fault detection methods are based on the knowledge of the dynamics of the system being monitored. This knowledge can be obtained either from the physical properties of the system (e.g., state-space model) or from learning the parameters of a designated model (e.g., a Markov or autoregressive model). Both Kobayashi et al. [7] and Da et al. [8] proposed centralized detection algorithms which assume a state-space model of the

system is available, and a bank of Kalman filters is used to detect faulty sensors. Both methods assume there is *at most* one faulty sensor at any given time and make use of the remaining sensors as detection references. A more detailed and quantitative comparison is given in Section IV. Li et al. [9] proposed an algorithm that requires fault-free sensors designated *a priori* as reference sensors, and the number of reference sensors is required to be more than the number of uncertain sensors (i.e., those in unknown fault state). Their algorithm constructs analytical relationships between the output of each uncertain sensor and that of all reference sensors, which is then used for detection. Riquebourg et al. [10] captured the sensor dynamics by a Markov chain under a transferable belief framework when the whole system is healthy. Once the model is established, any sensor outputs inconsistent with the model are further analyzed using predefined decision rules. Lo et al. [11] proposed a decentralized algorithm which is able to identify spike faults in addition to detecting general faults. Under this method pairs of sensors cross-validate each other using solely their measurements and an autoregressive with exogenous input model (ARX) trained *a priori*. This method does not require reference sensors or *a priori* knowledge of the system model.

Model-less fault detection methods do not require a dynamical model and usually rely on the assumption that sensors in close proximity observe similar dynamics. As a result, the density of the sensors needs to be high relative to the fluctuation of the signals being monitored. For instance, Ding et al. [12] and Chen et al. [13] suggested similar model-less sensor fault detection methods, where each sensor's output is compared with its neighbors' output. A sensor that deviates significantly from its neighbors is identified as faulty. Koushanfar et al. [14] proposed a cross-validation based fault detection algorithm that focuses on the impact of a particular sensor's measurement on the consistency of the entire network's measurement, on the assumption that an incorrect measurement will degrade the consistency. The algorithm removes one sensor at a time and evaluates how much the consistency of the system improves. The sensor whose removal improves the system most significantly is regarded as faulty and eliminated and the process repeats until the system consistency cannot be improved anymore.

All of the above mentioned fault detection methods require the number of tests at least on the order of the size of the

Method type & Where	Complexity	Notes
Model-based:		
Kobayashi et al. [7]	$\mathcal{O}(N)$	At most one faulty sensor
Da et al. [8]	$\mathcal{O}(N)$	At most one faulty sensor
Lo et al. [11]	$\mathcal{O}(N)$	
Li et al. [9]	$\mathcal{O}(N)$	Reference sensor required
Ricquebourg et al. [10]	$\mathcal{O}(N)$	
Model-less:		
Ding et al. [12]	$\mathcal{O}(mN)$	$m = \#$ of neighbors
Chen et al. [13]	$\mathcal{O}(mN)$	$m = \#$ of neighbors
Koushanfar et al. [14]	$\mathcal{O}(N^2)$	
Blough et al. [16]	$\mathcal{O}(N \log N)$	
Ruiz et al. [6]	$\mathcal{O}(N)$	Focused on sensor failure

TABLE I: Summary of various methods' complexity

network, i.e., $\mathcal{O}(N)$ tests are required, where N is the number of sensors in the network. Some methods even need $\mathcal{O}(mN)$ (where m is the number of neighbors of a sensor) or $\mathcal{O}(N^2)$ tests. A summary of the detection complexity is given in Table I. For applications using an extremely large number of sensors [15], running a fault detection algorithm can involve a large amount of resources and cause significant delay.

We observe that while certain regional effects or disaster events may result in a large number of faulty sensors at the same time, in the absence of systemic problems, during normal operation faults occur randomly and sporadically. This motivates us to seek lower complexity fault detection methods when faults may be rare and sparse.

Toward this end, we introduce a novel use of group testing techniques and Kalman filtering in detecting faulty sensors in a network, which monitors a linear dynamical system, when faults are relatively rare, with the goal of reducing the number of required tests. When the targets of a detection problem are sparse (i.e., they are few), group testing is a potential and sound framework to solve the problem. Usually, the main challenge in applying group testing is to propose a compatible detection method which can accurately evaluate whether a group of *arbitrary* items contains any faulty sensors. There are a few studies that adopted group testing to detect malfunctioning sensors. Specifically, Goodrich and Hirschberg [17] proposed a group testing based algorithm for detecting failure (dead) sensors. This algorithm evaluates a group of sensors by counting the number of responses from the group to a broadcast query. Tošić *et al.* [18] proposed a distributive sensor fault detection algorithm that measures a smooth phenomena (which implies neighboring sensors have similar measurements), while a group test is performed using an unspecified dissimilarity comparison of neighboring sensors' measurements. Our work is significantly different from these two studies because we focus on detecting faulty, though live, sensors deployed on linear dynamical systems, where neighboring sensors may not have similar measurements. Moreover, as the group detection method is problem specific, our main contributions lie in the creative use of two known, though often unrelated, techniques: combinatorial group testing and Kalman filtering on detecting rare faulty sensors.

This paper is organized as follows: Section II reviews the main concepts used in the proposed fault detection algorithm.

