Vertex-Connectivity for Node Failure Identification in Boolean Network Tomography*

Nicola Galesi[1], Fariba Ranjbar[1], and Michele Zito[2]

[1] Universitá La Sapienza, Roma, Italy
[2] University of Liverpool, United Kingdom

Abstract. In this paper we study the node failure identification problem in undirected graphs by means of Boolean Network Tomography. We argue that vertex connectivity plays a central role. We show tight bounds on the maximal identifiability in a particular class of graphs, the Line of Sight networks. We prove slightly weaker bounds on arbitrary networks. Finally we initiate the study of maximal identifiability in random networks. We focus on two models: the classical Erdős-Rényi model, and that of Random Regular graphs. The framework proposed in the paper allows a probabilistic analysis of the identifiability in random networks giving a tradeoff between the number of monitors to place and the maximal identifiability.

1 Introduction

A central issue in communication networks is to ensure that the structure works reliably. To this end it is of the utmost importance to discover as quickly as possible those components that develop some sort of failure. Network Tomography is a family of distributed failure detection algorithms based on the spreading of end-to-end measurements [8, 24] rather than directly measuring individual network components. Typically a network $G = (V, E)$ is given as a graph along with a collection of paths $\mathbb{P}$ in it and the goal is to take measurements along such paths to infer properties of the given network. Quoting from [12] "A key advantage of tomographic methods is that they require no participation from network elements other than the usual forwarding of packets. This distinguishes them from well-known tools such as `traceroute` and `ping`, that require ICMP responses to function. In some networks, ICMP response has been restricted by administrators, presumably to prevent probing from external sources. Another feature of tomography is that probing and the recovery of probe data may be embedded within transport protocols, thus co-opting suitably enabled hosts to form

* The first two authors kindly acknowledge the partial support by the MIUR under the grant Dipartimenti di eccellenza 2018-2022 of the Department of Computer Science of Sapienza University. The research was also partly supported by a visiting fellowship of the University of Liverpool and the Networks Sciences & Technologies (NeST) initiative of the University of Liverpool (`https://www.liverpool.ac.uk/network-science-technologies/`).

impromptu measurement infrastructures". The approach is strongly related to group testing [11] where, in general, one is interested in making statements about individuals in a population by taking group measurements. The main concern is to do so with the minimum number of tests. In our setting, the connectivity structure of the network constrains the set of feasible tests. Graph-constrained group testing has been studied before, starting with [7]. We are interested in using structural graph-theoretic properties to make statements about the quality of the testing process.

Research in Network Tomography is vast. The seminal works of Vardi [24], and Coates *et al.* [8], or more recent surveys like [6] each have more that 500 citations, according to Google Scholar. Methods and algorithms vary dramatically depending on the network property of interest, or the measurements one has to rely on. Boolean Network Tomography (BNT) aims to identify corrupted components in a network using boolean measurements (i.e. assuming that elementary network components can be in one of two states: "working" or "not-working"). Introduced in [12], the paradigm has recently attracted a lot of interest [14, 20] because of its simplicity. In this work we use BNT to identify failing nodes. Assume to have a set $\mathbb{P}$ of measurement paths over a node set V. We would like to know the state x_v (with $x_v = 0$ corresponding to "v in working order" and $x_v = 1$ corresponding to "v in a faulty state") of each node $v \in V$. The localization of the failing nodes in $\mathbb{P}$ is captured by the solutions of the system:

$$\bigwedge_{p \in \mathbb{P}} \left(\bigvee_{v \in p} x_v \equiv b_p \right) \tag{1}$$

where b_p models the (boolean) state of the path $p \in \mathbb{P}$. Of course, systems of this form may have several solutions and therefore, in general, the availability of a collection of end-to-end measurements does not necessarily lead to the unique identification of the failing nodes. We will investigate properties of the underlying network that facilitate the solution of this problem. In particular, we follow the approach initiated by Ma et. al. [19] based on the notion of *maximal identifiability* (see Section 2 for a precise definition). The metric aims to capture the maximal number of simultaneously failing nodes that can be uniquely identified in a network by means of measurement along a given path system. It turns out that the network maximal identifiability is an interesting combinatorial measure and several studies [2, 15, 19, 21] have investigated variants of this measure in connection with various types of path systems. However, it seems difficult to come up with simple graph-theoretic properties that affect the given network identifiability. We contend that the maximal identifiability using measures over the collection of all simple paths between two disjoint sets of vertices S and T enables us to make good progress on this issue. More specifically we show that the proposed approach provides an almost tight characterization of the maximal identifiability in *augmented hypergrids* (see definition in Section 2) and more general Line-of-Sight (LoS) networks. LoS networks were introduced by Frieze *et al.* in [13] and have been widely studied (see for instance [10, 9, 22, 23]) as models for communication patterns in a geometric environment containing obstacles. Like grids, LoS networks can be embedded in a finite cube of $\mathbb{Z}^d$, for some positive integer d. But LoS networks generalize grids in that edges are allowed between nodes that are not necessarily next to each other in the network embedding.

