IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 July 2023, accepted 7 September 2023, date of publication 15 September 2023, date of current version 26 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3315745

THEORY

Performance Limitation of Group Testing in Network Failure Detection

FANGYUAN XU[1], SHUN-ICHI AZUMA[2], (Senior Member, IEEE), RYO ARIIZUMI[3], (Member, IEEE), AND TORU ASAI[1], (Member, IEEE)

[1]Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
[2]Graduate School of Informatics, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan
[3]Department of Mechanical Systems Engineering, Faculty of Engineering, Tokyo University of Agriculture and Technology, Koganei-shi, Tokyo 184-8588, Japan

Corresponding author: Shun-Ichi Azuma (sazuma@i.kyoto-u.ac.jp)

This work was supported in part by the Grant-in-Aid for Scientific Research (A) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, under Grant 21H04558; and in part by the China Scholarship Council.

ABSTRACT In a network system, there inevitably be a few connection failures at nodes, such as delay. Once a failure occurs, the network administrator must detect failure sources as soon as possible to maintain communication over the network. Group testing is a method for detecting failure nodes in networks using a small number of measurements, provided that the measurement matrix is constructed appropriately. A promising method for constructing measurement matrices is given by the binary correlation matrices. This study analyzes the performance limitation of group testing based on the binary correlation measurement matrix. We derive the upper and lower bounds of the minimum number of measurements needed for network detection. Moreover, we propose a sufficient condition of network topology, under which the failure vertices in the network can be detected with optimal performance, and we also provide a detection scheme with guaranteed exactness for the network. Numerical example indicates that for the network that satisfies the proposed sufficient condition, the administrator can exactly detect the failure vertices with optimal performance by using our proposed detection scheme.

INDEX TERMS Network, failure detection, group testing, sparse reconstruction, graph theory.

I. INTRODUCTION

Network management refers to the process of configuring, monitoring, and managing the performance of a network. It is one of the most important components of network operation [1]. Monitoring and detecting abnormal network characteristics, such as delay at nodes, is an indispensable research topic of network management [2], [3], [4], [5]. In this paper, we focus on the diagnosis of the nodes with abnormal characteristics in a network.

A typical network system consists of an administrator and nodes, as shown in Fig. 1. It is modeled as an undirected graph, which consists of two sets called vertices, edges, and an incidence relation between them. In a network, the vertices communicate with others over the edges [6]. The network administrator is responsible for network management and maintaining the stability of the network. In practice, it is inevitable to occur some failure nodes in the communication network. Thus, the administrator has to locate the failure sources and subsequently repair them as soon as possible to maintain the quality of communication over the network [7].

A straightforward approach for failure detection is directly measuring the health of individual nodes. However, such direct measurements and monitoring of all nodes are unwieldy due to the high costs of communication and detection [2], [3], [4]. Therefore, it is desirable to avoid employing such brute-force measurements for failure detection in networks.

Recently, several frameworks for failure detection in networks have been proposed [8], [9], [10], [11], and they are based on the idea of *group testing*. Group testing is an anomaly detection approach that divides the objects into several groups and identifies abnormal items on groups, rather than on individual ones [12]. When there are a few

The associate editor coordinating the review of this manuscript and approving it for publication was Anandakumar Haldorai.

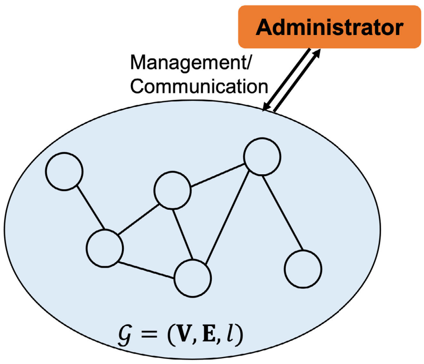


FIGURE 1. Network system.

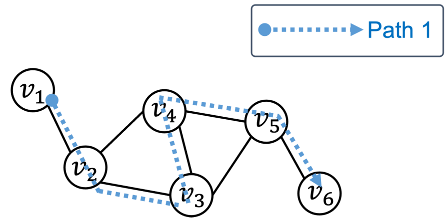


FIGURE 2. Network $\mathcal{G}$.

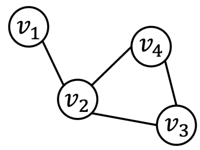


FIGURE 3. Induced subgraph $\mathcal{G}_{\mathbf{T}}$.

abnormal objects, it is expected to detect them with a few tests. Applying group testing to failure detection in networks, then a group corresponds to a simple path that is a sequence of distinct adjacent nodes in the network. The administrator sends test signals, called probes, along the pre-determined paths to measure the sum of faults over the paths, and such measurement is called analog measurement [13]. Then, the administrator estimates the failure nodes based on the measurement results of the probes.

For exact detection, it is important to appropriately choose probes, i.e., to design paths. However, in almost all of the existing results [8], [9], [10], [11], the probes are constructed based on random walks in networks. Although the results show that such random construction can detect failure sources with a high probability, it is never exactly equal to one. In the field of sparse reconstruction, there is a promising method that can be applied to the construction of probes for detection, called binary correlation construction [14], [15], [16], [17], [18]. Compared with the existing method in [8], [9], [10], and [11], the binary correlation construction of probes can guarantee the exactness of the failure detection in the network.

In the interest of minimizing the cost of failure detection, it is important to know the minimum number of measurements needed for failure detection in networks. Therefore, in this study, we analyze the performance limitation of group testing based on the binary correlation construction probes. The contributions of this paper are as follows.