The detailed methodology of the detection algorithm is then explained in Section III. The performance of the detection algorithm on bridge vibration monitoring sensors is presented in Section IV and the performance comparison between the proposed method and other existing methods are presented in Section V. Finally, Section VI is the conclusion and future work.

II. PRELIMINARIES

In this section we review two main concepts used in our fault detection algorithm. The first is combinatorial group testing, the goal of which is to identify sparse faulty items with a number of tests less than the total number of items. The second concept is Kalman filtering, which is able to produce optimal state estimation for a linear dynamical system.

Consider a large number of items of which a few are defective, and we wish to identify them. If all items are tested individually to check for defectiveness, the cost can be high (e.g., linear in the total number of items). However, if it is possible to determine the existence of any defective item in a group of items via a single *group test*, then performing a sequence of group tests over different subsets of these items can potentially lead to much fewer number of tests and thus much lower cost. This is the main idea of combinatorial group testing; it was first proposed by Dorfman [19] during World War II for detecting syphilis amongst soldiers.

Consider a length N signal \mathbf{s} which is d sparse: this means \mathbf{s} has at most d non-zero entries that correspond to the defective items and $d \ll N$. As the "true" signal dimension (i.e., d) is smaller than N , it is reasonable to believe signal \mathbf{s} can be acquired with $M < N$ measurements. In group testing paradigm, signal \mathbf{s} is measured M times in the form of $\mathbf{z} = \Phi\mathbf{s}$, where Φ is the measurement matrix of size $M \times N$. The goal is to design Φ such that \mathbf{s} can be reconstructed correctly from \mathbf{z} , that is, we can find the d defective items. The arithmetic is boolean which means multiplication is the logical AND and addition is the logical OR.

We describe this in the context of a network of sensors. Consider a network of N sensors, of which at most d are faulty. Let vector \mathbf{s} represent the *fault state* of the sensors in the network, where $s_i = 0$ if sensor i is normal and $s_i = 1$ if sensor i is faulty. Each row of the matrix Φ , which has $\{0, 1\}$ entries, represents the set of sensors involved in a test, while the number of rows equals the number of tests. Finally, the vector \mathbf{z} represents the result of the group test. Below is a toy example of $\Phi\mathbf{s} = \mathbf{z}$:

Example 1:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

In this example, there are 6 sensors; sensor 2 is faulty. A total of 3 group tests are performed: sensors $\{2, 5, 6\}$ are included in the first test (first row of Φ), and so on. The test result shows

that the first group contains at least one faulty sensor, while the other two groups have none. In a fault detection setting, note that the actual value of \mathbf{s} is unknown, while Φ is known and is used to set up the tests. The vector \mathbf{z} is obtained after the tests, and is then used to reconstruct \mathbf{s} using recovery algorithms. However, it must be noted that group-testing a set of sensors is far more complicated than a simple boolean operator as in the above equation. In other words, to use this group testing framework in practice, we must specify what a ‘‘group test’’ entails, and how to actually obtain values in the \mathbf{z} vector from such a test if the test output is not naturally boolean. This is addressed by a novel use of Kalman filtering detailed next.

The Kalman filter [20] is an algorithm which takes a series of noisy inputs and iteratively calculates a statistically optimal estimate of the state of an underlying linear dynamical system. More specifically, consider a linear dynamical system given by the following state-space model [20]:

$$\mathbf{X}_{k+1} = \mathbf{A}\mathbf{X}_k + \mathbf{B}\mathbf{U}_k + \mathbf{W}_k \quad (1)$$

$$\mathbf{Y}_k = \mathbf{C}\mathbf{X}_k + \mathbf{V}_k \quad (2)$$

The Kalman filter can be separated into two steps, a prediction step and an update step. In the prediction step, the predicted state (of time k based on the value at time $k-1$), $\hat{\mathbf{X}}_{k|k-1}$ and the corresponding uncertainty measure of the prediction, $\mathbf{P}_{k|k-1}$ are calculated:

$$\hat{\mathbf{X}}_{k|k-1} = \mathbf{A}\hat{\mathbf{X}}_{k-1|k-1} + \mathbf{B}\mathbf{U}_k \quad (3)$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^T + \mathbf{R}_w \quad (4)$$

Upon a measurement, \mathbf{Y}_k is observed, and the estimated state and uncertainty measure are updated as follows:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R})^{-1} \quad (5)$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{K}_k(\mathbf{Y}_k - \mathbf{C}\hat{\mathbf{X}}_{k|k-1}) \quad (6)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}_{k|k-1} \quad (7)$$

To summarize, we can use the Kalman filter to estimate the state of a system. In the next section we show how it can be used to perform a group test over a set of sensors.

III. A GROUP TESTING AND FAULT DETECTION METHOD

Consider a network of N sensors monitoring an underlying physical system that can be modeled as a linear dynamical system. Assume any sensor in the network can be faulty and that at most d of them are faulty at any given time. The dynamic evolution of the underlying system as well as observations by the sensors can be expressed similarly as in (2):

$$\mathbf{Y}_k = \mathbf{C}\mathbf{X}_k + \mathbf{V}_k + \mathbf{E}_k \quad (8)$$

where the additional vector \mathbf{E}_k is an unknown sensor fault vector. If there is no fault on sensor i , the i^{th} component of \mathbf{E}_k is zero.

