

Fault Tolerability Analysis of Split-Star Networks Based on Component Fault Pattern

Xiuzhen Zhu, Yanze Huang*, *Member, IEEE*, Limei Lin*, *Senior Member, IEEE*, Xiaoding Wang, *Senior Member, IEEE*, Sun-Yuan Hsieh, *Fellow, IEEE*, Jie Wu, *Fellow, IEEE*

Abstract—The growing reliance on multiprocessor systems in high-performance computing underscores the need for robust fault diagnosis, particularly under processor failures that disrupt network connectivity. Component diagnosability, denoted $ct_r(G)$, measures the maximum number of faulty processors a network G can tolerate while remaining diagnosable. This study improves fault tolerance through a rigorous analysis of component diagnosability, providing key insights to design large-scale resilient systems. This paper improves fault tolerance through a rigorous analysis of component diagnosability, offering valuable information for the design of resilient large-scale computing systems. We make two key contributions: first, we establish tight bounds for the $\{2, 3\}$ -component diagnosability of split-star networks (S_n^2) under the PMC model using innovative linear analysis techniques and a detailed examination of fault set indistinguishability patterns. Second, we introduce NCFDSSN, a novel algorithm that integrates structural graph theory with machine learning to achieve the exact g -component t -diagnosability. The algorithm is rigorously validated through extensive experiments in both synthetic and real-world network traffic datasets. NCFDSSN demonstrates exceptional diagnostic performance, achieving perfect detection (all metrics equal to 1) for fault thresholds of up to 30. At a threshold of 35, it maintains an accuracy of 0.9994, precision of 0.9982, $F1$ -score of 0.999, and G -Mean of 0.9991. Performance degradation is gradual even at higher thresholds; at a threshold of 70, it still attains an accuracy of 0.9597, precision of 0.9706, $F1$ -score of 0.9653, and G -Mean of 0.9653, while retaining low error rates (FPR: 0.0407; FNR: 0.04). These results substantially surpass existing methods, particularly in large-scale fault scenarios. Our contributions advance both the theoretical understanding of network diagnosability and practical fault detection mechanisms for mission-critical computing systems requiring ultra-high reliability.

Index Terms—Reliability, Component diagnosability, Interconnection networks, Split-star networks.

I. INTRODUCTION

IN the field of parallel computing, the reliability of multiprocessor systems has emerged as a critical concern, since

Xiuzhen Zhu, Limei Lin (*Corresponding author) and Xiaoding Wang are with the College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China. E-mail: fjnuzxz0715@163.com, linlimei@fjnu.edu.cn, wangdin1982@fjnu.edu.cn.

Yanze Huang (*Corresponding author) is with the School of Computing and Data Science, Fujian University of Technology, Fuzhou 350118, China. E-mail: yzhuang@fjut.edu.cn.

Sun-Yuan Hsieh is with the Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan. E-mail: hsiehsy@mail.ncku.edu.tw.

Jie Wu is with the China Telecom Cloud Computing Research Institute, Beijing, 100088, China and with the Department of Computer and Information Sciences, Temple University, 1925 N. 12th St., Philadelphia, PA 19122, USA. E-mail: jiewu@temple.edu.

all processors operate concurrently [1]. As the scale of nodes in a multiprocessor system continues to expand, accurately detecting malfunctioning nodes becomes essential to maintain both reliability and parallelism in computation [2]. The process of identifying and separating faulty nodes from those that operate correctly in a system is known as fault diagnosis [3], [4]. After a faulty node is identified, it can be replaced with a functioning one to ensure the continued reliability of the system [5].

The Preparata-Metze-Chien (PMC) model, proposed in a seminal work by Preparata, Metze, and Chien, is a commonly adopted framework for system-level diagnosis. In this model, diagnostic tests are carried out between processors directly connected [6]. A network is said to be t -diagnosable if all faulty processors, up to a maximum of t , can be diagnosed without replacement [7]. Traditional diagnosability does not impose any constraints on the type of fault distribution. However, certain fault distribution patterns are highly improbable. When the set of faulty nodes is large, removing it from the multiprocessor system leads to many components. In this context, diagnosability is closely tied to the number of resulting components [8]. To address this issue, Zhang et al. introduced a new concept called r -component diagnosability [9]. For a given graph G , this parameter, denoted as $ct_r(G)$, is defined as the maximum number of faulty nodes in a set X that can be correctly identified, under the condition that removing X from G leaves at least r different components.

Since its introduction, component diagnosability has attracted considerable attention in the fault diagnosis of interconnection networks. Zhang et al. [9] showed that for hypercubes Q_n with $n \geq 7$, the $(g + 1)$ -component diagnosability is $ct_{g+1}(Q_n) = -\frac{1}{2}g^2 + (n - \frac{3}{2})g + n$ for $g \leq n - 1$, and $ct_{n+1}(Q_n) = \frac{n}{2} + \frac{n}{2} - 2$ for $n \geq 7$. In a subsequent work, Zhang et al. [10] initially established general properties characterizing component diagnosability in regular networks under both the PMC and MM* models. Liu et al. [17] explored the relationship between component connectivity and component diagnosability. Sun et al. [18] analyzed how component connectivity correlates with component diagnosability in certain regular networks. Zhuang et al. [19] determined that under both the PMC and MM* models, the $(h + 1)$ -component diagnosability of the Bicube network BQ_n is $(h + 1)(n - 1) - \frac{h(h+1)}{2} + 1$ for $n \geq 7$ and $1 \leq h \leq n - 3$. Chen et al. [20] demonstrated that for $m \geq 4$, an m -dimensional component-composition graph G is $(\Omega(h), \kappa(G))$ -diagnosable, where $h = \frac{2^{m-2} \cdot \log(m-1)}{m-1}$ when $2^{m-1} \leq |V(G)| < m!$, and

TABLE I: Existing diagnosability results for the split-star network and known component diagnosability results for representative networks.

Diagnosability type	Network	Result	Conditions	Ref.
A. Existing diagnosability results for the split-star network				
Pessimistic diagnosability	Split-star network	$4n - 9$	$n \geq 4$	[10]
Strong diagnosability	Split-star network	$2n - 3$	$n = 4$	[11]
Conditional diagnosability	Split-star network	$6n - 16$	$n = 5$	[12]
Intermittent fault diagnosability	Split-star network	$2n - 4$	$n \geq 3$	[13]
Component diagnosability	Split-star network	—	—	—
B. Known component diagnosability results for representative networks				
Component diagnosability	Star graph network	$2rn - 2r$	$r = 2, 3, n \geq 4$	[14]
Component diagnosability	Alternating group networks	$2rn - (r + 1)^2 + r - 3$	$r = 2, 3, n \geq 4$	[15]
Component diagnosability	Hierarchical cubic network	$rn - \frac{1}{2}(r - 1)r + 1$	$n \geq 2, 1 \leq r \leq n - 1$	[15]
Component diagnosability	Hypercube network	$-\frac{1}{2}g^2 + (n - \frac{3}{2})g + n$	$g \leq n - 1$	[16]
Component diagnosability	Split-star network	—	—	—

$h = 2^{m-2}$ when $|V(G)| \geq m!$. Huang et al. [15] proved that under both diagnostic models, the r -component diagnosability of the hierarchical cubic network HCN_n is $rn - \frac{r(r-1)}{2} + 1$ for $n \geq 2$ and $1 \leq r \leq n - 1$. Tian et al. [16] investigated the relationship between component diagnosability and component connectivity of hypercubes under the PMC model.

Among various interconnection network topologies, the split-star network S_n^2 has attracted considerable attention due to its attractive properties such as symmetry, regularity, and high fault tolerance. Consequently, a substantial body of research has been devoted to analyzing its connectivity, reliability, and diagnosability. In terms of connectivity, Chen et al. [21] studied restricted connectivity for several families of interconnection networks, including split-star networks, thereby providing an early connectivity-based evaluation of their fault-tolerance capability. Subsequently, Lin et al. [22] examined the g -extra connectivity for $1 \leq g \leq 3$. Zhao et al. [23] investigated the structure connectivity and substructure connectivity of S_n^2 when the faulty structures are modeled as paths or cycles. Li et al. [24] proved that the split-star network possesses super spanning connectivity. Gu et al. [25] determined the $\{1, 2, 3\}$ -extra edge connectivity of S_n^2 , which are $4n - 8$, $6n - 15$, and $8n - 20$, respectively. Regarding diagnosability, Lin et al. [22] also demonstrated that the conditional diagnosability of S_n^2 under the PMC model is $6n - 16$. Song et al. [13] investigated the intermittent fault diagnosability of S_n^2 under the PMC model. Lin et al. [26] derived an upper bound for the $\{1, 2, 3\}$ -good-neighbor diagnosability under the MM* model. Kung et al. [27] adopted a combinatorial method to evaluate the subsystem reliability of the split-star network.

In summary, the above studies show that the split-star network has been extensively investigated in terms of connectivity, reliability, and several diagnosis-related measures, while component diagnosability has been established for a number of classical interconnection networks. However, to the best of our knowledge, the r -component diagnosability of the split-star network S_n^2 under the PMC model has not yet been studied. For clarity, we summarize the existing diagnosability-related results of the split-star network and the known component diagnosability results of several other networks in Table I. The major contributions of our work are listed as follows:

- We demonstrate that if a subset of vertices D with size

not exceeding $4n - 9$ is deleted from S_n^2 , the resulting network has at most two components. Likewise, when a vertex set D of size no greater than $6n - 14$ is removed from S_n^2 , the resulting graph is divided into no more than four disconnected components.