1) Based on the knowledge of sparse reconstruction, we analyze the performance limitation of group testing based on the binary correlation construction, where the performance is evaluated by the number of needed measurements for networks. We derive the upper and lower bounds of the minimum number of measurements needed for failure detection in networks.
2) We consider the network topological constraint on the design of probes and provide a sufficient condition of network topology, under which the failure vertices in the network can be detected with optimal performance.
3) We provide detection schemes with guaranteed exactness for the networks that satisfy the condition in 2).

To the best of our knowledge, there is no result that guarantees the exactness of group testing with analog measurements in network detection.

Our result is expected to provide a solution for the network administrator to configure a network with a low failure detection cost, and further construct a detection scheme with guaranteed exactness for the network.

The rest of the paper is organized as follows: In the rest of this section, we introduce some preliminaries on graph theory and notations. Section II presents the failure detection problem in networks and introduces group testing with analog measurements, then formulates the problem of analyzing the performance limitation of group testing based on the binary correlation measurement matrix. Our main results are provided in Section III. Finally, a numerical example of the application of our results to network detection is shown in Section IV.

Preliminaries on Graph Theory and Notations: We next introduce some graph theory preliminaries [19], [20] and notations that will be used in this paper.

Consider an undirected graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V} = \{v_1, v_2, \ldots, v_n\}$ is the set of vertices and $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$ is the set of edges. The *degree* of a vertex is the number of edges with the vertex as an end-point. Let δ_i denote the degree of the vertex $v_i \in \mathbf{V}$. Then the minimum degree of the vertices in $\mathcal{G}$ is denoted by

$$\delta(\mathcal{G}) = \min_{i \in \{1,2,\ldots,n\}} \delta_i.$$

For example, consider the network $\mathcal{G}$ with six vertices and seven edges in Fig. 2. The degree of the vertex v_3 is $\delta_3 = 3$, and the minimum degree of the vertices in $\mathcal{G}$ is $\delta(\mathcal{G}) = 1$.

A *simple path* is a sequence of distinct adjacent vertices in graph. A *Hamiltonian path* is a path that visits each vertex of the graph exactly once. For example, consider the network in Fig. 2. Path 1 is a simple path, and it is a Hamiltonian path of $\mathcal{G}$.

Next, a class of subgraphs, called *induced subgraph*, is introduced. Consider a set of vertices in $\mathcal{G}$, denoted by

$\mathbf{T} \subseteq \mathbf{V}$. The subgraph of $\mathcal{G}$ induced by $\mathbf{T}$ is the graph that has $\mathbf{T}$ as its set of vertices and contains all the edges of $\mathcal{G}$ that have both endpoints in $\mathbf{T}$, it is denoted by $\mathcal{G}_{\mathbf{T}}$. For example, Fig. 3 shows the subgraph of the network in Fig. 2 induced by $\mathbf{T} = \{v_1, v_2, v_3, v_4\}$.

Finally, we introduce some notations. Let $\mathbf{R}$ be the real number field, and let $\mathbf{Z}_+$ denote the set of positive integers. We use $\|x\|_p$ to represent the ℓ_p-norm of the vector x. The cardinality of the set $\mathbf{P}$ is denoted by $|\mathbf{P}|$. We use $\lfloor x \rfloor$ to denote the floor of a real number x, it is the greatest integer less than or equal to x. Similarly, $\lceil x \rceil$ denotes the ceiling of x, it is the least integer greater than or equal to x.

II. FAILURE DETECTION IN NETWORKS

A. FAILURE-VERTEX DETECTION PROBLEM

Consider a communication network system as exemplified in Fig. 1, which consists of an administrator and nodes. It is modeled as an undirected graph $\mathcal{G} = (\mathbf{V}, \mathbf{E}, l)$, where $\mathbf{V}$ is the set of vertices, $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$ is the set of edges, and $l : \mathbf{V} \to [0, \infty)$ is a function defining vertex labels.

The elements of $\mathbf{V}$ are denoted by $v_1, v_2, \ldots, v_n$, and the label of vertex v_j is given by $l(v_j) \in [0, \infty)$. We use $l(v_j)$ to represent the *failure status* at the vertex v_j $(j = 1, 2, \ldots, n)$. The value $l(v_j)$ represents the delay time in communication at vertex v_j, and thus $l(v_j) \neq 0$ and $l(v_j) = 0$ respectively mean that there is a fault and that there is no fault. In this paper, we refer to a vertex with a nonzero label as a *failure vertex*.

Once a vertex fails, the administrator has to detect it as soon as possible to maintain the quality of communication over the network. In general, failure detection can be easily performed by directly monitoring the status of all vertices. However, the administrator should avoid such brute-force diagnosis. Thus, we expect a method that can detect the failure vertices of a network via a few measurements.

B. GROUP TESTING WITH ANALOG MEASUREMENTS

Group testing [12] is a method for detecting failure vertices with a small number of measurements. In this method, test signals, called probes, are sent by the administrator to measure the failure status on pre-determined paths in the network. This study focuses on the group testing with analog measurements [9], and it is detailed as follows.

Consider a network $\mathcal{G}$ with n vertices. First, the administrator specifies m simple paths in the network and calls them path $1, 2, \ldots, m$. The set of the vertices in path i is denoted by $\mathbf{P}_i \subseteq \{v_1, v_2, \ldots, v_n\}$. Next, the administrator sends probes along each path to measure the failure status of vertices. The *measurement* for the probe along path i is given by

$$y_i = \sum_{v_j \in \mathbf{P}_i} l(v_j). \tag{1}$$

The measurement $y_i \in [0, \infty)$ denotes the sum of the labels of the vertices in path i. If there is no failure vertex in path i, we have $y_i = 0$; otherwise, $y_i > 0$.