The rest of this section details how to perform group selection (i.e., how to design Φ and how many rows in Φ), fault detection within a group and fault state reconstruction.

A. Group Selection and number of group tests

Recall the fault detection problem represented as $\mathbf{z} = \Phi\mathbf{s}$, where \mathbf{s} represents the fault state of sensors (‘‘1’’ means faulty). One important question in group testing is how to decide the entries of Φ , i.e., which sensors to be included in a particular test. The performance of the detection method largely depends on the form of Φ . This study only focuses on non-adaptive testing, i.e., the entire matrix Φ is decided before performing the first test. One way to ensure the group testing’s ability to identify the faulty sensors is to design a disjunct measurement matrix. A d -disjunct matrix has the property that for any $d+1$ columns, there is always a row with entry 1 in a column and zeros in all the other d columns. For instance, the measurement matrix in Example 1 is 1-disjunct (since any two columns differ in at least one row) but is not 2-disjunct. One simple method to generate a d -disjunct measurement matrix Φ with high probability is to generate each entry randomly such that $\Phi(i, j) = 1$ has probability $1/2$.

One direct question is how many group tests (the number of rows in Φ) is required to gain enough information for correctly revealing the faulty state, \mathbf{s} . For noise-free detections (i.e., the group test result is always correct) and fixed d , if it is reasonable to assume that the faulty sensors in the network are distributed uniformly at random, the necessary and sufficient number of rows in Φ are $\mathcal{O}(d \log(N/d))$ and $\Omega(d \log(N))$, respectively [21]. For the worse case distribution of faults (i.e., adversarial fault model), the necessary and sufficient number of rows in Φ are $\mathcal{O}(\frac{d^2 \log(N)}{\log(d)})$ and $\Omega(d^2 \log(N))$, respectively [21].

B. Fault Detection of a Group

From the previous subsection, we can see that the fault detection method should be able to identify whether an arbitrary group of sensors contains any faulty member. The idea of using Kalman filtering for group testing lies in its ability to estimate the state of the underlying system from the observations of almost arbitrary group of sensors. Specifically, after selecting an arbitrary group of sensors ϕ , we will further split this set into two subgroups A and B , and use the observations from each subset to estimate the state of the underlying system (thus it is required that a group contain at least two sensors). If the estimated states do not agree with each other, the group is regarded as containing at least one faulty sensor and the corresponding entry of \mathbf{z} is set to 1.

Denote the estimated states of the system, computed from observations of the subgroups A and B , as $\hat{\mathbf{X}}_{k|k-1}^A$ and $\hat{\mathbf{X}}_{k|k-1}^B$, respectively. The difference between the two estimated states is given by:

$$\mathbf{e}_k = \hat{\mathbf{X}}_{k|k-1}^A - \hat{\mathbf{X}}_{k|k-1}^B \quad (9)$$

As all states estimated from the Kalman filter are unbiased (i.e., $E[\hat{\mathbf{X}}_{k|k-1}] = \mathbf{X}_k$) [20], the expected difference $E[\mathbf{e}_k] = E[\hat{\mathbf{X}}_{k|k-1}^A] - E[\hat{\mathbf{X}}_{k|k-1}^B] = 0$ if neither A nor B contains any faulty sensor (i.e., the corresponding components in \mathbf{E}_k are zero). Otherwise this expectation is non-zero. Therefore, a threshold can be used to decide whether a group of sensors,

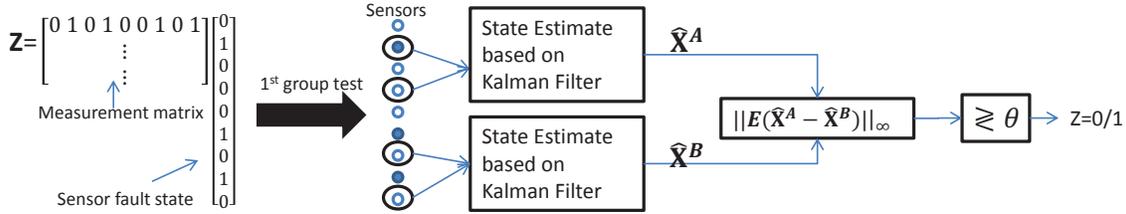


Fig. 1: State diagram of the proposed sensor fault detection method.

ϕ , contain any faulty sensors. If $\|E[\mathbf{e}_k]\|$ is larger than this threshold, the group, ϕ , will be regarded as having at least one faulty sensor and the corresponding entry of \mathbf{z} will be set to 1. Otherwise, the corresponding entry of \mathbf{z} will be set to 0. Fig. 1 gives an overview of this approach. In all the algorithm performs M group tests. When all M tests are done, the algorithm reconstructs the fault state of the sensors from the test results.

It should be mentioned that the system of a group test needs to be detectable in order for Kalman filter to perform properly. Whether a system is detectable depends on the sensors chosen in the group and how the group is partitioned into two sets. If Φ has a group test that cannot satisfy the detectable property, a new group test can always be generated to replace it. The detectability of the system within a group test can be determined by checking whether all of its unobservable modes are stable [20]. This verification process can be done before performing any group test and thus does not affect the number of tests.