- By utilizing the indistinguishability properties of the constructed fault sets along with linear analysis methods for multiple faults, we establish the $\{2, 3\}$ -component diagnosability of S_n^2 under the PMC model. Specifically, we prove that $ct_2(S_n^2) = 4n - 9$ and $ct_3(S_n^2) = 6n - 14$.
- We propose a novel r -component t -diagnosable algorithm called NCFDSSN (Network Component Fault Diagnosis for Split-Star Networks) to accurately identify faulty processors in multiprocessor systems. This algorithm uniquely integrates structural graph theory with combinatorial optimization techniques to achieve precise fault detection under the PMC model. Rigorously validated on both synthetic benchmarks and real-world network topologies, NCFDSSN demonstrates superior diagnosability compared to existing methods.
- The proposed algorithm achieves outstanding diagnostic performance, delivering flawless fault detection (all evaluation metrics equal to 1) for failure thresholds as high as 30. Even at the more challenging threshold of 35, it sustains near-perfect accuracy (0.9994), precision (0.9982), $F1$ -score (0.999), and G -Mean (0.9991). As the threshold increases further, performance declines gracefully while maintaining high effectiveness—at a threshold of 70, it still achieves strong results (accuracy: 0.9597; precision: 0.9706; $F1$ -score: 0.9653; G -Mean: 0.9653) with minimal error rates (FPR: 0.0407; FNR: 0.04). These results represent a substantial improvement over state-of-the-art methods, particularly in large-scale fault scenarios. The advancements not only enhance theoretical understanding of network diagnosability but also enable highly reliable fault detection for mission-critical computing systems.

The remainder of this paper is organized as follows. Section II presents fundamental terms and definitions. Section III demonstrates the linear fault diagnosis analysis. Section IV proposes the component diagnosis of diagnosability of split-star network. The NCFDSSN algorithm is comprehensively described in Section V. Finally, Section VI concludes the paper

TABLE II: Summary of basic terms and notations

G	an undirected graph
V_G	a set of nodes of G
E_G	a set of edges of G
S_n^2	an n -dimensional split-star network
AG_n	an n -dimensional alternating group graph
$G - F$	a subgraph removing all nodes and edges with its end-node in F from G
$F_1 \triangle F_2$	$\{x \mid x \in F_1 \cup F_2, x \notin F_1 \cap F_2\}$
$E[F_1, F_2]$	the set of all edges with one end in F_1 , and the other end in F_2
$N_G(w)$	a set of nodes adjoining to a node w in G
$N_G(F)$	a set of neighbors of F in V_G not in G
$ct_r(G)$	the r -component diagnosability of G
$S_{n,E}^2$	an $(n-1)$ -subgraph of S_n^2 induced by the set of all even permutations
$S_{n,O}^2$	an $(n-1)$ -subgraph of S_n^2 induced by the set of all odd permutations

with a summary of findings.

II. PRELIMINARIES

This study starts by outlining several fundamental concepts from graph theory that are essential for the discussions that follow. After that, we establish the concept of component diagnosability within the framework of the PMC model. Next, we introduce the formal definition of the n -dimensional split-star network, denoted by S_n^2 . In addition, we explore the relationship between the split-star network and the alternating group graph, analyzing their structural connections.

A. Terminologies and Notations

An undirected graph can be denoted as $G = G(V_G, E_G)$, where V_G represents the set of vertices and E_G denotes the set of edges associated with graph G . Within a multiprocessor system, vertices symbolize individual processors, while edges indicate the communication channels that connect pairs of processors. Given any vertex $w \in V_G$, its neighborhood $N_G(w)$ is defined as the collection of all vertices directly connected to w by an edge, i.e., $N_G(w) = \{z \in V_G \mid wz \in E_G\}$. In a similar manner, for any vertex subset $F \subseteq V_G$, the neighborhood $N_G(F)$ is defined as the set of vertices outside F that share an edge with at least one vertex in F . Formally, this can be expressed as $N_G(F) = \{z \in V_G - F \mid \exists w \in F \text{ such that } wz \in E_G\}$. The expression $G - F$ refers to the subgraph obtained by deleting the vertex set F from G , along with all edges connected to any vertex in F . The degree of a vertex u in the graph G , written as $d(u)$, is defined as the total number of adjacent vertices connected to u , that is, $|N_G(u)|$. In addition, the collection of shared neighbors between two vertices w and z is denoted by $cn(G : w, z)$, which is defined as the intersection of their respective neighborhoods, i.e., $N_G(w) \cap N_G(z)$. Table II summarizes the basic terms and notations.

B. PMC Model and Component Diagnosability

Under the PMC model, for any edge x_1x_2 in graph G , processor x_1 is capable of testing x_2 , and vice versa. In this

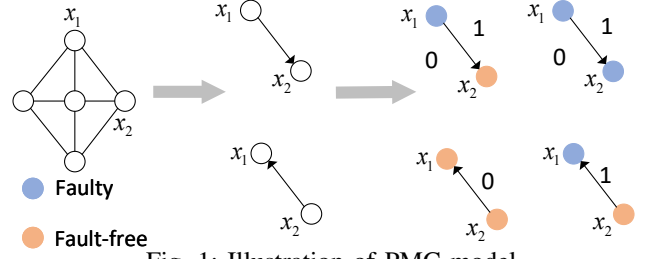


Fig. 1: Illustration of PMC model.

context, when x_1 performs the test on x_2 , x_1 is referred to as the tester, while x_2 is the node being tested. Further illustration of this mechanism is provided in Fig. 1.

Next, we will present several concepts associated with component diagnosability. Determining whether two sets of nodes are distinguishable plays a key role in performing fault diagnosis within the PMC framework. In real-world scenarios, a widely used approach—stated as Lemma 1—involves applying the following equivalent conditions to judge whether two vertex subsets can be distinguished under the PMC model.

Lemma 1. [8] *Under the PMC model, two distinct node sets X_1 and X_2 are said to be distinguishable if and only if there exists at least one edge between the set difference $V(G) - (X_1 \cup X_2)$ and the symmetric difference $X_1 \triangle X_2$; that is, $E[V(G) - (X_1 \cup X_2), X_1 \triangle X_2] \neq \emptyset$.*

Recently, Zhang introduced the concept of r -component diagnosability, which represents a new approach to fault diagnosis [9]. To begin with, we introduce the concept of an r -component cut in the context of a network G . This definition serves as a foundation for constructing distinguishable pairs within the network G and is subsequently employed to demonstrate the r -component diagnosability (see Definition 1). Furthermore, the concept of an r -component t -diagnosable network is defined by assessing the distinguishability of any two different r -component cuts, each involving at most t vertices. In addition, the r -component diagnosability of a network G is defined as the highest level of fault identification the system can inherently achieve based on r -component analysis.

Definition 1. [9] (1) *Let X be a subset of vertices in graph G . If removing X from G results in a subgraph with no fewer than r connected components, then X is defined as an r -component cut of the graph.*

(2) *A graph G is considered r -component t -diagnosable if and only if every pair of distinct r -component cuts D_1 and D_2 (where $D_1, D_2 \subseteq V(G)$, and both satisfy $|D_1| \leq t$ and $|D_2| \leq t$) can be uniquely identified and distinguished from one another.*

(3) *The r -component diagnosability of a graph G , expressed as $ct_r(G)$, refers to the largest integer t such that the graph remains r -component t -diagnosable.*

C. Introduction to Split-Star Network

We present Definition 2, which formally introduces the n -dimensional split-star network S_n^2 , along with a discussion of several fundamental characteristics associated with S_n^2 .

Lemma 5. [30] (1) An n -split-star S_n^2 is $(2n-3)$ -regular, and its connectivity $\kappa(S_n^2)$ and minimum degree $\delta(S_n^2)$ are both equal to $2n-3$ for $n \geq 2$.

(2) For any vertex in S_n^2 , its two external neighbors are located in different $(n-1)$ -subgraphs and are adjacent to each other.

(3) For any pair of vertices within the same $(n-1)$ -subgraph, their external neighbors are unique.

(4) A one-to-one correspondence exists between $S_{n,O}^2$ and $S_{n,E}^2$.

(5) Between any two $(n-1)$ -subgraphs, there are $2(n-2)!$ edge-disjoint connections.

Lemma 6. [32] Let D be an independent set in $V(S_n^2)$ where $n \geq 4$. We obtain the following results:

(1) When $|D| = 2$, the number of neighbors of D is at least $4n-8$, i.e., $|N_{S_n^2}(D)| \geq 4n-8$.

(2) When $|D| = 3$, the set D has at least $6n-14$ neighbors, meaning that $|N_{S_n^2}(D)| \geq 6n-14$.

(3) When $|D| = 4$, the number of neighbors of D is at least $8n-20$, i.e., $|N_{S_n^2}(D)| \geq 8n-20$.

Lemma 7. [26] Let F be a subset of $V(S_n^2)$ where $n \geq 4$ and $|F| \leq 4n-9$. In this case, $S_n^2 - F$ will exhibit one of the following outcomes:

(1) A single connected component.

(2) The graph consists of two components, with the smaller component being just a single vertex.

(3) The graph is divided into two components, with the smaller component consisting of a single edge.

(4) When $n = 4$, the graph $S_n^2 - F$ is composed of two separate copies of $C_4 \times K_2$.

A. Fault Tolerate Properties of 3-component

We then demonstrate that when any subset of vertices D with $|D| \leq 6n-14$ is removed from S_n^2 , the resulting network will contain no more than four components.

Theorem 1. For $D \subseteq V(S_n^2)$ with $n \geq 5$ and $|D| \leq 6n-14$, the subgraph $S_n^2 - D$ will have one of the following structures:

(1) A single connected component.

(2) The subgraph consists of two components, with the smaller component being either a single vertex, a 2-path, an edge, or a 3-cycle.

(3) The subgraph is divided into three components, where the two smaller components are either both vertices or consist of one vertex and one edge.

(4) The subgraph consists of four components, with the three smallest components each being a single vertex.

Proof. Let $D_O = D \cap V(S_{n,O}^2)$ and $D_E = D \cap V(S_{n,E}^2)$. We classify the situation into three cases based on the sizes of D_O and D_E .

Case 1. $|D_O| \leq 6n-19$ and $|D_E| \leq 6n-19$.