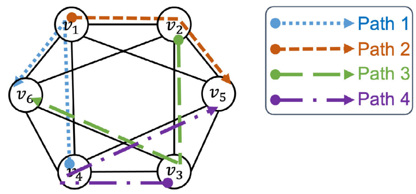


FIGURE 4. Failure detection of network by using group testing.

By sending the m probes, we obtain the vector $y := [\, y_1 \ y_2 \ \cdots \ y_m \,]^\top \in [0, \infty)^m$. Then, from (1), we have

$$y = Cx, \tag{2}$$

where $x := [\, l(v_1) \ l(v_2) \ \cdots \ l(v_n) \,]^\top \in [0, \infty)^n$, and C is an $m \times n$, $\{0, 1\}$-valued matrix, called the *measurement matrix*, whose (i, j)-element c_{ij} represents whether v_j is in path i or not, i.e.,

$$c_{ij} = \begin{cases} 1 & v_j \in \mathbf{P}_i, \\ 0 & v_j \notin \mathbf{P}_i. \end{cases}$$

For example, consider the network in Fig. 4. Suppose that the administrator specifies four paths, for which $\mathbf{P}_1 = \{v_1, v_4, v_6\}$, $\mathbf{P}_2 = \{v_1, v_2, v_5\}$, $\mathbf{P}_3 = \{v_2, v_3, v_6\}$, $\mathbf{P}_4 = \{v_3, v_4, v_5\}$. Then the measurement matrix is given by

$$C = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}. \tag{3}$$

Since the administrator has information on C and y, (2) can be regarded as a linear equation with the unknown $x \in [0, \infty)^n$. If we can uniquely determine x by solving (2), the administrator can locate the failure vertices in the network.

In general, the number of measurements is less than the number of vertices in group testing, that is, $m < n$. It implies that the equation is underdetermined, and there are an infinite number of solutions of (2). Meanwhile, it is reasonable to assume that only a few vertices fail simultaneously, that is, x has very few nonzero elements. By assuming this, the failure vertex vector x may be uniquely determined from (2). The group testing is to find the failure vertices by solving the linear equation in (2) under the assumption that there exist at most f failure vertices.

For exact detection, it is important to appropriately choose probes, i.e., to design C. A promising method is given by the notion of f-identifiability for the matrix C. If a vector $x \in \mathbf{R}^n$ has at most f nonzero elements, the vector x is said to be *f-sparse*. Let $\mathbf{S}(f)$ denote the set of f-sparse vectors in $\mathbf{R}^n$. The following notion is concerned with the measurement matrix C.

Definition 1 (f-Identifiable Matrix): Consider the linear equation (2). The matrix $C \in \mathbf{R}^{m \times n}$ is said to be f-identifiable if (2) has a unique solution on $\mathbf{S}(f)$. □

If we can construct an f-identifiable matrix and set it to C, the failure vertices are uniquely determined. Now, how do we construct an f-identifiable matrix? An answer is given by the binary correlation matrices [14], [15], [16], [17], [18].

Definition 2 (Binary Correlation Matrix): Consider a matrix $C \in \{0,1\}^{m\times n}$, and let $c_i \in \{0,1\}^m$ denote the i-th column vector of C. The matrix C is called a binary correlation matrix if $c_i^\top c_j \leq 1$ holds for every $(i,j) \in \{1,2,\ldots,n\}^2$ such that $i \neq j$. □

We obtain a sufficient condition for a binary correlation matrix to be f-identifiable [21].

Lemma 1: Consider a binary correlation matrix $C \in \{0,1\}^{m\times n}$ and a nonnegative integer f. Let $d(C) := \min_{i\in\{1,2,\ldots,n\}} \|c_i\|_2$. If

$$d(C) > f, \tag{4}$$

then the matrix C is f-identifiable. □

Therefore, if C is a binary correlation matrices satisfying (4), C is f-identifiable.

C. PROBLEM FORMULATION

This paper aims at analyzing the performance limitation of group testing based on the binary correlation measurement matrix. Our problem is formulated as follows.

Let $\mathbf{G}(n)$ be the set of the networks with n vertices, consider a network $\mathcal{G} \in \mathbf{G}(n)$, and assume that there are at most f failure vertices in $\mathcal{G}$. Let $\mathbf{C}(\mathcal{G})$ be the set of all binary correlation measurement matrices of $\mathcal{G}$. Let $m(C)$ denote the row size of C. This is the number of measurements of group testing when employing measurement matrix $C \in \mathbf{C}(\mathcal{G})$ in network $\mathcal{G}$. Then, the *performance index* of group testing is defined as the minimum number of measurements of group testing for a given network $\mathcal{G}$, i.e.,

$$m(\mathcal{G}, f) := \min_{C\in\mathbf{C}(\mathcal{G})} m(C) \;\; \text{s.t. (4)}. \tag{5}$$

Thus, the *performance limitation* of group testing in network detection is represented as follows:

$$m^*(n, f) := \min_{\mathcal{G}\in\mathbf{G}(n)} m(\mathcal{G}, f). \tag{6}$$

Let $\mathcal{G}^*$ denote the optimal solution to the minimization problem in (6), and $m^*(n,f)$ is given by the minimum number of measurements of network $\mathcal{G}^*$.