Also, notice that the Kalman filter based group detection method may make mistakes, i.e., the result z_i can be equal to 1 even when there is no faulty sensor in test i , and vice versa. Therefore, our study is a noisy group testing problem, instead of the well studied noise-free case. A recent study [22] has been conducted to evaluate the number of tests required for two noisy group testing scenarios: 1) Additive model: the group result, 0, may change to 1 with probability p ; and 2) Dilution model: a faulty sensor may act like a normal sensor (diluted) with probability q in a group test. The sufficient number of tests for the additive model and dilution model, under worst case distribution of faults, are $\mathcal{O}(\frac{d^2 \log(N)}{1-p})$ and $\mathcal{O}(\frac{d^2 \log(N)}{(1-q)^2})$, respectively. However, for group tests that can have both false alarm and miss detection, as in our proposed algorithm, the requirement on the number of tests is still an open question.

C. Fault State Recovery

After the group test results \mathbf{z} are calculated, the sensor fault state is recovered by a straightforward most-likely (ML) decoding. The recovery algorithm evaluates all $\binom{N}{d}$ possible fault states and chooses the one such that the group testing result \mathbf{z} is most likely, i.e., choose ν^* if

$$p(\mathbf{z}|L_\nu^*) > p(\mathbf{z}|L_\nu) \quad \forall \nu \neq \nu^* \quad (10)$$

where L_ν denotes any possible fault state and $\nu \in \{1, 2, \dots, \sum_0^d \binom{N}{d}\}$. In some cases, and in particular in the experiments shown in the next section, the probability measure in Eq. (10) is difficult to obtain and depends on the threshold

used in group testing. This study simply assumes each group test has the same faulty probability. For each possible fault state L_ν , the recovery algorithm calculates the Hamming distance, defined as the number of distinct entries, between the predicted output ΦL_ν and the detection outcome \mathbf{z} . Fault states with smaller Hamming distance is preferred. Among fault states having the same Hamming distance from \mathbf{z} , states with a smaller support are preferred as the probability of a sensor being faulty is $< 1/2$. If this still results in a tie, then the recovery algorithm will choose randomly. The recovery algorithm as implemented here does not scale with the number of sensors N but there are faster algorithms [23] that do scale, (some sublinearly with N). For our parameters used in simulation, the Kalman filter dominates the complexity.

D. Practical Implementation

The method outlined above can be implemented in two ways. The first is as a post processing of data already collected at a cluster head or central location, which is rather straightforward. The second is a form of real-time sequential detection process, where a control center solicits input from a single group of sensors at a time. A single group test is then performed over this group of input. This is followed by soliciting input from the next group, and so on. Note that as long as the fault state of the underlying system remains unchanged, the fault state estimate can be done over different segments of observations over time. In other words, the data provided by each group need not be synchronized and can be generated on demand.

IV. EXPERIMENTAL RESULT

In this section we evaluate our detection method by applying it to a set of real bridge vibration sensing data collected from the New Carquinez Bridge in California. In the next section we further compare our method with a few existing methods. We begin with a list of common sensor fault types considered in our study. We then discuss the nature of our sensing data followed by detection results.

A. Sensor Fault Types

We consider four different fault types: spike, non-linear transduction, mean drift and excessive noise. These are illustrated in Fig. 2 on a sinusoidal signal. More specifically, a spike fault is an impulse superimposed on normal sensor measurements. They are assumed to occur randomly in time with constant or varying magnitudes (consistent with a random signal model). Moreover, the occurrence of these spikes is assumed sparse. A non-linearity fault represents an abnormal discrepancy between

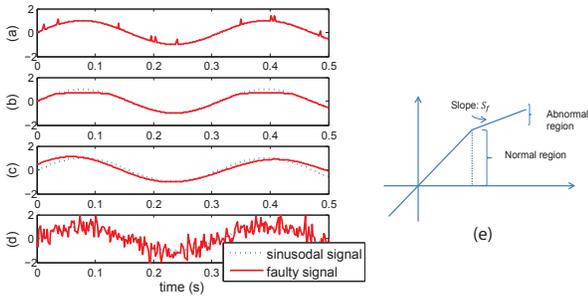


Fig. 2: Illustration of different faults on a sinusoidal signal: (a) Spike, (b) Non-linearity, (c) mean-drift, (d) Excessive noise and (e) non-linear fault model

the sensor input and output. This fault usually happens when the measurement falls outside a certain dynamic range. In this study, a simple non-linear fault model is used as shown in Fig. 2(e): when the measurement is within the normal region, the sensor output reflects the measurement; otherwise the output follows the slope S_f . A mean drift fault preserves the output dynamics but not its mean value. This type of fault generates outputs whose mean drifts away from the true mean of the signal slowly compared to the output dynamics. Finally, excessive noise refers to a large amount of Gaussian noise in the output of a sensor. Compare to regular measurement noise, this fault has much higher amplitude such that the output signal is highly corrupted. Note that only the non-linearity fault is a function of the measured signal but the other fault types are not.