According to Lemma 4 (3), $S_{n,O}^2 - D_O$ (respectively, $S_{n,E}^2 - D_E$) has the following possible structures: (a) A single connected component, (b) two components, where the smaller component is either a vertex or an edge, or (c) three components, with the two smallest being vertices. Let Z_O (or

Z_E) denote the union of the smaller components in $S_{n,O}^2 - D_O$ (or $S_{n,E}^2 - D_E$), where $|Z_O| \leq 3$ and $|Z_E| \leq 3$. According to Lemma 5 (4), there are $\frac{n!}{2}$ edge-disjoint connections between $S_{n,O}^2$ and $S_{n,E}^2$. Given that $\frac{n!}{2} > 4n-9$ for $n \geq 5$, it follows that $S_{n,O}^2 - D_O - Z_O$ is connected to $S_{n,E}^2 - D_E - Z_E$.

If $|Z_O| + |Z_E| = 6$, then it follows that $|Z_O| = 3$ and $|Z_E| = 3$. From Lemma 4 (2), we know that for $n \geq 5$, $|D_O| \geq 6n-19$ and $|D_E| \geq 6n-19$. Therefore, for $n \geq 5$,

$$|D| = |D_O| + |D_E| \geq 2(6n-19) > 6n-14 \geq |D|,$$

which is inconsistent.

If $|Z_O| + |Z_E| = 5$, by the symmetry of S_n^2 , we have $|Z_O| = 2$ and $|Z_E| = 3$. According to Lemma 4 (1), for $n \geq 5$, $|D_O| \geq 4n-11$, and by Lemma 4 (2), $|D_E| \geq 6n-19$ for $n \geq 5$. Therefore, for $n \geq 5$,

$$|D| = |D_O| + |D_E| \geq 4n-11 + 6n-19 > 6n-14 \geq |D|,$$

which is inconsistent.

If $|Z_O| + |Z_E| = 4$, then $|Z_O| = |Z_E| = 2$. According to Lemma 4 (1), for $n \geq 5$, both $|D_O| \geq 4n-11$ and $|D_E| \geq 4n-11$. Therefore, for $n \geq 5$,

$$|D| = |D_O| + |D_E| \geq 4n-11 + 4n-11 > 6n-14 \geq |D|,$$

which is inconsistent.

If $|Z_O| + |Z_E| = 4$ and $|Z_O| \neq |Z_E|$, then by the symmetry of S_n^2 , we have $|Z_O| = 1$ and $|Z_E| = 3$. From Lemma 2 (1), it follows that $|D_O| \geq 2n-4$. Additionally, by Lemma 4 (2), $|D_E| \geq 6n-19$ for $n \geq 5$,

$$|D| = |D_O| + |D_E| \geq 2n-4 + 6n-19 > 6n-14 \geq |D|,$$

which is inconsistent.

If $|Z_O| + |Z_E| = 3$, then by the symmetry of S_n^2 , assume that $|Z_O| = 1$ and $|Z_E| = 2$. We can prove that the Theorem 1 (1) - (4) holds.

Case 2. $|D_O| \leq 6n-19$ and $|D_E| \geq 6n-18$.

Given that $|D| \leq 6n-14$, it follows that

$$|D_O| = |D| - |D_E| \leq (6n-14) - (6n-18) = 4.$$

Since $S_{n,E}^2$ is isomorphic to AG_n , and based on Lemma 1 (1), it follows that $S_{n,E}^2 - D_O$ remains connected. This is further supported by the fact that $2n-4 > 3$ for $n \geq 5$. Let H represent the union of all smaller components in $S_n^2 - D$, where $V(H)$ does not intersect with $V(S_{n,E}^2)$. As a result, $S_n^2 - D - V(H)$ remains connected, where $V(H) \subseteq V(S_{n,O}^2 - D_O)$. According to the definition of H , the neighbors of $V(H)$ in $S_{n,E}^2$ belong to D_E , while the neighbors of $V(H)$ in $S_{n,O}^2$ are contained in D_O . As stated in Lemma 5 (4), each vertex in $V(H)$ has a unique neighbor in $S_{n,O}^2$, and all these neighbors are distinct. Given that $|D_O| \leq 3$, it follows that $|V(H)| \leq 3$. When $|V(H)| \leq 3$, Theorem 1 (1)-(4) holds.

Case 3. $|D_O| \geq 6n-18$ and $|D_E| \geq 6n-18$.

For $n \geq 5$, we have

$$|D| = |D_O| + |D_E| \geq 6n-18 + 6n-18 > 6n-14 \geq |D|,$$

its a inconsistent.

From the above, Theorem 1 holds. ■

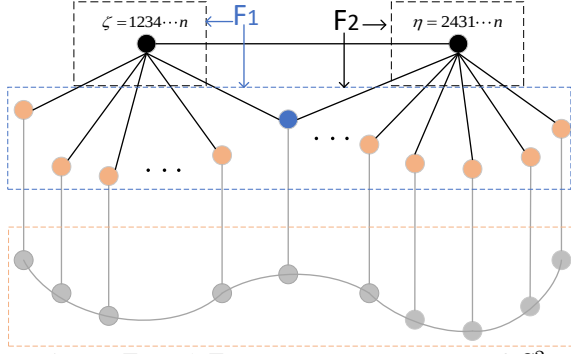


Fig. 3: F_1 and F_2 are 2-component-cuts of S_n^2 .

IV. COMPONENT DIAGNOSIS OF DIAGNOSABILITY OF SPLIT-STAR NETWORK

Theorem 2. *The 2-component diagnosability of S_n^2 is $4n - 9$ for $n \geq 4$ under PMC model.*

Proof. Let $X = \{\zeta, \eta\}$, where $\zeta = 1234 \cdots n$ and $\eta = 2431 \cdots n$, so that $\zeta\eta \in E(S_n^2)$ and $|X| = 2$.

It follows that $S_n^2 - N(X)$ consists of at least two components, with the smaller component being the isolated edge $\zeta\eta$. Indeed, by Lemma 7, $S_n^2 - N(X)$ contains precisely two components. Hence, by Definition 1, for $n \geq 4$, $N(X)$ is a 2-component cut of S_n^2 (as shown in Fig. 3).

Let $F_1 = N_{S_n^2}(X) \cup \{\zeta\}$, $F_2 = N_{S_n^2}(X) \cup \{\eta\}$. By Definition 2 and Lemma 5(1), the split-star network S_n^2 is $(2n - 3)$ -regular. Moreover, for the chosen adjacent vertices ζ and η , one can verify directly that they have exactly one common neighbor. Hence,

$$|N_{S_n^2}(X)| = (2n - 3 - 1) + (2n - 3 - 1) - 1 = 4n - 9.$$

Therefore,

$$|F_1| = |F_2| = |N_{S_n^2}(X)| + 1 = 4n - 8.$$

Note that $S_n^2 - F_1$ contains the isolated vertex η , and

$$|V(S_n^2) - F_1 - \{\eta\}| = n! - (4n - 8) - 1 > 0.$$

Thus, $S_n^2 - F_1$ has at least two components. Similarly, $S_n^2 - F_2$ contains the isolated vertex ζ , and

$$|V(S_n^2) - F_2 - \{\zeta\}| = n! - (4n - 8) - 1 > 0.$$

Hence, $S_n^2 - F_2$ also has at least two components. By Definition 1, both F_1 and F_2 are 2-component cuts of S_n^2 .

Since

$$F_1 \triangle F_2 = \{\zeta, \eta\} \quad \text{and} \quad F_1 \cup F_2 = N_{S_n^2}(X) \cup X,$$

every neighbor of ζ or η belongs to $F_1 \cup F_2$. Therefore,

$$\begin{aligned} & E[F_1 \triangle F_2, V(S_n^2 - (F_1 \cup F_2))] \\ &= E[\{\zeta, \eta\}, V(S_n^2 - (F_1 \cup F_2))] \\ &= \emptyset. \end{aligned}$$

According to Lemma 1, F_1 and F_2 cannot be distinguished under the PMC model. Hence, S_n^2 is not 2-component $(4n - 8)$ -diagnosable under the PMC model, and therefore $ct_r(S_n^2) \leq 4n - 9$, for $r = 2$ and $n \geq 4$.

Then, we aim to demonstrate that $ct_2(S_n^2) \geq 4n - 9$ for

$n \geq 4$. By Definition 1, it suffices to prove that S_n^2 is 2-component $(4n - 9)$ -diagnosable under the PMC model. We proceed by contradiction. Assume that there exist two distinct 2-component cuts F_1 and F_2 such that

$$|F_1| \leq 4n - 9, \quad |F_2| \leq 4n - 9,$$

but F_1 and F_2 are indistinguishable under the PMC model. Without loss of generality, assume that $F_1 - F_2 \neq \emptyset$.

Since $|F_1|, |F_2| \leq 4n - 9$, by Lemma 7, the graph $S_n^2 - F_1$ consists of one large component C_1 and at least one small component, whose total number of vertices is at most 2. Let

$$U_1 = V(S_n^2 - F_1 - C_1),$$

and let $nc(U_1)$ denote the number of connected components in the subgraph induced by U_1 . Then $nc(U_1) \geq 1, |U_1| \leq 2$, and

$$\begin{aligned} |V(C_1)| &= |V(S_n^2)| - |F_1| - |U_1| \\ &\geq n! - (4n - 9) - 2 \\ &= n! - 4n + 7. \end{aligned}$$

Moreover,

$$\begin{aligned} |V(C_1) - F_2| &\geq |V(C_1)| - |F_2| \\ &\geq (n! - 4n + 7) - (4n - 9) \\ &= n! - 8n + 16 \\ &> 2. \end{aligned}$$

Suppose that $F_2 - F_1 \not\subseteq U_1$. Then there exists a vertex

$$y \in (F_2 - F_1) \cap V(C_1).$$

Since C_1 is connected and $|V(C_1) \setminus F_2| > 2$, there exists an edge $yz \in E(S_n^2)$ such that $z \in V(C_1) \setminus F_2 \subseteq V(S_n^2 - (F_1 \cup F_2))$. Hence $E[F_1 \triangle F_2, V(S_n^2 - (F_1 \cup F_2))] \neq \emptyset$, which contradicts Lemma 1. Therefore, $F_2 - F_1 \subseteq U_1$, and consequently, $|F_2 - F_1| \leq |U_1| \leq 2$.