Using the network vertex-connectivity, $\kappa(G)$, (i.e. the size of the minimal set of nodes disconnecting the graph) we are able to prove the following:

Theorem 1. *Let $\mathcal{H}$ be an augmented hypergrid. For every pair of disjoint $S, T \subseteq V(\mathcal{H})$, the maximal identifiability of $\mathcal{H}$, $\mu(\mathcal{H})$ using measures over simple paths between S and T satisfies: $\mu(\mathcal{H}) \leq \kappa(\mathcal{H})$. Furthermore, there is a way to choose S and T that guarantees $\mu(\mathcal{H}) \geq \kappa(\mathcal{H}) - 1$.*

The result on hypergrids immediately suggests the related question about general graphs. In this work we prove upper and lower bounds on the maximal identifiability of any network G. The following statement summarizes our findings (here $\kappa_{ST}(G)$ is the size of smallest set of vertices separating S and T):

Theorem 2. *Let $G = (V, E)$ be an arbitrary graph. For every pair of disjoint $S, T \subseteq V(G)$, the maximal identifiability of G, $\mu(G)$ using measures over simple paths between S and T satisfies: $\mu(G) \leq \min(\delta(G), \kappa_{ST}(G))$. Furthermore, there is a way to choose S and T that guarantees $\mu(G) \geq \lfloor \kappa(G)/2 \rfloor - 1$.*

In both results, the upper bound is proved by showing that there are sets of $\kappa(G) + 1$ vertices that cannot be identified. The lower bounds which require the construction of paths separating large sets of nodes in the graph, are based on a well-known relationship between $\kappa(G)$ and the existence of collections of vertex-disjoint paths between certain sets of nodes in G. In fact a much higher lower bound can be proved for graphs with low connectivity. The following result applies to arbitrary LoS networks, and to many topologies studied in relation to communication problems including various types of grids, butterflies, hypercubes, and sparsely connected sensor networks.

Theorem 3. *Let $G = (V, E)$ be an arbitrary network with $\kappa(G) \leq |V|/3$. Let $\mu(G)$ denote the maximal identifiability of G using measures over simple paths between two disjoint sets of vertices S and T.*

1. *For all pairs of disjoint $S, T \subseteq V$, $\mu(G) \leq \kappa(G)$.*
2. *There is a pair of disjoint $S, T \subseteq V(G)$ such that $\mu(G) \geq \kappa(G) - 2$.*

Finally, we look at random networks (Erdős-Rényi and Random Regular Graphs). In these structures we are able to show a trade-off between the success probability of the relevant path construction processes and the size of the sets S and T defining the path set $\mathbb{P}$. Random graphs also give us alternative constructions of networks with large identifiability.

The rest of the paper is organized as follows. After a section devoted to preliminaries and important definitions, we have a section that focuses on Theorem 1. Section 4 focuses on arbitrary graphs. First we look at the proof of Theorem 2. Then describe a different construction that leads to the proof of Theorem 3. Finally Section 5 is dedicated to the analysis of the maximal identifiability in random graphs. First we look at Erdős-Rényi graphs, then random regular graphs.

2 Preliminaries

Sets, Graphs, Paths, and Connectivity. If U and W are sets, $U \triangle W = (U \setminus W) \cup (W \setminus U)$ is the symmetric difference between U and W. Graphs (we will use the terms network and graph interchangeably) in this paper will be undirected, simple and loop-less. A path (of length k) in a graph $G = (V, E)$ from a node u to a node v is a sequence of nodes $p = u_1, u_2, \ldots, u_{k+1}$ such that $u_1 = u$, $u_{k+1} = v$ and $\{u_i u_{i+1}\} \in E$ for all $i \in [k]$. The path p is *simple* of no two u_i and u_j in p are the same. Any sub-sequence $u_x, \ldots, u_{x+y}$ ($x \in \{1, \ldots, k+1\}$, $y \in \{0, \ldots, k+1-x\}$) is said to be *contained in* p, and dually we say that p *contains* the sequence or *passes through* it. We say that path p and q *intersect* if they contain a common sub-sequence. The intersection of a path p and an arbitrary set of nodes W is the set of elements of W that are contained in p. When p intersect W sometimes we say that p *touches* W. For an arbitrary $U \subseteq V(G)$, $N(U)$ is the set of *neighbours* of $u \in U$. If $U = \{u\}$ we write $N(u)$ instead of $N(\{u\})$. The *degree* of u, $\deg(u)$, is the cardinality of $N(u)$, and let $\delta(G) = \min_{u \in V} \deg(u)$ be the minimum degree of G.