B. Bridge Vibration Data and State Estimation

We evaluate our detection method using bridge vibration data collected by a network of 18 vibration sensors deployed on the New Carquinez Bridge in California. The New Carquinez Bridge is a 1056 meter long suspension bridge which connects Crockett and Vallejo. The locations of these 18 sensors are shown in Fig. 3. They monitor the bridge vibration in the direction perpendicular to the bridge surface. Fig. 4 shows an example of the output of a sensor when vehicles pass through. We took 18 data traces at the beginning of the deployment and performed manual inspection. All tests, including frequency spectrum analysis and the bridge mode-shapes calculation from the data suggested the data traces are correct. Two experiments are conducted to evaluate the performance of the proposed algorithm. The first one is a control experiment, which allows us to evaluate the performance under specific types of faults. In this experiment, different fault types (described in Section IV-A) are superimposed over a random subset of data traces. The resulting data are then used for evaluation purposes. On the other hand, the second experiment is a direct application of the algorithm to detecting potential real faults within the traces.

As shown in Section III, to use Kalman filtering for a group test requires a model representing the dynamics of the system and of the sensors. In practice, it is difficult to obtain the true dynamical model of a bridge. One solution is to learn the model from measured data. One of the commonly used

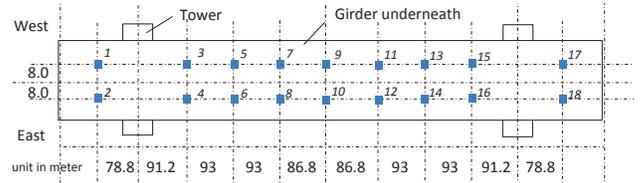


Fig. 3: Plan map of the deployed sensors. Credit: Yilan Zhang

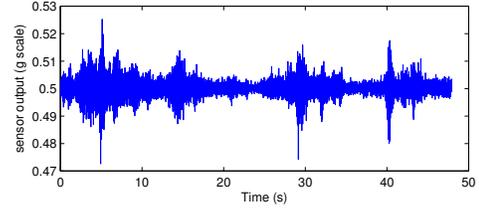


Fig. 4: Vibration measurement of a sensor

methods is the subspace method [24] which utilizes measured output (and input, if available) to calculate model parameters such as matrix \mathbf{A} , (\mathbf{B} if input data is available) and \mathbf{C} in the state-space model (8). Notice that the excitation input to the bridge is in general not available. While input is not necessary for learning the system model by the subspace method, prior study suggests the input can be assumed to be a Gaussian signal for large structures with complex excitations, and that this leads to a better learned system model in terms of output prediction [25]. For our study, we take 50 seconds of vibration data sampled at $200Hz$, from each of the 18 sensors. Half of the vibration data is then used to learn the bridge dynamical model, while the remaining half is used for evaluating the group testing method. The order of the dynamical model is set to 162 (A study of the bridge, [26], indicates that a 162-order state space model is sufficient to capture the bridge dynamics), i.e., the length of the state vector is 162. The excitation inputs are assumed to be 18 degree-of-freedom Gaussian signals and each degree-of-freedom input has zero mean and variance equal to the variance of the output of the sensors. Moreover, a zero-mean Gaussian noise with variance equals to 1% of the variance of sensor measurement was added to each sensor to model the measurement noise. This is in addition to what noise may have already been embedded in the data we collected.

C. Performance of the Group-testing Based Detection Method

For the control experiment, we add different types of faults to the bridge data by randomly selecting up to two sensors (a number α is first chosen uniformly from $\{0, 1, 2\}$, and then α number of intended faulty sensors are chosen uniformly among the 18 sensors). According to our experience, a reasonable assumption on the percentage of faulty sensors is about 10% and thus we set $d = 2$. A total of 100 random runs are conducted (over the choice of the number and identity of faulty sensors, as well as over the random injection of faults and the generation of the Φ matrix) for each number reported in the figures shown in this section. We first examine the performance as a function of the detection threshold used in each group test and the number of tests performed. Fig. 5a shows the

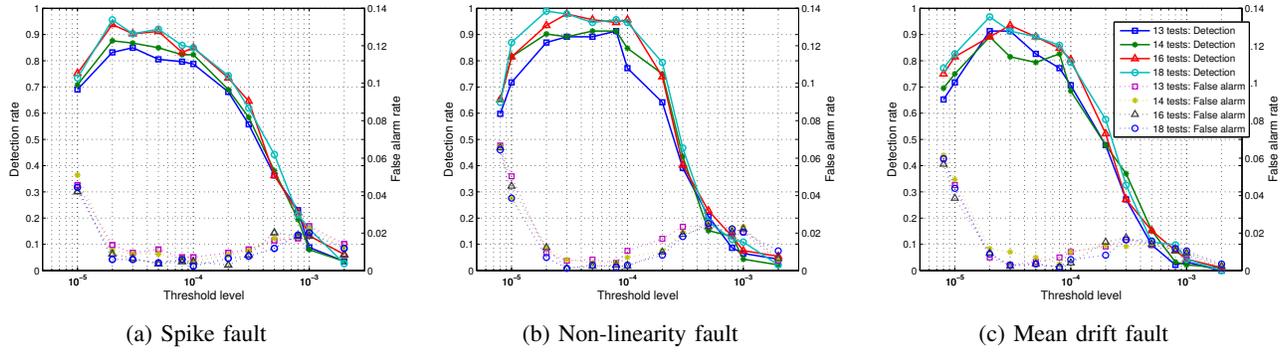


Fig. 5: Detection rate and false alarm on detecting (a) spike fault, (b) non-linearity fault and (c) mean drift fault.

detection rate (the number of detected faulty sensors over the total number of faulty sensors) and false alarm in detecting spike fault under different number of tests and threshold levels. The spike fault was set to appear at 5% of the samples and have mean amplitude equal to the variance of the sensor output, which is common among spike fault in sensors. As can be seen, when the number of tests increases, the detection rate increases while false alarm decreases. When 14 tests are used, the detection rate is above 85% and false alarm is below 1%, with a threshold of 2×10^{-5} . Similarly, when 16 tests are used, the accuracy is over 93% and remains above 80% with a threshold less than 2×10^{-4} .