Since $E[U_1, C_1] = \emptyset$, it follows that $E[F_2 - F_1, C_1] = \emptyset$. Moreover, because F_1 and F_2 are indistinguishable under the PMC model, we also have $E[F_1 - F_2, C_1] = \emptyset$. Hence C_1 is a component of both $S_n^2 - F_2$ and $S_n^2 - (F_1 \cap F_2)$. Since F_2 is also a 2-component cut of S_n^2 , the graph $S_n^2 - F_2$ contains the large component C_1 and at least one additional component. Let $U_2 = V(S_n^2 - F_2 - C_1)$. Then $nc(U_2) \geq 1, |U_2| \geq 1$. By symmetry, we also obtain $F_1 - F_2 \subseteq U_2$.

Since $|F_1 \cap F_2| = |F_1| - |F_1 - F_2| \leq 4n - 10$, Lemma 7 implies that $S_n^2 - (F_1 \cap F_2)$ has the large component C_1 and at most one small component, and any such small component has at most two vertices. Now we consider two cases.

Case 1. $U_1 - (F_2 - F_1) \neq \emptyset$.

Choose $x \in U_1 - (F_2 - F_1)$. Then $x \notin F_1$ and $x \notin F_2$, so $x \in V(S_n^2 - (F_1 \cup F_2))$. On the other hand, since $F_1 - F_2 \neq \emptyset$, choose $y \in F_1 - F_2$. Then $y \in F_1 \triangle F_2$. If $xy \in E(S_n^2)$, then $E[F_1 \triangle F_2, V(S_n^2 - (F_1 \cup F_2))] \neq \emptyset$, contrary to Lemma 1. Hence $xy \notin E(S_n^2)$.

Furthermore, x has no edge to C_1 because $x \in U_1$, and y has no edge to C_1 because $E[F_1 - F_2, C_1] = \emptyset$. Therefore, in $S_n^2 - (F_1 \cap F_2)$, the vertices x and y lie in two different components outside C_1 . Hence $S_n^2 - (F_1 \cap F_2)$ has at least

two small components, contradicting Lemma 7.

Case 2. $U_1 - (F_2 - F_1) = \emptyset$.

Then $U_1 = F_2 - F_1$. Thus, outside the large component C_1 , the graph $S_n^2 - (F_1 \cap F_2)$ contains the vertices of $U_1 \cup (F_1 - F_2)$, where both sets are nonempty.

If $|U_1| + |F_1 - F_2| \geq 3$, then $S_n^2 - (F_1 \cap F_2)$ has either at least two small components or one small component with at least three vertices, both of which contradict Lemma 7. Hence we must have $|U_1| = |F_1 - F_2| = 1$. Let $U_1 = \{u\}$, $F_1 - F_2 = \{v\}$. If $uv \notin E(S_n^2)$, then u and v form two distinct small components in $S_n^2 - (F_1 \cap F_2)$, again contradicting Lemma 7. Therefore, we must have $uv \in E(S_n^2)$. Hence uv is an isolated edge of $S_n^2 - (F_1 \cap F_2)$, which implies $N_{S_n^2}(\{u, v\}) \subseteq F_1 \cap F_2$. Since u and v are adjacent in S_n^2 , each has degree $2n - 3$, and they have exactly one common neighbor, we obtain

$$|N_{S_n^2}(\{u, v\})| = (2n - 3 - 1) + (2n - 3 - 1) - 1 = 4n - 9.$$

Thus, $|F_1 \cap F_2| \geq 4n - 9$, which contradicts $|F_1 \cap F_2| \leq 4n - 10$.

Both cases lead to contradictions. Therefore, S_n^2 is $(4n - 9)$ -diagnosable under the PMC model. Hence, $ct_2(S_n^2) \geq 4n - 9$.

Thus, it can be concluded that S_n^2 is $(4n - 9)$ -diagnosable under the PMC model. ■

Theorem 3. Under the PMC model, the 3-component diagnosability of S_n^2 is $6n - 14$ for $n \geq 5$.

Proof. Let $X = \{\zeta, \eta, \xi\}$ with $\zeta = 1234l \cdots n$, $\eta = 2431 \cdots n$, $\xi = 3412 \cdots n$ such that $\zeta\eta \in E(S_n^2)$, $\xi\zeta \notin E(S_n^2)$, and $\xi\eta \notin E(S_n^2)$.

According to Theorem 1, for $n \geq 5$, it can be demonstrated that $S_n^2 - N(X)$ contains at least three components, the two smaller components being the isolated vertex ξ and the edge $\zeta\eta$. In fact, by Theorem 1, $S_n^2 - N(X)$ consists of exactly three components (see Fig. 4). According to Definition 1, for $n \geq 4$, $N(X)$ forms a 3-component cut of S_n^2 . Let $F_1 = N_{S_n^2}(X) \cup \{\zeta\}$, $F_2 = N_{S_n^2}(X) \cup \{\eta\}$.

By Definition 2 and Lemma 5(1), the split-star network S_n^2 is $(2n - 3)$ -regular. Moreover, for the chosen vertices ζ, η, ξ , one can verify directly that $|N_{S_n^2}(X)| = 6n - 14$.

Therefore, $|F_1| = |F_2| = |N_{S_n^2}(X)| + 1 = 6n - 13$.

Note that $S_n^2 - F_1$ contains the isolated vertices η and ξ , and $|V(S_n^2) - F_1 - \{\eta, \xi\}| = n! - (6n - 13) - 2 > 0$. Hence, $S_n^2 - F_1$ has at least 3 components. Similarly, $S_n^2 - F_2$ contains the isolated vertices ζ and ξ , and $|V(S_n^2) - F_2 - \{\zeta, \xi\}| = n! - (6n - 13) - 2 > 0$. Hence, $S_n^2 - F_2$ also has at least 3 components. By Definition 1, both F_1 and F_2 are 3-component cuts of S_n^2 .

Since $F_1 \triangle F_2 = \{\zeta, \eta\}$ and $F_1 \cup F_2 = N_{S_n^2}(X) \cup \{\zeta, \eta\}$, since the only component of $S_n^2 - N(X)$ containing ζ or η is the edge $\zeta\eta$, every neighbor of ζ or η belongs to $F_1 \cup F_2$. Therefore,

$$\begin{aligned} & E[F_1 \triangle F_2, V(S_n^2 - (F_1 \cup F_2))] \\ &= E[\{\zeta, \eta\}, V(S_n^2 - (F_1 \cup F_2))] \\ &= \emptyset. \end{aligned}$$

By Lemma 1, F_1 and F_2 cannot be distinguished under

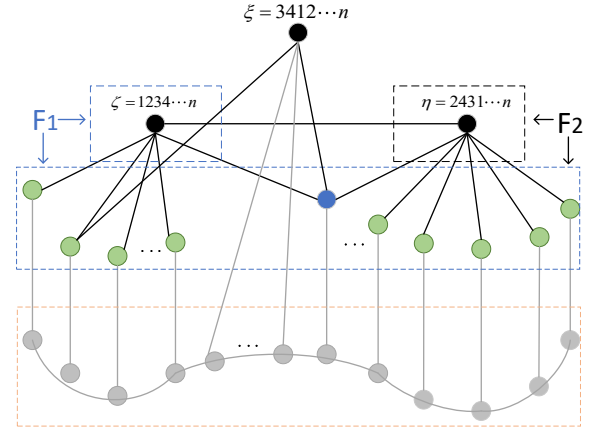


Fig. 4: F_1 and F_2 are 3-component-cuts of S_n^2 .

the PMC model. Hence, S_n^2 is not 3-component $(6n - 13)$ -diagnosable under the PMC model. Based on Definition 1, we have $ct_r(S_n^2) \leq 6n - 14$ for $r = 3$ $n \geq 5$.

Next, we aim to demonstrate that $ct_3(S_n^2) \geq 6n - 14$ for $n \geq 5$. Based on Definition 1, it suffices to show that S_n^2 is 3-component $(6n - 14)$ -diagnosable under the PMC model. We proceed by contradiction. Assume that there exist two distinct 3-component cuts F_1 and F_2 such that $|F_1| \leq 6n - 14$, $|F_2| \leq 6n - 14$, but F_1 and F_2 are indistinguishable under the PMC model. Without loss of generality, assume that $F_1 - F_2 \neq \emptyset$.

Since $|F_1|, |F_2| \leq 6n - 14$, by Theorem 1, the graph $S_n^2 - F_1$ contains one large component C_1 and at least two smaller components, each of size at most 3. Let $U_1 = V(S_n^2 - F_1 - C_1)$. Then $nc(U_1) \geq 2$, $|U_1| \leq 3$, and

$$\begin{aligned} |V(C_1)| &= |V(S_n^2)| - |F_1| - |U_1| \\ &\geq n! - (6n - 14) - 3 \\ &= n! - 6n + 11. \end{aligned}$$

Moreover,

$$\begin{aligned} |V(C_1) - F_2| &\geq |V(C_1)| - |F_2| \\ &\geq (n! - 6n + 11) - (6n - 14) \\ &= n! - 12n + 25. \end{aligned}$$

In particular, $V(C_1) - F_2 \neq \emptyset$.

Suppose that $F_2 - F_1 \not\subseteq U_1$. Then there exists a vertex $y \in (F_2 - F_1) \cap V(C_1)$. Since C_1 is connected and $V(C_1) - F_2 \neq \emptyset$, there exists an edge $yz \in E(S_n^2)$ such that $z \in V(C_1) - F_2 \subseteq V(S_n^2 - (F_1 \cup F_2))$. Hence, $E[F_1 \triangle F_2, V(S_n^2 - (F_1 \cup F_2))] \neq \emptyset$, which contradicts Lemma 1. Therefore, $F_2 - F_1 \subseteq U_1$, and consequently, $|F_2 - F_1| \leq |U_1| \leq 3$.

Since $E[U_1, C_1] = \emptyset$, it follows that $E[F_2 - F_1, C_1] = \emptyset$. Moreover, because F_1 and F_2 are indistinguishable under the PMC model, we also have $E[F_1 - F_2, C_1] = \emptyset$. Hence, C_1 is a component of both $S_n^2 - F_2$ and $S_n^2 - (F_1 \cap F_2)$.