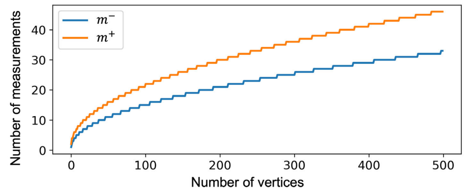
The problem we would like to address in this paper is as follows:

Problem 1: Consider the set of networks with n vertices, that is, $\mathbf{G}(n)$, and assume that there are at most f failure vertices in each network in $\mathbf{G}(n)$.

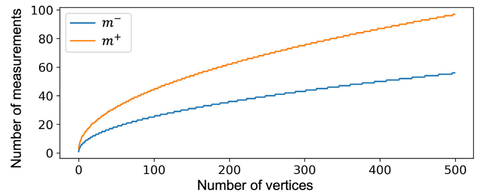
1) Derive $m^*(n,f)$.
2) Find $\mathcal{G}^*$.
3) For the network $\mathcal{G}^*$, find an f-identifiable measurement matrix $C \in \{0,1\}^{m^*(n,f)\times n}$. □

III. PERFORMANCE LIMITATION OF GROUP TESTING IN NETWORK FAILURE-VERTEX DETECTION

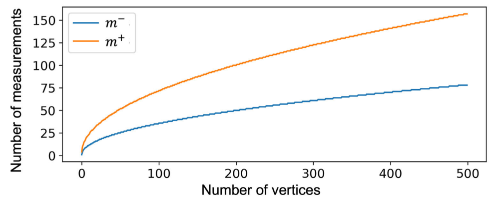
In this section, we present a solution to Problem 1 in Section II-C. Our main result is as follows.



(a) Case of $f = 1$



(b) Case of $f = 2$



(c) Case of $f = 3$

FIGURE 5. Performance limitation of group testing based on the binary correlation matrix in network failure-vertex detection.

Theorem 1: Consider the networks in $\mathbf{G}(n)$, and assume that there are at most f failure vertices in a network. Let

$$m^- = \left\lceil \sqrt{(f+1)nf + \frac{1}{4}} + \frac{1}{2} \right\rceil, \tag{7}$$

$$m^+ = \left\lceil (f+1)\sqrt{nf + \frac{1}{4}} + \frac{1}{2}(f+1) \right\rceil. \tag{8}$$

1) The relation

$$m^- \leq m^*(n,f) \leq m^+ \tag{9}$$

holds. If there is a binary correlation matrix $C \in \{0,1\}^{m^-\times n}$ with $d(C) = f+1$, then $m^*(n,f) = m^-$, i.e., m^- is the solution of 1) of Problem 1.

2) Consider a network $\mathcal{G} \in \mathbf{G}(n)$. If

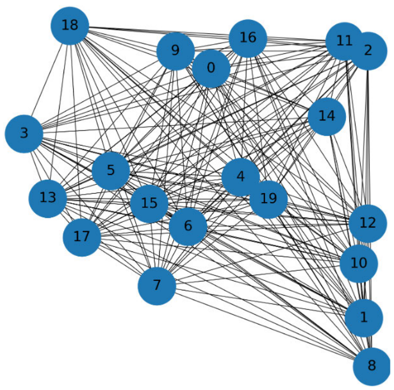
$$\delta(\mathcal{G}) \geq n - \frac{1}{2}\left\lfloor \frac{(f+1)n}{m^+} \right\rfloor, \tag{10}$$

then $\mathcal{G}$ is the solution to 2) of Problem 1.

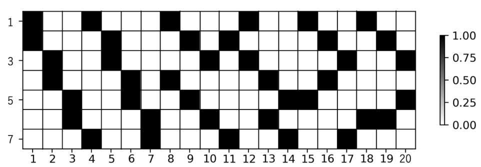
3) Let $\mathcal{G}^*$ be a network satisfying (10). A binary correlation measurement matrix $C \in \{0,1\}^{m^*(n,f)\times n}$ with $d(C) = f+1$ is a solution to 3) of Problem 1.

Proof of Theorem 1: See Appendix. ■

Theorem 1 1) gives the performance limitation of group testing based on the binary correlation matrix, whose upper and lower bounds in three cases of different f are shown in Fig. 5. It demonstrates that for a fixed f, the detection efficiency, evaluated by $n/m^*(n,f)$, increases as the network scale, i.e., n. Theorem 1 2) provides a sufficient condition for