In all cases we see a fairly wide region of threshold values within which the method enjoys high detection rate ($> 80\%$) and low false alarm ($< 2\%$). This is clearly a desired operating regime for the detection method. In addition, the detection rate first increases with the threshold and then drops slowly with further increase in the threshold. When the threshold increases beyond a certain value (e.g., 3×10^{-4}), the detection rate quickly drops and eventually reaches zero. The false alarm moves in exactly the opposite direction. To explain this phenomenon we note there are two sources of error at play, one due to Kalman filtering and the other due to the recovery algorithm. When the threshold is very low the group test is highly sensitive to small errors in the estimate comparison which could be attributed to measurement noise or inaccuracy in the model rather than actual sensor fault. This high sensitivity leads to high false alarm, but not high detection rate because incorrect group testing results cause the recovery algorithm to err on the fault state. With high false alarm and low detection rate, this is clearly an undesirable threshold region. As the threshold increases the error from recovery decreases, which more than compensates for the decreased sensitivity in the group testing, achieving an overall better tradeoff. When the threshold increases beyond a certain level, the group test becomes insensitive to faults and eventually declares all groups normal, resulting in deteriorating detection rate and false alarm.

The same evaluation is done for the other fault types; these are shown in Figs. 5b and 5c. In Fig. 5b, results of detecting non-linearity fault are shown. The normal dynamic range is set to 80% of the output maximum, with a slope in the abnormal region of 0.3. The result for mean-drift error is presented in

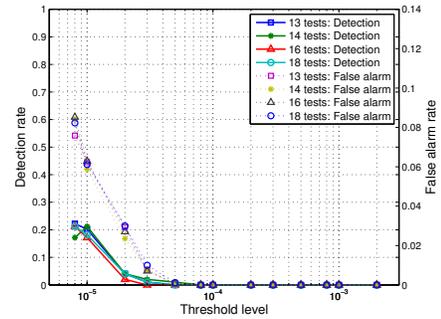


Fig. 6: Detection and false alarm rate on detecting excessive-noise fault.

Fig. 5c. The mean-drift has a maximum frequency of $5Hz$ and a magnitude of 50% of the sensor output variance. All these results show similar behavior to those observed in the spike fault case. Within the preferred threshold range, the detection rate generally exceeds 80% in accuracy while false alarm remains low. Furthermore, the preferred threshold range is smaller when the fault is less pronounced. Finally, the detection performance of the proposed method is tested when the sensor is corrupted by excessive Gaussian noise with zero mean and variance equal to 50% of the variance of sensor output. The result presented in Fig. 6 shows that the proposed method is not recommended for detecting this type of fault. The poor detection performance in this case is due to the fact that Kalman filtering, in computing statistically optimal estimates of the system state, tends to eliminate noise variance existing in the sensor measurement. Consequently, zero-mean excessive noise is sufficiently suppressed in the estimate and does not get reflected in the residual of a group test.

In addition to the control experiment, we also evaluated the algorithm performance on real sensor faults. Several weeks after deployment, sensor 11 appears to start having errors (this is done by manual and visual inspection of its data). As shown in Fig. 7, the output of sensor 11 has obvious spikes beyond normal fluctuation, and possibly has a shift on the mean amplitude and a small mean-drift error as well. It should be noted that this observation is not the absolute “ground truth” but is the closest we can possibly get under the circumstances (the alternative is to take the sensor off the bridge and calibrate

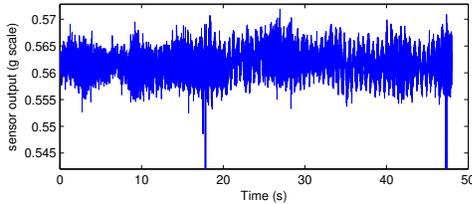


Fig. 7: Abnormal vibration measurement of sensor 11.

it in a lab; even if we could do so the result is only valid if the same type of faults persists in the lab setting). We used our algorithm on the 18 sensors with 6 and 8 tests respectively. Under the same preferred threshold range (between 3×10^{-3} and 1×10^{-4}) as in the control experiment, our algorithm was able to identify the faults in sensor 11, with a detection rate $> 78\%$ ($> 92\%$) and false alarm $< 1.8\%$ ($< 0.7\%$) when using 6 (resp. 8) tests.

V. COMPARISON WITH OTHER METHODS

In this section we compare our group-testing based detection method with two existing methods both in terms of their complexity and in accuracy. As our method is model-based, our comparison focuses on other model-based methods. Among the model-based methods listed in Table I, two of them are directly comparable, which are Kobayashi et al. [7] and Da et al. [8].