Since F_2 is also a 3-component cut of S_n^2 , the graph $S_n^2 - F_2$ contains the large component C_1 and at least two other components. Let $U_2 = V(S_n^2 - F_2 - C_1)$. Then $nc(U_2) \geq 2$, $|U_2| \geq 2$. By symmetry, we also obtain $F_1 - F_2 \subseteq U_2$.

Furthermore, $|F_1 \cap F_2| = |F_1| - |F_1 - F_2| \leq 6n - 15$. Hence, by Lemma 7 and Theorem 1, the graph $S_n^2 - (F_1 \cap F_2)$ has the large component C_1 and at most two smaller components

whose total size is at most 3.

If $U_1 - (F_2 - F_1) \neq \emptyset$, then $nc(U_1 \cup (F_1 - F_2)) \geq 3$. This implies that $S_n^2 - (F_1 \cap F_2)$ has a large component C_1 and at least three small components. On the other hand, $|F_1 \cap F_2| = |F_1| - |F_1 - F_2| \leq 6n - 15$. By Lemma 7 and Theorem 1, $S_n^2 - (F_1 \cap F_2)$ contains a large component, C_1 , and no more than two smaller components, which results in a contradiction.

If $U_1 - (F_2 - F_1) = \emptyset$, then $|F_1 - F_2| \geq 2$. This suggests that $S_n^2 - (F_1 \cap F_2)$ includes a large component, C_1 , along with a union of smaller components whose total size is at least 4. $|F_1 \cap F_2| = |F_1| - |F_1 - F_2| \leq 6n - 15$. According to Lemma 7 and Theorem 1, $S_n^2 - (F_1 \cap F_2)$ consists of a large component, C_1 , and a union of smaller components whose total size is at most 3, which leads to a contradiction. ■

V. A NOVEL COMPONENT FAULT DIAGNOSIS ALGORITHM IN SPLIT-STAR NETWORK INTERCONNECTION NETWORKS

A. Component Diagnosis Algorithm

This section presents a novel algorithm, NCFDSSN, designed for r -component t -diagnosability in the split-star network S_n^2 . Test outcomes are first utilized to construct an “image graph” by removing edges whose results are suspicious (i.e., outcome 1), so that connectivity among fault-free nodes is preserved while connections around faulty nodes are cut. Within this modified graph, the unique largest connected component is identified and denoted as C_1 , which serves as the fault-free core. All nodes in C_1 are declared healthy and added to the set $IS_{\text{fault-free}}$. For any node outside C_1 , if it can be reliably tested by a node in C_1 , its status is determined by the corresponding test outcome; otherwise, if it is completely disconnected from C_1 , it is directly classified as faulty and added to IS_{faulty} . By leveraging the diagnosability bound of S_n^2 , this procedure ensures effective and correct differentiation between faulty and fault-free nodes (see Algorithm 1).

B. Complexity Analysis

We analyze the time complexity of the proposed algorithms, assuming the n -dimensional split-star network S_n^2 has n vertices.

Step 1: Both the faulty node set and the fault-free node set are initialized as empty sets, with a time complexity of $\mathcal{O}(1)$.

Step 2: The algorithm ImageGenerate, with a time complexity of $\mathcal{O}(N)$.

Steps 3-5: A depth-first traversal is performed on the graph, with a time complexity of $\mathcal{O}(N + E)$, E is the edges of the network.

Step 6: The components are sorted based on their sizes, with a time complexity of $\mathcal{O}(1)$.

Steps 7-8: The largest branch is classified as the fault-free node set, with a time complexity of $\mathcal{O}(1)$.

Steps 9-18: Nodes within the smaller components are tested to identify faulty and fault-free nodes, with a time complexity of $\mathcal{O}(1)$.

Overall time complexity: The total time complexity of the algorithm is $\mathcal{O}(N + E)$. To evaluate the efficiency of our proposed NCFDSSN algorithm, we compare its time complexity with several representative fault diagnosis methods based

Algorithm 1 Novel Component Fault Diagnosis for Split-Star Network Interconnection Network (for short, NCFDSSN)

Input: An n -dimensional split-star network S_n^2 ; A t -fault set with $t \leq ct_g(S_n^2)$;

Output: An identified faulty node-set IS_{faulty} ; An identified fault-free node-set $IS_{\text{fault-free}}$.

- 1: Let IS_{faulty} and $IS_{\text{fault-free}}$ be the empty sets;
 - 2: Executing ImageGenerate() (see Algorithm 2);
 - 3: **for** each $i \in V_{S_n^2}$ **do**
 - 4: **if** !visited[i] **then**
 - 5: Find all components through Deep First Search algorithm $DFS(i)$;
 - 6: Count the total amount of nodes in each component C_i and sort the components from large to small;
 - 7: Set the C_1 as the largest component and others as small components;
 - 8: Add all nodes in $V(C_1)$ into the set $IS_{\text{fault-free}}$;
 - 9: **for** node v_i in all small components **do**
 - 10: **if** node v_i connected with nodes $V(C_1)$ in largest component **then**
 - 11: Some node in $V(C_1)$ can test v_i under PMC model;
 - 12: **if** the test result is 1 **then**
 - 13: v_i is faulty and add it into the set IS_{faulty} ;
 - 14: **else**
 - 15: v_i is fault-free and add it into the set $IS_{\text{fault-free}}$;
 - 16: **else**
 - 17: v_i is faulty and add it into the set IS_{faulty} ;
 - 18: **return** $IS_{\text{faulty}}, IS_{\text{fault-free}}$.
-

Algorithm 2 Optimized Construction of the Image Graph

Input: An n -dimensional split-star network S_n^2 ; A t -fault set with $t \leq ct_g(S_n^2)$ for $g = 2, 3$.

Output: $V_{S_n^2}$ after executing the PMC model.

- 1: **for** each node $v_i \in V_{S_n^2}$ **do**
 - 2: **for** each neighbor v_j of v_i **with** $i < j$ **do**
 - 3: v_i tests v_j ;
 - 4: **if** v_i is fault-free **and** v_j is fault-free **then**
 - 5: Test result $\leftarrow 0$;
 - 6: **else if** v_i is fault-free **and** v_j is faulty **then**
 - 7: Test result $\leftarrow 1$;
 - 8: **else if** v_i is faulty **then**
 - 9: Test result $\leftarrow 0$ or 1 (uncertain);
 - 10: **if** Test result is 1 **then**
 - 11: Remove edge between v_i and v_j in $V_{S_n^2}$;
 - 12: **return** $V_{S_n^2}$.
-

on the PMC model. The time complexity of the NCFDSSN algorithm is $\mathcal{O}(N + E)$, where N is the total number of nodes. To comprehensively evaluate the performance of our proposed NCFDSSN algorithm, we compare it with five representative fault diagnosis methods in Table III and Fig. 5. Huang et al. [33] proposed an algorithm GMISKDIAGMM* for the interconnection network with a time complexity of $\mathcal{O}(Nr^2)$, where N denotes the number of nodes in the network and r represents the degree of network G . Although the algorithm

TABLE III: Summary of different time complexity of fault diagnosis algorithm.

Algorithm	Time Complexity	Model	Network Type
GMISKDI-AGMM* [33]	$\mathcal{O}(Nr^2)$	MM*	Interconnection Networks
Ftk-DIA G-MM* [34]	$\mathcal{O}(N \log_2 N)^2$	MM*	Folded Hypercube
FIFPD PMC [13]	$\mathcal{O}(Ndx)$	MM*	Split Star Network
FHCD-ADM U-CM [35]	$\mathcal{O}(N + 3 \log_2(N)/2)$	Comparison Model	Human Connection Network
FCFDSn [14]	$\mathcal{O}(N^2)$	PMC	Star Network
NCFDSSN (Ours)	$\mathcal{O}(N + E)$	PMC	Split Star Network

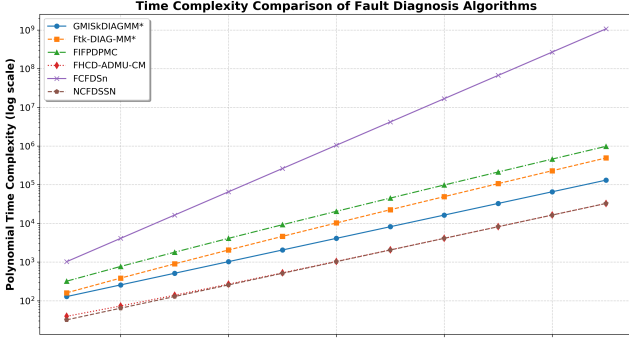


Fig. 5: Comparative analysis of time complexities for fault diagnosis algorithms.

avoids a quadratic dependency on N , it can still incur significant computational costs in dense graphs due to the increased number of connections.

Lin et al. [34] developed an adaptive system-level fault self-diagnosis algorithm Ftk-DIAG-MM* for the FQ_n network with a time complexity of $\mathcal{O}(N \log_2 N)^2$, where N is the total number of nodes. Although this method offers relatively low time complexity in theory, its computational overhead can increase substantially when connectivity among network devices becomes complex.

Song et al. [13] proposed the FIFPDPMC algorithm for split-star networks, which has a time complexity of $\mathcal{O}(Ndx)$, where x denotes the number of diagnostic rounds. Although the iterative process improves fault tolerance, the cumulative cost increases as x increases.

Another algorithm, FHCD-ADMU-CM, was introduced by Lin et al. [35] for detecting malicious nodes in HCN networks. Its time complexity is $\mathcal{O}(N + 3 \log_2(N)/2)$, where N indicates the number of nodes. This algorithm is suitable for human contact networks with relatively regular structures; however, when the network becomes more complex, the computational cost can rise significantly.

Additionally, Wan et al. [14] presented the FCFDSn algorithm, which adopts a component-based diagnosis strategy with a time complexity of $\mathcal{O}(N^2)$. Although the method is effective, it may incur high computational costs in dense or large-scale environments.