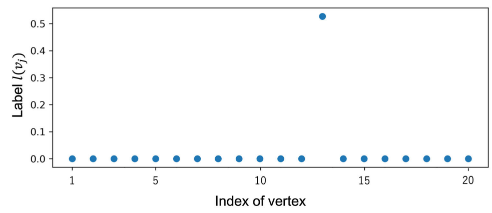


(a) Network $\mathcal{G}$ with $n = 20$ and $\delta(\mathcal{G}) = 18$

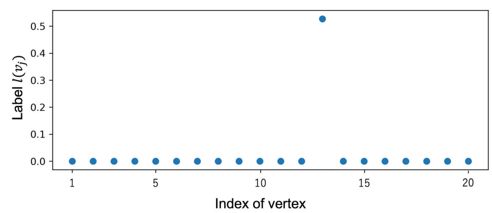


(b) Measurement matrix $C \in \{0, 1\}^{7\times 20}$ of network $\mathcal{G}$

FIGURE 6. **Network $\mathcal{G}$ and its detection scheme.**



(a) True delay vector



(b) Estimated delay vector

FIGURE 7. **Delay vector of network $\mathcal{G}$.**

networks, under which the administrator can construct an f-identifiable measurement matrix with optimal performance for the network, and the measurement matrix is given in Theorem 1 3). For example, consider the case of $n = 20$ and $f = 1$. From Theorem 1, we have $m^- = 7$ and $m^+ = 10$. Fig. 6 shows an example of a network $\mathcal{G}$ satisfying (10) and its measurement matrix $C \in \{0, 1\}^{7\times 20}$, where the black square in the i-th row and j-th column in Fig. 6 represents $c_{ij} = 1$, and the blank represents $c_{ij} = 0$. The matrix is constructed by the progressive edge-growth (PEG) algorithm, which is a promising method for constructing binary correlation matrices [14], [15], [16], [17], [18].

IV. NUMERICAL EXAMPLE

In this section, we show an example of the application of Theorem 1 to network failure detection. More specifically, we show a detection example that performs group testing with optimal performance for a network satisfying (10) by using the proposed measurement matrix.

A. SIMULATION CONDITIONS

Consider the network in Fig. 6(a). In this example, there is a failure vertex v_{13} with $l(v_{13}) = 0.55$ as shown in Fig. 7(a), where each dot represents the delay at a vertex.

Then, perform the group testing with the measurement matrix in Fig. 6(b), and the estimated results are obtained by solving the following optimization problem:

$$\min_{x\in\mathbf{R}^n} \|x\|_1 \quad \text{s.t. (2)}. \tag{11}$$

It is a popular approach for inferring a sparse vector from a linear equation, which can be cast as a linear programming problem and efficiently solved [22], [23], [24], [25].

The simulation was coded by Python and executed by the personal computer with CPU Intel (R) Core (TM) i7-1065G7, 1.30 [GHz] and memory 16.0 [GB].

B. DETECTION RESULT

Fig. 7(b) presents the estimated delay vector by the group testing. The computation time to estimate x was 0.003 seconds. From this result and Fig. 7(a), we see that the failure vertex can be identified accurately and efficiently.

This example suggests that by using our proposed measurement matrix, the administrator can exactly detect the failure vertices in the network with optimal performance.

V. CONCLUSION

In this study, the performance limitation of group testing based on the binary correlation measurement matrix was investigated, where the performance is evaluated by the number of needed measurements for network failure detection. We derived the upper and lower bounds of the minimum number of measurements needed for network detection. Moreover, we proposed a sufficient condition of network topology, under which the failure vertices in the network can be detected with optimal performance, and we also provided a detection scheme with guaranteed exactness for the network. Numerical example indicates that for the network that satisfies the proposed sufficient condition, the administrator can exactly detect the failure vertices with optimal performance by using our proposed detection scheme.

APPENDIX PROOF OF THEOREM 1

Theorem 1 is proved in this section. Table 1 summarizes the notations about the number of nonzero entries in matrix C that will be used in the proof.

TABLE 1. Summary of notations about the number of nonzero entries in matrix $C \in \{0,1\}^{m\times n}$.

Notation	Meaning
$d_{cj}(C)$	Number of nonzero entries in the j-th column vector of C
$d_{ri}(C)$	Number of nonzero entries in the i-th row vector of C
$t(C)$	Total number of nonzero entries in C
$d(C)$	Minimum number of nonzero entries in a column of C, i.e., $\min_{j\in\{1,2,\ldots,n\}} \|c_j\|_2 = \min_{j\in\{1,2,\ldots,n\}} d_{cj}(C)$
$p(C)$	Minimum number of nonzero entries in a row of C, i.e., $\min_{i\in\{1,2,\ldots,m\}} d_{ri}(C)$

A. THEOREM 1 1): DERIVING $m^*(n, f)$

1) PREPARATION

We first prepare the following lemmas from [17], [26].

Lemma 2: Consider a binary correlation matrix $C \in \{0,1\}^{m\times n}$. Let $d_{ri}(C)$ and $d_{cj}(C)$ denote the number of nonzero entries in the i-th row vector and j-th column vector of C, respectively, and let $t(C)$ denote the total number of nonzero entries in C, i.e., $t(C) = \sum_{i=1}^{m} d_{ri}(C) = \sum_{j=1}^{n} d_{cj}(C)$. Let $p(C) = \min_{i\in\{1,2,\ldots,m\}} d_{ri}(C)$. If $d(C) \geq 2$ and $p(C) \geq 2$, then

$$m \geq \sum_{k=0}^{2}\left(\left(\frac{t(C)}{n}-1\right)^{\lceil \frac{k}{2}\rceil}\left(\frac{t(C)}{m}-1\right)^{\lfloor \frac{k}{2}\rfloor}\right). \tag{12}$$

□

For example, consider the binary correlation matrix C in (3), in which $m = 4$, $n = 6$, and $t(C) = 12$. The term on the right side of (12) is equal to 4. It is clear that the relation in (12) holds.