In what follows $\kappa(G)$ denotes the vertex-connectivity of the given graph $G = (V, E)$, namely $\kappa(G)$ is the size of the minimal subset K of V, such that removing K from G disconnects G. In particular it is well-known (see for example [16], Theorem 5.1, pag 43) that

$$\kappa(G) \leq \delta(G). \tag{2}$$

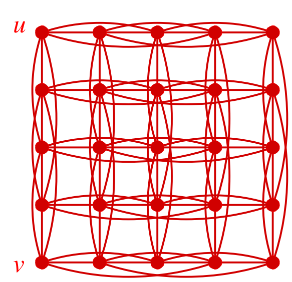
It will also be convenient to work with sets of vertices disconnecting particular parts of G. If $S, T \subseteq V$, then $\kappa_{ST}(G)$ is the size of the smallest vertex separator of S and T in G, i.e. the smallest set of vertices whose removal disconnects S and T (set $\kappa_{ST}(G) = \infty$ if $S \cap T \neq \emptyset$ or there are $s \in S$ and $t \in T$ such that $\{s, t\} \in E$). Notice that $\kappa_{ST}(G) \geq \kappa(G)$.

Grids and LoS networks. For positive integers d, and $n \geq 2$, let $\mathbb{Z}_n^d$ be the d-dimensional cube $\{1, \ldots, n\}^d$. We say that distinct points P_1 and P_2 in a cube *share a line of sight* if their coordinates differ in a single place. A graph $G = (V, E)$ is said to be a *Line of Sight (LoS) network of size n, dimension d, and range parameter* ω if there exists an embedding $f_G : V \to \mathbb{Z}_n^d$ such that $\{u, v\} \in E$ if and only if $f_G(u)$ and $f_G(v)$ share a line of sight and the (Euclidean) distance between $f_G(u)$ and $f_G(v)$ is less than ω. In the rest of the paper a LoS network G is always given along with some embedding f_G in $\mathbb{Z}_n^d$ for some d and n, and with slight *abus de langage* we will often refer to the vertices of G, $u, v \in V$ in terms of their corresponding points $f_G(u), f_G(v), \ldots$ in $\mathbb{Z}_n^d$, and in fact the embedding f_G will not be mentioned explicitly. Note that d-dimensional hypergrids, $\mathcal{H}_{n,d}$, as defined in [15] are particular LoS networks with $\omega = 2$ and all possible n^d vertices. In the forthcoming sections we will study *augmented* hypergrids $\mathcal{H}_{n,d,\omega}$ (or simply $\mathcal{H}_{n,\omega}$ in the 2-dimensional case), namely d-dimensional LoS networks with range parameter $\omega > 2$ containing all possible n^d nodes.

Paths, Monitors and Identifiability. In BNT one takes measurements along paths, and the quality of the monitoring scheme depends on the choice of such paths. Let $\mathbb{P}$ be a set of paths over some node set V. For a node $v \in V$, let $\mathbb{P}(v)$ be the set of paths in $\mathbb{P}$ passing through v. For a set of nodes U, $\mathbb{P}(U) = \bigcup_{u \in U} \mathbb{P}(u)$. Hence if $U \subseteq V$, $\mathbb{P}(U) \subseteq \mathbb{P}(V)$. Crucially, we identify two disjoint sets of vertices S and T, and assume that $\mathbb{P}$ is the set of all $S - T$ paths in G, i.e. simple paths with one end-point in S and the other one in T. This is similar to the CSP probing scheme analyzed in [18], but the scheme in that paper does not assume $S \cap T = \emptyset$.