In this section, we examine the effectiveness and efficiency of the 2-component and 3-component diagnosis strategies

through simulation experiments. The simulation experiments were implemented in Python and conducted on a computing platform equipped with an Intel Core i5-13500HX CPU and an NVIDIA RTX 4050 GPU with 16 GB memory. Initially, we evaluate the performance of the NCFDSSN algorithm using synthetic data. From the theoretical analysis, it follows that all faulty nodes can be detected simultaneously when the number of faulty nodes t meets the conditions $t \leq ct_2(S_n^2) = 4n - 9$ for 2-component diagnosis, and $t \leq ct_3(S_n^2) = 6n - 14$ for 3-component diagnosis in the n -dimensional split-star network S_n^2 . Next, we compute several detection metrics for $t \geq 4n - 9$ and $t \geq 6n - 14$ to conduct a more in-depth performance analysis. For example, we run the program with the number of faulty nodes ranging from 10 to 70 for S_5^2 . Finally, we apply the NCFDSSN algorithm to a real-world network.

C. Evaluation Metrics

In this subsection, we begin by presenting several metrics to assess the performance of the NCFDSSN algorithm for r -component t -diagnosability.

Define TPN (resp., TNN) as the count of faulty (resp., fault-free) nodes correctly classified. Meanwhile, FPN (resp., FNN) represents the count of fault-free (resp., faulty) nodes that are misclassified.

Accuracy: It is calculated as the proportion of correctly diagnosed nodes out of all nodes. The formula for accuracy is
$$\text{Accuracy} = \frac{\text{TPN} + \text{TNN}}{\text{TPN} + \text{TNN} + \text{FPN} + \text{FNN}}.$$

Precision: It refers to the proportion of correctly identified faulty nodes among all nodes diagnosed as faulty. Precision is defined by
$$\text{Precision} = \frac{\text{TPN}}{\text{TPN} + \text{FPN}}.$$

True Positive Rate (TPR): It measures the proportion of correctly diagnosed faulty nodes out of all actual faulty nodes. TPR is given by
$$\text{TPR} = \frac{\text{TPN}}{\text{TPN} + \text{FNN}}.$$

F1-score: It denotes the harmonic average of precision and TPR. The expression for F1-score is
$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{TPR}}{\text{Precision} + \text{TPR}}.$$

G-Mean: It represents the geometric mean of precision and TPR. The G-Mean is defined as
$$G\text{-Mean} = \sqrt{\text{TPR} \times \text{Precision}}.$$

False Positive Rate (FPR): It refers to the fraction of fault-free nodes that are mistakenly identified as faulty among all fault-free nodes. The FPR is computed as
$$\text{FPR} = \frac{\text{FPN}}{\text{FPN} + \text{TNN}}.$$

True Negative Rate (TNR): It represents the ratio of fault-free nodes that are correctly identified as fault-free to the total number of fault-free nodes. The TNR is defined as
$$\text{TNR} = \frac{\text{TNN}}{\text{FPN} + \text{TNN}}.$$

False Negative Rate (FNR): It represents the fraction of fault nodes that are mistakenly identified as fault-free among all faulty nodes. The FNR is expressed as
$$\text{FNR} = \frac{\text{FNN}}{\text{FNN} + \text{TPN}}.$$

D. Applications to Abstract Data

We execute the program with the number of faulty nodes varying from 10 to 70, performing experiments every 5 nodes. We run the experiment 1,000 times to compute the average values of the detection indicators in S_5^2 . The 5-dimensional split-star network contains 120 nodes and 840 edges. The results of the experiment are presented in Table IV.

TABLE IV: Summary of evaluation metrics under different number of fault nodes in S_n^2 .

No. of faults	10	15	20	25	30	35	40	45	50	55	60	65	70
Accuracy	1	1	1	1	1	0.9994	0.9991	0.9985	0.9968	0.9958	0.9887	0.9829	0.9597
Precision	1	1	1	1	1	0.9982	0.9982	0.9975	0.9949	0.9952	0.9874	0.9843	0.9706
$F1$ -score	1	1	1	1	1	0.999	0.9987	0.998	0.9962	0.9955	0.9887	0.9843	0.9653
G -Mean	1	1	1	1	1	0.9991	0.9987	0.998	0.9962	0.9955	0.9887	0.9843	0.9653
FPR	0	0	0	0	0	0.0008	0.0009	0.0015	0.0037	0.0041	0.0127	0.0186	0.0407
TNR	1	1	1	1	1	0.9992	0.9991	0.9985	0.9963	0.9959	0.9873	0.9814	0.9593
FNR	0	0	0	0	0	0	0.0008	0.0014	0.0025	0.0043	0.01	0.0157	0.04
TPR	1	1	1	1	1	1	0.9992	0.9986	0.9975	0.9957	0.99	0.9843	0.96

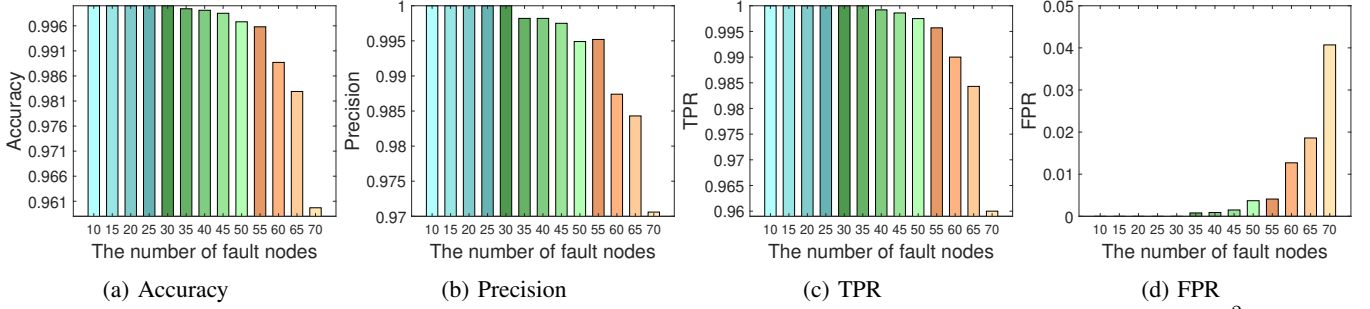
Fig. 6: Comparisons of accuracy, precision, TPR, FPR under the different numbers of fault nodes in S_5^2 .

TABLE V: Summary of evaluation metrics under different number of fault nodes in the real-world datasets.

No. of faults	10	15	20	25	30	35	40	45	50	55	60	65	70
Accuracy	0.9993	0.9991	0.9988	0.9985	0.9984	0.9984	0.9981	0.9995	0.9972	0.9968	0.9865	0.9864	0.9556
Precision	0.9058	0.9176	0.9154	0.9185	0.9232	0.9219	0.9127	0.9126	0.9197	0.9189	0.9186	0.9208	0.9116
$F1$ -score	0.9502	0.9567	0.9553	0.9568	0.9593	0.9583	0.9533	0.9529	0.9568	0.9963	0.9559	0.9572	0.9519
G -Mean	0.9513	0.9576	0.9563	0.9576	0.9601	0.9591	0.9543	0.9539	0.9576	0.957	0.9566	0.958	0.9528
FPR	0.0006	0.0008	0.0011	0.0013	0.0015	0.0018	0.0024	0.0027	0.0027	0.0031	0.0034	0.0036	0.0043
TNR	0.9994	0.9992	0.9989	0.9987	0.9985	0.9982	0.9976	0.9973	0.9973	0.9969	0.9866	0.9964	0.9957
FNR	0.0009	0.0007	0.0011	0.0016	0.0016	0.0023	0.0023	0.0013	0.0030	0.0032	0.0038	0.0039	0.0041
TPR	0.9991	0.9993	0.9989	0.9984	0.9984	0.9977	0.9977	0.9977	0.9970	0.9968	0.9962	0.9961	0.9959

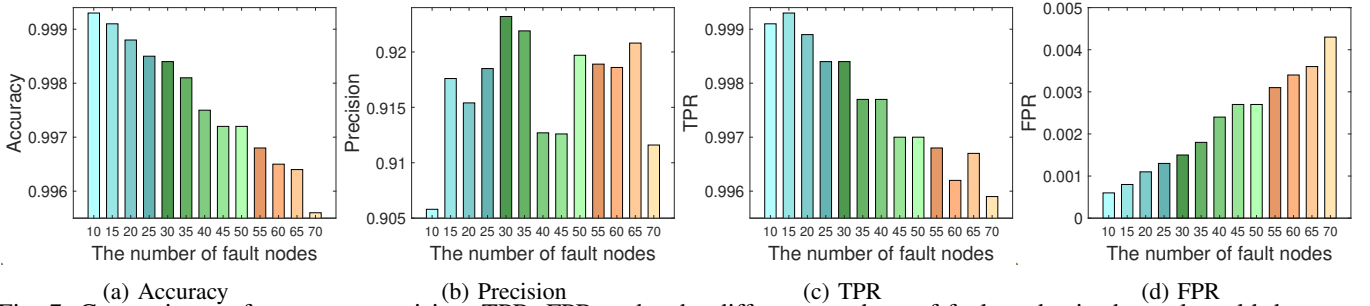


Fig. 7: Comparisons of accuracy, precision, TPR, FPR under the different numbers of fault nodes in the real-world datasets.

As shown in Fig. 6a, the accuracy values remain at 1 when the number of faulty nodes is between 10 and 30. This suggests that our ways can accurately diagnose faulty nodes with a high degree of precision. As the number of faulty nodes increases, the accuracy values gradually decrease. Even when the number of faulty nodes reaches 70, the accuracy remains above 0.95. This indicates that our algorithm is still capable of diagnosing all faulty nodes with high precision, even when the number of faulty nodes exceeds the range initially validated.

From Fig. 6b, the precision values are all 1 when the number of faulty nodes is within the range of 10-30. This suggests that our ways can effectively diagnose faulty nodes with a very high level of accuracy. As the number of faulty nodes increases, the precision values gradually decrease. However, even when the number of faulty nodes reaches 70, the precision metric remains above 0.97. This indicates that our algorithm can still diagnose all faulty nodes with a high level

of precision even when the faulty nodes exceeds the range initially validated.