Lemma 3: Consider three positive integers d_r, d_c, and m. If

$$\frac{\log\left(md_r - \frac{md_r}{d_c} - m + 1\right)}{\log\left((d_r-1)(d_c-1)\right)} - 1 \geq 1, \tag{13}$$

then there exist an $m \times n$ binary correlation matrix with

$$\max_{i\in\{1,2,\ldots,m\}} d_{ri}(C) = d_r, \tag{14}$$

$$\max_{j\in\{1,2,\ldots,n\}} d_{cj}(C) = d_c. \tag{15}$$

□

For example, consider $d_r = 3$ and $d_c = 2$, there exist an 6×7 binary correlation matrix given as follows

$$C = \begin{bmatrix} 1&0&0&1&0&0&1\\ 1&0&0&0&1&0&0\\ 0&1&0&1&0&0&0\\ 0&1&0&0&0&1&0\\ 0&0&1&0&1&0&0\\ 0&0&1&0&0&1&1 \end{bmatrix}, \tag{16}$$

and it is constructed by the PEG algorithm.

2) PROOF

Let $\mathbf{C}(n) = \{C \in \bigcup_{i=1}^{n}\{0,1\}^{i\times n} \mid C \text{ is a binary correlation matrix}\}$, and let

$$d_{min} = \min_{C\in\mathbf{C}(n)} d(C) \text{ s.t. (4)}. \tag{17}$$

Then, the statement 1) of Theorem 1 is the consequence of the following four facts:

(a) $d_{min} = f + 1$

(b) $m^*(n,f) \geq \left\lceil \sqrt{d_{min}n(d_{min}-1)+\frac{1}{4}}+\frac{1}{2}\right\rceil$

(c) $m^*(n,f) \leq \left\lceil d_{min}\sqrt{n(d_{min}-1)+\frac{1}{4}}+\frac{1}{2}d_{min}\right\rceil$

(d) If there is a $m \times n$ binary correlation matrix C with $m = \left\lceil \sqrt{d_{min}n(d_{min}-1)+\frac{1}{4}}+\frac{1}{2}\right\rceil$ and $d(C) = d_{min}$, then the equality holds in (b).

Next, we prove the four facts.

Proof of (a): Consider a matrix $C \in \mathbf{C}(n)$. Since $d(C) \in \mathbf{Z}_+$, then (4) implies that

$$d(C) \geq f + 1. \tag{18}$$

Thus, we obtain fact (a).

Proof of (b): Consider an $m \times n$ matrix $C \in \mathbf{C}(n)$ with $d(C) \geq 2$ and $p(C) \geq 2$. It is clear that

$$t(C) \geq nd(C). \tag{19}$$

From Lemma 2, we have

$$m \geq 1 + (d(C)-1) + (d(C)-1)\left(\frac{nd(C)}{m}-1\right), \tag{20}$$

which implies

$$m \geq \sqrt{d(C)n(d(C)-1)+\frac{1}{4}}+\frac{1}{2}. \tag{21}$$

From (21), a small value of $d(C)$ improves the bound of m. Because d_{min} is the smallest integer that satisfies (4), then by regarding d_{min} as $d(C)$ in (21), from Lemma 1, we have fact (b).

Proof of (c): The condition in (13) can be simplified to

$$m \geq d_c + d_c(d_c-1)(d_r-1), \tag{22}$$

from which small values of d_r and d_c improve the bound of m. Then, let us consider a binary correlation matrix $C \in \{0,1\}^{m\times n}$ with d_c nonzero entries per column and at most $\left\lceil \frac{nd_c}{m}\right\rceil$ nonzero entries in each row. From Lemma 3, a sufficient condition for the existence of such C is

$$m \geq d_c + d_c(d_c-1)\frac{nd_c}{m}, \tag{23}$$

that is,

$$m \geq d_c\sqrt{n(d_c-1)+\frac{1}{4}}+\frac{1}{2}d_c, \tag{24}$$

Therefore, by regarding d_{min} as d_c in (24), we have a sufficient condition for there exist an $m \times n$, f-identifiable matrix $C \in \mathbf{C}(n)$ is

$$m \geq d_{min}\sqrt{n(d_{min}-1)+\frac{1}{4}}+\frac{1}{2}d_{min}. \tag{25}$$

This proves fact (c).

Proof of (d): Let $\mathbf{C}(n, d_{min}) = \{C \in \mathbf{C}(n) \mid d(C) \geq d_{min}\}$. In other words, $\mathbf{C}(n, d_{min})$ is the set of binary correlation matrices of column size n that satisfy (4). By the definition of $\mathbf{C}(n, d_{min})$ and $m^*(n,f)$, we have

$$m^*(n,f) = \min_{C\in\mathbf{C}(n,d_{min})} m(C). \tag{26}$$

Next, let us consider an $m \times n$ binary correlation matrix C with $m = \left\lceil \sqrt{d_{min}n(d_{min}-1)+\frac{1}{4}}+\frac{1}{2} \right\rceil$ and $d(C) = d_{min}$. It is clear that the matrix $C \in \mathbf{C}(n, d_{min})$. Then, from fact (b), we have fact (d).

B. THEOREM 1 2): FINDING $\mathcal{G}^*$

1) PREPARATION

We prepare a lemma from [27] and [28] that will be used to prove our result.

Lemma 4: Consider a graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ with $|\mathbf{V}| = n$ and a nonnegative integer q. If

$$\delta(\mathcal{G}) \geq \frac{n+q}{2}, \tag{27}$$

then for each set of vertices $\mathbf{Q} \subseteq \mathbf{V}$ such that $|\mathbf{Q}| \leq q$, the subgraph of $\mathcal{G}$ induced by $\mathbf{V} \setminus \mathbf{Q}$ has a Hamiltonian path. □

2) PROOF

The statement 2) of Theorem 1 is the consequence of the following two facts:

(e) Consider a network $\mathcal{G} \in \mathbf{G}(n)$ and a matrix $C \in \mathbf{C}(n)$. If

$$\delta(\mathcal{G}) \geq n - \frac{p(C)}{2}, \tag{28}$$

then $C \in \mathbf{C}(\mathcal{G})$.