Both Kobayashi and Da are based on a bank of Kalman filters. Specifically, with N sensors in the network, N fault detection tests (N Kalman filters) are required to evaluate all sensors in the network. In each test, all sensors but one are involved, i.e., test i uses $N - 1$ sensors and exclude sensor i . A key assumption in this method is that there is only one faulty sensor in the network, thus the test which does not contain the faulty sensor will have different characteristics than the other $N - 1$ tests, and thus the single faulty sensor can be identified.

The difference between these two methods lies in how to compare the test outcomes to determine the difference characteristics with and without the faulty sensor. Under the method by Kobayashi the estimated sensor output from the Kalman filter is compared to the corresponding observed sensor output. The test which does not contain the faulty sensor will have higher consistency result than the other tests. Under the method by Da, there is a reference system which gives a reference system state estimate from N sensors by the Kalman filter. Each test compares the estimated system state (from $N - 1$ sensors) from the Kalman filter to this reference system state. The test that does not contain the faulty sensor is supposed to have lower consistency result because the reference system contains faulty sensor but the test does not.

Fig. 8 shows the detection rate of the three methods under different types of faults, different measurement noises, and with a single faulty sensor, using the same set of bridge data as in the previous section. As we can see, Kobayashi and Da's method achieve similar performance as our proposed method when 8 to 10 tests are used. This result is to be expected when the assumption of no more than one faulty sensor holds, since all the methods are based on Kalman filter. Moreover, because

they use a fixed number of sensors ($N - 1$) in a test (while our method involves different (and fewer) number of sensors in different group tests, ranging from 1 to $N/2$ in a test), their estimation accuracy is generally better because the sensitivity to fault of Kalman filtering is slightly different when different number of sensors are used. As shown in Section II, the the complexity of Kalman filtering largely depends on the size of the system state \mathbf{s} , rather than the number of sensors used in state estimation. *One* detection test of Da's and Kobayashi's algorithms has similar complexity as *one* group detection test of the proposed group-testing based detection method if the sensor network size remains the same. The results in Fig. 8 indicate our proposed method is able to achieve similar, and sometimes better, accuracy when around 8 to 10 tests are used, which is about half of the complexity of Kobayashi's and Da's method (18 tests).

When the system has more than one faulty sensor, the performance of Kobayashi and Da's method deteriorates sharply as all the reference system are contaminated by faulty sensor observations. If the false alarm rate is restricted to a reasonable level (5%), the accuracy of Da's method dropped to about 55% and Kobayashi's method dropped to about 50% for non-linearity fault and to about 20% for spike and mean drift fault (Fig. 9). At the same time, the proposed algorithm maintains over 85% of accuracy for all fault types as shown in the simulation results in section IV.

To summarize, compared to other model-based methods, our proposed method has fewer assumptions on the underlying system and the nature of the faults. It achieves high accuracy with much lower complexity than existing methods. This is particularly relevant for very large sensor networks. Moreover, the proposed method provides a simple trade off between complexity and fault detection accuracy. When the network lacks resources such as energy or communication bandwidth, fewer tests need be done by lowering the target detection accuracy. Furthermore, the above comparison shows that the proposed method is insensitive to measurement noise.

VI. CONCLUSION

This paper presented the first study on the asymptotic fault detection problem where the number of faulty sensors is much smaller than the size of the sensor network. By introducing how to detect whether a random group of sensors contain any faulty sensor and how to form a sensor fault detection problem in a group testing framework, the proposed method successfully reduced the number of detection tests required to identify faulty sensors, thus The reduction in complexity helps to preserve power and communication bandwidth in resources-limited wireless sensor networks. Detailed performance analysis also shows the proposed method is able to achieve comparable, sometimes even higher, accuracy that the other existing methods while reducing the detection complexity. Future works of this study include developing an adaptive compressed sensing algorithm, where the decision on which sensors being tested is based on the previous results.

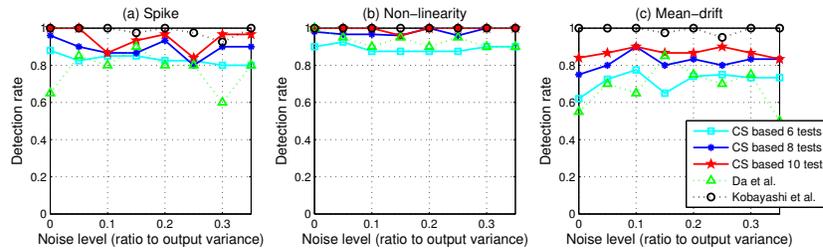


Fig. 8: Detection rate under different measurement noises and fault types with non-adaptive threshold.

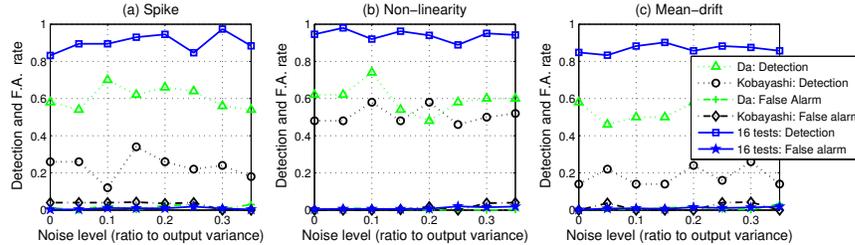


Fig. 9: Detection rate under different measurement noises and fault types with two faulty sensors.