As seen in Fig. 6c, when the number of faulty nodes falls between 10 and 35, the TPR values remain at 1. When the number of faulty nodes is 40, the TPR value is 0.9992, and as the number of faulty nodes increases, the TPR starts to decrease. When the number of faulty nodes reaches 70, the TPR value is 0.96, indicating that even when significantly exceeding the validated quantity, our algorithm can still diagnose the majority of faulty nodes.

Fig. 6d reveals that when the number of faulty nodes is within the range of 10-30, the FPR value is 0. When the number of faulty nodes is within this range, the count of non-faulty nodes misdiagnosed as faulty is 0, meaning there are few misdiagnosed nodes. As the number of faulty nodes exceeds 30, the FPR value starts to decrease, indicating instances where non-faulty nodes are diagnosed as faulty. This is because, as

the number of faulty nodes surpasses a certain threshold, there is a situation where small components around non-faulty nodes consist mostly of faulty nodes. With an increasing number of faulty nodes, such small components become more prevalent, leading to an increase in misdiagnosis rates.

E. Application to Real-World Datasets

In this subsection, we evaluate the performance of the NCFDSSN algorithm for component fault diagnosis using real-world data, specifically the Internet of Things Network (IOTN), which was sourced from <https://networkrepository.com/inf.php>. First, we preprocess the IOTN dataset and extract the connection patterns of the edges. The processed IOTN dataset consists of 1608 nodes and 27,051 edges, where each node corresponds to an IoT device, and each edge represents a communication link between the devices.

To ensure accuracy, we run the NCFDSSN algorithm for 100,000 iterations and compute the average results to minimize errors. We randomly create various sets of faulty nodes within the IOTN and use the NCFDSSN algorithm to diagnose both faulty and fault-free nodes. The number of faulty nodes is varied from 10 to 70, with experiments conducted at intervals of 5 nodes to facilitate comparison with theoretical data. The experimental outcomes are detailed in Table V.

As shown in Fig. 7a, our accuracy tends to decrease as the number of faulty nodes increases. With 10 faulty nodes, the accuracy of fault node diagnosis is 0.9993, which is nearly 1. When there are 70 faulty nodes, the accuracy is 0.9956, also close to 1. Therefore, our NCFDSSN algorithm performs well on real-world data.

Fig. 7b shows that with an increase in faulty nodes, the precision of our algorithm fluctuates between 0.9058 and 0.9232. Through data analysis, we find that this is due to the large number of nodes in the actual data. When the system's faulty nodes are set to 10, the system may misdiagnose one or two non-faulty nodes as faulty. As the number of faulty nodes increases, the number of non-faulty nodes misdiagnosed as faulty also increases, with both TPN and FPN values increasing. Experimental data shows that when the number of faulty nodes is 70, an average of 6 non-faulty nodes are diagnosed as faulty, accounting for 0.37% of the total experimental nodes. This error is negligible. The precision value fluctuates due to the calculation formula, but from the experimental results, the accuracy of the NCFDSSD algorithm is above 0.9, indicating good performance.

From Fig. 7c, it can be observed that as the number of faulty nodes increases, the TPR decreases. When there are 10 faulty nodes, the TPR is 0.9991, and when there are 70 faulty nodes, the TPR is 0.9959, both close to 1. Experimental results demonstrate that on a real dataset, our algorithm can diagnose almost all faulty nodes.

From Fig. 7d, it can be observed that as the number of faulty nodes increases, the False Positive Rate (FPR) increases. When there are 70 faulty nodes, the FPR is 0.004, very low. This indicates that the proportion of misdiagnosing faulty nodes as non-faulty nodes is very low, demonstrating the effectiveness of our algorithm.

F. Comparison of Experimental Results Between Abstract Data and Real Data

From Fig. 8a, it can be observed that with an increase in the number of faulty nodes, the accuracy decreases in both the abstract dataset and the real dataset. When the number of faulty nodes is small, the accuracy in the abstract dataset tends to be slightly better than in the real dataset, with an advantage of 0.007. However, as the number of faulty nodes increases, the accuracy in the real dataset surpasses that in the abstract dataset. When there are more faulty nodes, the accuracy in the real dataset is higher by 0.0359 compared to the abstract dataset.

Fig. 8b indicates that as the number of faulty nodes increases, precision decreases in the abstract dataset, while it fluctuates in the real dataset. When there are 10 faulty nodes, precision in the abstract dataset is 0.0942 higher than in the real dataset, showing the maximum difference. When there are 30 faulty nodes, precision in the abstract dataset is 0.0749 higher than in the real dataset, with the minimum difference.

From Fig. 8c, it can be seen that as the number of faulty nodes increases, the TPR decreases in both the real and abstract datasets. When the number of faulty nodes is less than 50, the TPR in the abstract dataset performs better; however, when the number of faulty nodes exceeds 50, the TPR in the real dataset performs better.

Fig. 8d illustrates that as the number of faulty nodes increases, in the abstract dataset, both $F1$ -score and G -Mean values decrease, while in the real dataset, they fluctuate. As analyzed earlier, this is due to the fluctuation in precision. However, $F1$ -score and G -Mean values in the abstract dataset are larger than those in the real dataset. When there are 70 faulty nodes, $F1$ -score in the abstract dataset is 0.0498 higher, and G -Mean is 0.0487 higher compared to the real dataset.

From Fig. 8e, it can be observed that as the number of faulty nodes increases, the FPR decreases in both the abstract and real datasets. When the number of faulty nodes is less than or equal to 50, the FPR in the abstract dataset is smaller. When there are 10 faulty nodes, the difference between the two is 0.0006. However, when the number of faulty nodes exceeds 50, the FPR in the real dataset becomes smaller. When there are 70 faulty nodes, the difference is 0.0364.

VI. CONCLUSION

This research centers on the component fault diagnosis of S_n^2 . It was demonstrated that removing any subset of vertices D with $|D| \leq 4n - 9$ from S_n^2 leaves the network with at most 2 components, and removing any subset D with $|D| \leq 6n - 14$ results in at most 4 components. Under the PMC model, the component fault diagnosabilities of S_n^2 were established as $ct_2(S_n^2) = 4n - 9$ and $ct_3(S_n^2) = 6n - 14$ for $n \geq 5$. Furthermore, a new algorithm, NCFDSSN, for r -component t -diagnosability was introduced to identify all nodes in S_n^2 . The algorithm was evaluated using both synthetic and real data. Its performance and efficiency were assessed through various metrics, including Accuracy, Precision, $F1$, G -Mean, FPR, and TPR. The results indicated that as the split-star network S_n^2 and the Internet of Things Network (IOTN) expand, the

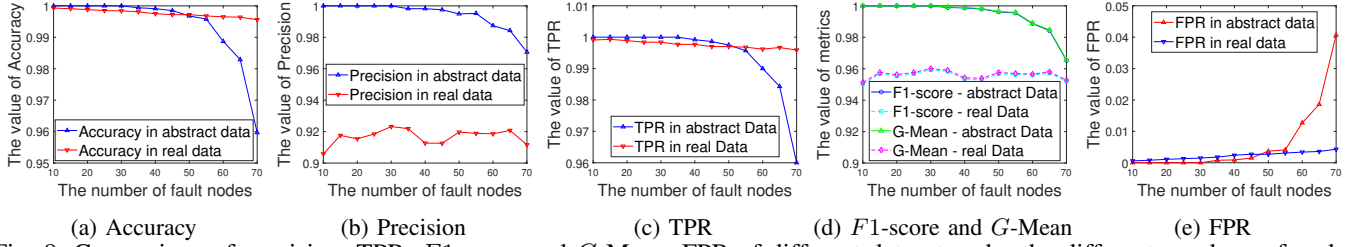


Fig. 8: Comparison of precision, TPR, $F1$ -score and G -Mean, FPR of different dataset under the different numbers of nodes.

NCFDSSN algorithm maintains high performance in terms of these detection metrics, accurately diagnosing most faulty and fault-free nodes. Looking ahead, we intend to explore other types of fault diagnosis in broader real-world networks under specific diagnostic models.

ACKNOWLEDGMENT

The authors declare that there is no conflict of interest regarding the publication of this paper. This work was supported in part by the Natural Science Foundation of Fujian Province under Grants 2024J09032, 2025J01379.