(f) There is an $m^*(n,f) \times n$ matrix $C \in \mathbf{C}(n, d_{min})$ with

$$p(C) \geq \left\lfloor \frac{n(f+1)}{m^+} \right\rfloor. \tag{29}$$

Facts (e) and (f) are proved as follows, respectively.

Proof of (e): Consider a matrix $C \in \mathbf{C}(n)$, let $\mathbf{S}_i(C)$ denote the index set of the nonzero entries in the i-th row of C. Regard C as the measurement matrix of network $\mathcal{G} \in \mathbf{G}(n)$, and let $\mathbf{T}_i(C) = \{v_j \in \mathbf{V} \mid j \in \mathbf{S}_i(C)\}$. In other words, $\mathbf{T}_i(C)$ is the set of the vertices to be measured by probe i. From Lemma 4, if (28) holds, then for each set of vertices $\mathbf{T} \subseteq \mathbf{V}$ such that $|\mathbf{T}| \geq p(C)$, there is a Hamiltonian path in $\mathcal{G}_{\mathbf{T}}$. Because $|\mathbf{S}_i(C)| = |\mathbf{T}_i(C)| \geq p(C)$ holds for each $i \in \{1, 2, \ldots, m\}$, we have for each $i \in \{1, 2, \ldots, m\}$, there is a Hamiltonian path in $\mathcal{G}_{\mathbf{T}_i(C)}$. By the definition of the Hamiltonian path, there is a path that traverses each vertex in $\mathbf{T}_i(C)$ exactly once. The administrator can send a probe along this path and the measurement for the probe is the sum of the delay of the vertices in the path. It indicates that the matrix C is feasible in the network $\mathcal{G}$, that is, $C \in \mathbf{C}(\mathcal{G})$. This proves fact (e).

Proof of (f): From [18], there is an $m^*(n,f) \times n$ matrix $C \in \mathbf{C}(n, d_{min})$ with

$$d(C) = d_{min} \tag{30}$$

$$p(C) = \left\lfloor \frac{t(C)}{m^*(n,f)} \right\rfloor. \tag{31}$$

By the definition of $\mathbf{C}(n, d_{min})$ and fact (a), we have for the matrix C,

$$t(C) \geq n(f+1), \tag{32}$$

which implies

$$p(C) \geq \left\lfloor \frac{n(f+1)}{m^*(n,f)} \right\rfloor. \tag{33}$$

Then fact (c) and (33) prove fact (f).

C. THEOREM 1 3)

In this section, we find an f-identifiable measurement matrix for network $\mathcal{G}^*$.

From facts (a), (e), and (f), there is a binary correlation matrix $C \in \{0, 1\}^{m^*(n,f)\times n}$ with $d(C) = f+1$ in $\mathbf{C}(\mathcal{G}^*)$. From Lemma 1, the matrix C is f-identifiable, and this completes the proof of Theorem 1 3).

REFERENCES

[1] M. Subramanian, T. A. Gonsalves, and N. U. Rani, *Network Management: Principles and Practice*. New Delhi, India: Dorling Kindersley (India) Pvt. Ltd, 2010.
[2] L. Ma, T. He, A. Swami, D. Towsley, and K. K. Leung, "Network capability in localizing node failures via end-to-end path measurements," *IEEE/ACM Trans. Netw.*, vol. 25, no. 1, pp. 434–450, Feb. 2017.
[3] J. Tapolcai, L. Rónyai, É. Hosszu, L. Gyimóthi, P.-H. Ho, and S. Subramaniam, "Signaling free localization of node failures in all-optical networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2527–2538, Jun. 2016.
[4] W. Xu, M. Wang, E. Mallada, and A. Tang, "Recent results on sparse recovery over graphs," in *Proc. Conf. Rec. Forty 5th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Nov. 2011, pp. 413–417.
[5] J. Staddon, D. Balfanz, and G. Durfee, "Efficient tracing of failed nodes in sensor networks," in *Proc. 1st ACM Int. Workshop Wireless Sensor Netw. Appl.*, Sep. 2002, pp. 122–130.
[6] J. L. Gross, J. Yellen, and M. Anderson, *Graph Theory and Its Applications*. New York, NY, USA: Chapman & Hall, 2018.
[7] R. Ramaswami, K. N. Sivarajan, and G. Sasaki, *Optical Networks: A Practical Perspective*. Burlington, MA, USA: Morgan Kaufmann, 2009.
[8] M. Cheraghchi, A. Karbasi, S. Mohajer, and V. Saligrama, "Graph-constrained group testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 248–262, Jan. 2012.
[9] M. Wang, W. Xu, E. Mallada, and A. Tang, "Sparse recovery with graph constraints," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 1028–1044, Feb. 2015.
[10] H. Zheng, W. Guo, and N. Xiong, "A kernel-based compressive sensing approach for mobile data gathering in wireless sensor network systems," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 48, no. 12, pp. 2315–2327, Dec. 2018.
[11] W. Xu, E. Mallada, and A. Tang, "Compressive sensing over graphs," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2087–2095.
[12] D. Du and F. K. Hwang, *Combinatorial Group Testing and Its Applications*. Singapore: World Scientific, 2000.