REFERENCES

- [1] J. Lynch, "An overview of wireless structural health monitoring for civil structures," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 345–372, 2007.
- [2] T. Wark, P. Corke, P. Sikka, L. Klingbeil, Y. Guo, C. Crossman, P. Valencia, D. Swain, and G. Bishop-Hurley, "Transforming agriculture through pervasive wireless sensor networks," *Pervasive Computing, IEEE*, vol. 6, no. 2, pp. 50–57, april-june 2007.
- [3] B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden, "Cartel: a distributed mobile sensor computing system," in *Proceedings of the 4th international conference on Embedded networked sensor systems*. ACM, 2006, pp. 125–138.
- [4] N. Ramanathan, K. Chang, R. Kapur, L. Girod, E. Kohler, and D. Estrin, "Sympathy for the sensor network debugger," in *Proceedings of the 3rd international conference on Embedded networked sensor systems*, ser. SenSys '05. New York, NY, USA: ACM, 2005, pp. 255–267.
- [5] J. Staddon, D. Balfanz, and G. Durfee, "Efficient tracing of failed nodes in sensor networks," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, ser. WSNA '02. New York, NY, USA: ACM, 2002, pp. 122–130.
- [6] L. B. Ruiz, I. G. Siqueira, L. B. e. Oliveira, H. C. Wong, J. M. S. Nogueira, and A. A. F. Loureiro, "Fault management in event-driven wireless sensor networks," in *Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, ser. MSWiM '04. New York, USA: ACM, 2004, pp. 149–156.
- [7] T. Kobayashi and D. Simon, "Application of a bank of kalman filters for aircraft engine fault diagnostics," DTIC Document, Tech. Rep., 2003.
- [8] R. Da and C. Lin, "Sensor failure detection with a bank of kalman filters," in *American Control Conference, 1995. Proceedings of the*, vol. 2. IEEE, 1995, pp. 1122–1126.
- [9] Z. Li, B. Koh, and S. Nagarajaiah, "Detecting sensor failure via decoupled error function and inverse input–output model," *Journal of engineering mechanics*, vol. 133, no. 11, pp. 1222–1228, 2007.
- [10] V. Riquebourg, D. Menga, M. Delafosse, B. Marhic, L. Delahoche, and A. Jolly-Desodt, "Sensor failure detection within the tbm framework: A markov chain approach," in *Proceedings of IPMU*, vol. 8, 1991, p. 323.
- [11] C. Lo, J. Lynch, and M. Liu, "Reference-free detection of spike faults in wireless sensor networks," in *Resilient Control Systems (ISRCs), 2011 4th International Symposium on*, aug. 2011, pp. 148–153.
- [12] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized fault-tolerant event boundary detection in sensor networks," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 2, march 2005, pp. 902–913 vol. 2.
- [13] J. Chen, S. Kher, and A. Somani, "Distributed fault detection of wireless sensor networks," in *Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks*. ACM, 2006, pp. 65–72.
- [14] F. Koushanfar, M. Potkonjak, and A. Sangiovanni-Vincentelli, "On-line fault detection of sensor measurements," in *Sensors, 2003. Proceedings of IEEE*, vol. 2. IEEE, 2003, pp. 974–979.
- [15] S. Cho and A. Chandrakasan, "Energy efficient protocols for low duty cycle wireless microsensor networks," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol. 4, 2001, pp. 2041–2044 vol.4.
- [16] D. Blough, G. Sullivan, and G. Masson, "Fault diagnosis for sparsely interconnected multiprocessor systems," in *Fault-Tolerant Computing, 1989. FTCS-19. Digest of Papers., Nineteenth International Symposium on*. IEEE, 1989, pp. 62–69.
- [17] M. Goodrich and D. Hirschberg, "Efficient parallel algorithms for dead sensor diagnosis and multiple access channels," in *Proceedings of the eighteenth annual ACM symposium on Parallelism in algorithms and architectures*. ACM, 2006, pp. 118–127.
- [18] T. Tosic and P. Frossard, "Distributed group testing detection in sensor networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3097–3100.
- [19] R. Dorfman, "The detection of defective members of large populations," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 436–440, 1943.
- [20] P. Maybeck, *Stochastic Models, Estimation, and Control: Vol.: 1*. Academic Press, 1979.
- [21] A. Gilbert, B. Hemenway, A. Rudra, M. Strauss, and M. Wootters, "Recovering simple signals," in *Information Theory and Applications Workshop (ITA), 2012*, feb. 2012, pp. 382–391.
- [22] G. Atia and V. Saligrama, "Boolean compressed sensing and noisy group testing," *Information Theory, IEEE Transactions on*, vol. 58, no. 3, pp. 1880–1901, march 2012.
- [23] H.-B. Chen and F. K. Hwang, "A survey on nonadaptive group testing algorithms through the angle of decoding," *Journal of Combinatorial Optimization*, vol. 15, no. 1, pp. 49–59, 2008.
- [24] T. Katayama, *Subspace methods for system identification*. Springer Verlag, 2005.
- [25] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1951–1968, 1998.
- [26] M. Kurata, J. Kim, J. Lynch, G. v. d. Linden, H. Sedarat, E. Thometz, P. Hipley, and L.-H. Sheng, "Internet-enabled wireless structural monitoring systems: Development and permanent deployment at the new carquinez suspension bridge," *Journal of Structural Engineering*, 2012.