REFERENCES

- [1] X. Liu, J. Fan, B. Cheng, Y. Wang, B. Yin, and X. Jia, "A family of general architectures toward interconnection networks and data center networks," *IEEE/ACM Transactions on Networking*, vol. 32, no. 5, pp. 4099–4113, 2024.
- [2] G. Wang, J. Yu, Y. Zou, J. Fan, and W. Cheng, "A new measure of fault-tolerance for network reliability: Double-structure connectivity," *IEEE/ACM Transactions on Networking*, vol. 32, no. 1, pp. 874–889, 2023.
- [3] L. Mei, G. Chen, Q. Liu, and L. Ming, "The cyclic diagnosability of star graphs under the pmc and mm* models," *Discrete Applied Mathematics*, vol. 364, pp. 60–73, 2025.
- [4] Q. Zhang, S. Zhou, and S.-Y. Hsieh, "A probabilistic approach for local diagnosis in large multiprocessor systems," *IEEE Transactions on Networking*, vol. 33, no. 3, pp. 1086–1096, 2025.
- [5] N.-W. Chang and S.-Y. Hsieh, "Conditional diagnosability of enhanced hypercubes under the pmc model," *IEEE Transactions on Networking*, vol. 33, no. 4, pp. 1603–1613, 2025.
- [6] W. Zheng, S. Zhou, E. Cheng, and Q. Zhang, "The cyclic diagnosability of Cayley graphs generated by transposition trees," *Discrete Applied Mathematics*, vol. 372, pp. 109–123, 2025.
- [7] L. Lin, K. Guan, Y. Huang, S.-Y. Hsieh, and G. Chen, "Local fault diagnosis analysis based on block pattern of regular diagnosable networks," *IEEE Transactions on Networking*, vol. 33, no. 2, pp. 849–864, 2024.
- [8] L. Lin, Y. Huang, Y. Lin, S.-Y. Hsieh, and L. Xu, "FFNLFD: Fault diagnosis of multiprocessor systems at local node with fault-free neighbors under PMC model and MM* model," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 7, pp. 1739–1751, 2021.
- [9] S. Zhang, D. Liang, L. Chen, R. Li, and W. Yang, "The component diagnosability of hypercubes with large-scale faulty nodes," *The Computer Journal*, vol. 65, no. 5, pp. 1129–1143, 2022.
- [10] H. Zhang, S. Zhou, E. Cheng, and S.-Y. Hsieh, "Characterization of component diagnosability of regular networks," *Discrete Applied Mathematics*, vol. 322, pp. 253–267, 2022.
- [11] P. Han, Y. Xiao, C. Ye, H. Li et al., "Fractional strong matching preclusion of split-star networks," *Journal of Mathematics Research*, vol. 11, no. 4, pp. 32–42, 2019.
- [12] S. Zhao and W. Yang, "Conditional connectivity of folded hypercubes," *Discrete Applied Mathematics*, vol. 257, pp. 388–392, 2019.
- [13] J. Song, L. Lin, Y. Huang, and S.-Y. Hsieh, "Intermittent fault diagnosis of split-star networks and its applications," *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 4, pp. 1253–1264, 2023.
- [14] Z. Wan, L. Lin, Y. Huang, and S.-Y. Hsieh, "Component diagnosis strategy of star graphs interconnection networks," *IEEE Transactions on Reliability*, vol. 73, no. 4, pp. 1907–1917, 2024.
- [15] Y. Huang, L. Lin, E. Cheng, and L. Xu, "Component fault diagnosis and fault tolerance of alternating group graphs," *The Computer Journal*, vol. 66, no. 5, pp. 1184–1194, 2023.
- [16] Y. Tian and Q. Zhu, " r -component diagnosability of hypercubes under the PMC model," *Theoretical Computer Science*, vol. 933, pp. 114–124, 2022.
- [17] J. Liu, S. Zhou, D. Wang, and H. Zhang, "Component diagnosability in terms of component connectivity of hypercube-based compound networks," *Journal of Parallel and Distributed Computing*, vol. 162, pp. 17–26, 2022.
- [18] X. Sun, J. Fan, B. Cheng, J. Zhou, and Y. Wang, "Relationship between component connectivity and component diagnosability of some regular networks," *The Computer Journal*, vol. 66, no. 8, pp. 2033–2042, 2023.
- [19] H. Zhuang, W. Guo, X.-Y. Li, X. Liu, and C.-K. Lin, "The component connectivity, component diagnosability, and t/k -diagnosability of bicube networks," *Theoretical Computer Science*, vol. 896, pp. 145–157, 2021.
- [20] C.-A. Chen and S.-Y. Hsieh, "Component-composition graphs: (t,k) -diagnosability and its application," *IEEE Transactions on Computers*, vol. 62, no. 6, pp. 1097–1110, 2013.
- [21] M. Chen, D. F. Hsu, and C.-K. Lin, "A new measure for locally t -diagnosable under PMC model," in *Computing and Combinatorics: 27th International Conference, COCOON 2021, Tainan, Taiwan, October 24–26, 2021, Proceedings*, ser. Lecture Notes in Computer Science, vol. 13025, 2021, pp. 306–316.
- [22] L. Lin, L. Xu, S. Zhou, and S.-Y. Hsieh, "The extra, restricted connectivity and conditional diagnosability of split-star networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 533–545, 2015.
- [23] L. Zhao and S. Wang, "Structure connectivity and substructure connectivity of split-star networks," *Discrete Applied Mathematics*, vol. 341, pp. 359–371, 2023.
- [24] J. Li, X. Li, and E. Cheng, "Super spanning connectivity of split-star networks," *Information Processing Letters*, vol. 166, p. 106037, 2021.
- [25] M.-M. Gu, R.-X. Hao, and J.-M. Chang, "Reliability analysis of alternating group graphs and split-stars," *The Computer Journal*, vol. 64, no. 9, pp. 1425–1436, 2021.
- [26] L. Lin, Y. Huang, X. Wang, and L. Xu, "Restricted connectivity and good-neighbor diagnosability of split-star networks," *Theoretical Computer Science*, vol. 824, pp. 81–91, 2020.
- [27] T.-L. Kung, Y.-H. Teng, C.-K. Lin, and Y.-L. Hsu, "Combinatorial analysis of the subsystem reliability of the split-star network," *Information Sciences*, vol. 415, pp. 28–40, 2017.
- [28] Y. Huang, L. Lin, L. Xu, and S.-Y. Hsieh, "Adaptive system-level fault diagnosis of bijective connection networks," *IEEE Transactions on Reliability*, vol. 74, no. 2, pp. 2916–2926, 2024.
- [29] X. Li, S. Zhou, X. Ren, and X. Guo, "Structure and substructure connectivity of alternating group graphs," *Applied Mathematics and Computation*, vol. 391, p. 125639, 2021.
- [30] Y. Lv, C.-K. Lin, D. F. Hsu, and J. Fan, "An algorithm for conditional-fault local diagnosis of multiprocessor systems under the mm* model," *Theoretical Computer Science*, vol. 987, p. 114372, 2024.
- [31] Q. Wang, J. Liu, D. Wang, W. Liu, J. Wu, and G. Li, "A novel conditional diagnostic scheme for hypercube-based multiprocessor systems," *IEEE Transactions on Networking*, vol. 34, pp. 2092–2102, 2025.
- [32] W. Liu, J. Liu, E. Cheng, S.-Y. Hsieh, J. Wu, and G. Li, "A multi-attribute adaptive fault diagnosis framework for star networks," *IEEE Transactions on Computers*, vol. 75, no. 1, pp. 114–126, 2026.
- [33] Y. Huang, L. Lin, and S.-Y. Hsieh, "A fast $f(r, k + 1)/k$ -diagnosis for interconnection networks under mm* model," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 7, pp. 1593–1604, 2021.
- [34] Y. Lin, L. Lin, Y. Huang, L. Xu, and S.-Y. Hsieh, "Endogenous security of FQ_n networks: Adaptive system-level fault self-diagnosis," *IEEE Transactions on Reliability*, vol. 73, no. 3, pp. 1659–1668, 2024.
- [35] L. Lin, Y. Huang, L. Xu, and S.-Y. Hsieh, "Better adaptive malicious users detection algorithm in human contact networks," *IEEE Transactions on Computers*, vol. 71, no. 11, pp. 2968–2981, 2022.



Xiuzhen Zhu received the M.S. degree in Cyberspace Security from Fujian Normal University, Fuzhou, China, in 2025. He is currently working toward the Ph.D. degree in Cyberspace Security with Fujian Normal University, Fuzhou, China. His research interests include network and information security, machine learning, and fault diagnosis.



Yanze Huang (Member, IEEE) received the Ph.D. degree in Mathematics from the Hong Kong University of Science and Technology, Hong Kong, in 2015. He is currently an Associate Professor with the School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou, China. He has published more than 40 papers in the IEEE TPDS/TC/TMC/TON/TR. His current research interests include fault diagnosis, network reliability and graph learning.



RSMIN 2021, CCNFF 2023, ICCNC 2025.

Limei Lin (Senior Member, IEEE) received the Ph.D. degree in Applied Mathematics from Fujian Normal University, Fuzhou, China, in 2016. She is a Professor with the Fujian Normal University, and associate editors of two SCI indexed Journals. Her research focuses on the graph learning, swarm intelligence, network security, fault tolerant computing. She has published over 80 papers as the first author or corresponding author in WWW, IJCAI, IEEE TMC/TDSC/TPDS/TC/TON. She has served on program committee chair of SocialSec 2021,



Xiaoding Wang (Senior Member, IEEE) is a Professor with the Fujian Normal University. He was a Global Top 2% Scientist from 2024 to 2025, specializes in data security and intelligent decision-making. He holds key roles in academic organizations (CCF, CIPS) and journals (Computer Engineering), while organizing international conferences like ICCNC 2025. With over 100 publications (more than 40 in CAS Q1/CCF A journals, e.g., IEEE/ACM Transactions) and 20 patents.



Sun-Yuan Hsieh (Fellow, IEEE) received the Ph.D. degree in Computer Science from National Taiwan University, Taipei, Taiwan, in June 1998. He is currently a Chair Professor with the National Cheng Kung University. His awards include the 2020 ACM Distinguished Scientist, 2019 Kwoh Ting Li Honorable Scholar Award. He is Fellow of the IEEE, the BCS, and the IET. He is an experienced editor with editorial services of a number of journals, including serving as and associate editor for the IEEE Transactions on Computers, Theoretical Computer Science,

Discrete Applied Mathematics, The Journal of Supercomputing, International Journal of Computer Mathematics, Fundamental Informaticae, Journal of Interconnection Networks, and the Discrete Mathematics Algorithms and Applications. His research interests include design and analysis of algorithms, fault-tolerant computing, bioinformatics, parallel and distributed computing, and algorithmic graph theory. He was on the organization committee and/or program committee of several dozens international conferences in computer science and computer engineering.



Jie Wu (Fellow, IEEE) is Chief Scientist of China Telecom and Director of China Telecom Cloud Computing Research Institute. Before he joined China Telecom, he was Laura H. Carnell Professor at Temple University and the Director of the Center for Networked Computing (CNC). His current research interests include mobile computing and wireless networks, routing protocols, network trust and security, distributed algorithms, applied machine learning, and cloud computing. He serves on several editorial boards, including IEEE/ACM Transactions on Networking. Dr. Wu is/was general chair/co-chair for IEEE IPDPS'23, ACM MobiHoc'23, and IEEE CCGrid'24 as well as program chair/cochair for IEEE INFOCOM'11 and CCF CNCC'13. He was an IEEE Computer Society Distinguished Visitor, ACM Distinguished Speaker, and chair for the IEEE Technical Committee on Distributed Processing (TCDP). Dr. Wu is a Fellow of the AAAS and a Fellow of the IEEE. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award. He is a Member of the Academia Europaea.