[13] F. Xu, S.-I. Azuma, R. Ariizumi, and T. Asai, "Minimum number of measurements in enhanced group testing for failure-edge detection in networks," in *Proc. 13th Asian Control Conf. (ASCC)*, May 2022, pp. 1822–1827.

[14] F. Tong, L. Li, H. Peng, and Y. Yang, "Flexible construction of compressed sensing matrices with low storage space and low coherence," *Signal Process.*, vol. 182, May 2021, Art. no. 107951.

[15] W. Lu, T. Dai, and S.-T. Xia, "Compressed sensing performance of binary matrices with binary column correlations," in *Proc. Data Compress. Conf. (DCC)*, Apr. 2017, pp. 151–160.

[16] W. Lu, K. Kpalma, and J. Ronsin, "Sparse binary matrices of LDPC codes for compressed sensing," in *Proc. Data Compress. Conf.*, Apr. 2012, p. 405.

[17] X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.

[18] H. Chen and Z. Cao, "A modified PEG algorithm for construction of LDPC codes with strictly concentrated check-node degree distributions," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2007, pp. 564–568.

[19] R. J. Wilson, *Introduction to Graph Theory*. Harlow, U.K.: Longman Group Ltd., 1979.

[20] C. Godsil and G. F. Royle, *Algebraic Graph Theory*. New York, NY, USA: Springer, 2001.

[21] S.-T. Xia, X.-J. Liu, Y. Jiang, and H.-T. Zheng, "Deterministic constructions of binary measurement matrices from finite geometry," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1017–1029, Feb. 2015.

[22] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[23] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

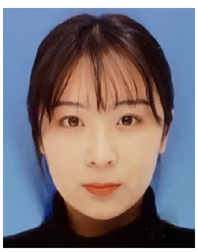
[24] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," in *Proc. PNAS*, vol. 100, no. 5. Washington, DC, USA, 2003, pp. 2197–2202.

[25] I. Rish and G. Grabarnik, *Sparse Modeling: Theory, Algorithms, and Applications*. Boca Raton, FL, USA: CRC Press, 2014.

[26] S. Hoory, "The size of bipartite graphs with a given girth," *J. Combinat. Theory, Ser. B*, vol. 86, no. 2, pp. 215–220, Nov. 2002.

[27] G. Chartrand, S. F. Kapoor, and D. R. Lick, "*n*-Hamiltonian graphs," *J. Comb. Theory*, vol. 9, no. 3, pp. 308–312, 1970.

[28] Y. Li, W. Liu, and L. Feng, "A survey on spectral conditions for some extremal graph problems," 2021, *arXiv:2111.03309*.



FANGYUAN XU was born in China, in 1995. She received the B.E. degree from the Shandong University of Science and Technology, Qingdao, Shandong, China, in 2017, and the M.E. degree from Nagoya University, Nagoya, Japan, in 2021, where she is currently pursuing the Ph.D. degree with the Graduate School of Engineering. Her research interests include graph theory and network systems.

SHUN-ICHI AZUMA (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from Hiroshima University, in 1999, and the M.Eng. degree in control engineering and the Ph.D. degree in mechanical and environmental informatics from the Tokyo Institute of Technology, in 2001 and 2004, respectively. He was an Assistant Professor and an Associate Professor with the Graduate School of Informatics, Kyoto University, from 2005 to 2011 and from 2011 to 2017, respectively, and a Professor with the Graduate School of Engineering, Nagoya University. He is currently a Professor with the Graduate School of Informatics, Kyoto University. His research interests include analysis and control of network systems and hybrid systems. He served as an Associate Editor for IEEE Transactions on Control of Network Systems, from 2013 to 2019, IEEE Control Systems Letters, from 2016 to 2019, and *International Journal of Control, Automation and Systems*, from 2018 to 2023. He has been serving as an Associate Editor for the IEEE CSS Conference Editorial Board, since 2011, *Automatica*, since 2014, *Nonlinear Analysis: Hybrid Systems*, since 2017, and IEEE Transactions on Automatic Control, since 2019, and a Senior Editor for *International Journal of Control, Automation and Systems*, since 2023.

RYO ARIIZUMI (Member, IEEE) was born in Japan, in 1986. He received the B.S., M.S., and Ph.D. degrees from Kyoto University, Japan, in 2010, 2012, and 2015, respectively. He was a Research Fellow with the Japan Society for the Promotion of Science, from 2014 to 2015, and an Assistant Professor with Nagoya University, Nagoya, Japan, from 2015 to 2023. He is currently an Associate Professor with the Tokyo University of Agriculture and Technology, Tokyo, Japan. His research interests include the control of redundant robots and the optimization of robotic systems. He has received awards, including the IEEE Robotics and Automation Society Japan Chapter Young Award (IROS 2014), in 2014, and the Best Paper Award from the Robotics Society of Japan (RSJ), in 2018.

TORU ASAI (Member, IEEE) was born in Japan, in 1968. He received the B.S., M.S., and Ph.D. degrees from the Tokyo Institute of Technology, Japan, in 1991, 1993, and 1996, respectively. He was a Research Fellow with JSPS, from 1996 to 1998. In 1999, he joined Osaka University, Suita, Japan, as a Faculty Member. Since 2015, he has been an Associate Professor with the Subdepartment of Mechatronics, Nagoya University, Nagoya, Japan. His research interests include robust control, switching control, and control applications to industry. He has received paper awards from the Society of Instrumental and Control Engineers (SICE), in 1997, 2005, and 2006, and the Institute of Systems, Control and Information Engineers (ISCIE), in 2003.