Traditionally in Network Tomography all measurements originate and end at special *monitoring stations* that are connected to the structure under observation. For any tomographic process to have any chance of succeeding one has to assume that such *monitors* are infallible. It is therefore customary to assume that the monitors are external to the given network, but connected to it through a designated set of nodes. $S \cup T$ is such set in our case. We call the pair (S, T) a *monitor placement*. In this settings, two sets of vertices U and W are *separable* if $\mathbb{P}(U) \triangle \mathbb{P}(W) \neq \emptyset$. A set of vertices N is *k-identifiable* (with respect to the probing scheme $(\mathbb{P}, S, T)$) if and only if any $U, W \subseteq N$, with $U \triangle W \neq \emptyset$ and $|U|, |W| \leq k$, U are separable. The *maximal identifiability* of N with respect to $(\mathbb{P}, S, T)$, $\mu(N, \mathbb{P}, S, T)$, is the largest k such that N is k-identifiable. For a graph $G = (V, E)$, we write $\mu(G, \mathbb{P}, S, T)$ to indicate the maximal identifiability of the set of nodes in V which are used in at least a path of $\mathbb{P}$. In what follows we usually omit the dependency of μ on the probing scheme $(\mathbb{P}, S, T)$ when this is clear from the context.

Note that k-identifiability is monotone: if G is k-identifiable then it is k'-identifiable for any $k' < k$. This implies that to prove that $\mu(N) \leq k - 1$ it is sufficient to show that N is **not** k-identifiable. By the definition given above this boils down to showing the existence of two distinct node sets U and W in N of cardinality at most k that are not separable.



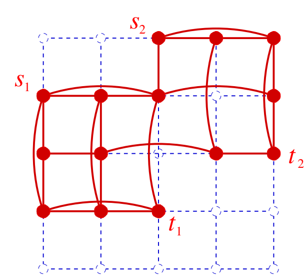


Fig. 1. On the left, the network $\mathcal{H}_{n,\omega}$ for $n = 5$ and $\omega = 4$ (note that vertices u and v are not adjacent); on the right a more general example of LoS network, having $\omega = 3$, embedded in $\mathbb{Z}_5^2$ (represented as a dashed grid).

Conversely, if we want to prove that $\mu(N) \geq k$ for some k, then it is enough to argue that all distinct node sets U and W of cardinality $|U|, |W| \leq k$ are separable. To prove this we have to show that for any two distinct node sets U and W of cardinality at most k there exists a path in $\mathbb{P}$ intersecting exactly one between U and W.

3 Failure Identifiability in Augmented Hypergrids

Let $\omega > 2$ be an integer. In this section we analyze the maximal identifiability of augmented hypergrids. To maximize clarity, we provide full details for the special case of $\mathcal{H}_{n,\omega}$, the 2-dimensional augmented hypergrid. The proof of the result for d-dimensional structures, which we state at the end of this section, is left for the full version of this work.

In [15] two of us showed that $\mu(G) \leq \delta(G)$ for any $(\mathbb{P}, S, T)$. In $\mathcal{H}_{n,\omega}$ each node u has $\omega - 1$ edges for each one of the possible directions (north, south, east, west). Hence the minimal degree in $\mathcal{H}_{n,\omega}$ is reached at the corner nodes and it is $2(\omega - 1)$. Thus $\mu(\mathcal{H}_{n,\omega}) \leq 2(\omega - 1)$ for any $(\mathbb{P}, S, T)$. In the remainder of this section we pair this up with a tight lower bound for a specific monitor placement. Note that these results readily imply the upper bound in Theorem 1 as in augmented hypegrids the vertex connectivity is actually equal to the network's minimum degree. The rest of this section focuses on the second inequality in that theorem.

We say that nodes with coordinates $(1, j)$ in $\mathcal{H}_{n,\omega}$, for some $j \in \{1, \ldots, n\}$, are *on the north border* of $\mathcal{H}_{n,\omega}$. Analogously we can define nodes on the south, west and east borders of $\mathcal{H}_{n,\omega}$. Given a node u of $\mathcal{H}_{n,\omega}$, identified as a pair $(i, j) \in \mathbb{Z}_n^2$, and a positive integer k, we define:

$$SE_k(u) = \{(i', j') \in \mathbb{Z}_n^2 : i + k \geq i' \geq i \ \wedge \ j + k \geq j' \geq j\}$$

and

$$NW_k(u) = \{(i', j') \in \mathbb{Z}_n^2 : i - k \leq i' \leq i \ \wedge \ j - k \leq j' \leq j\}.$$

In particular we denote by $SE(u)$ (resp. $NW(u)$) the union of all $SE_k(u)$ (resp. $NW_k(u)$). Furthermore $\partial SE_k(u)$ (resp. $\partial NW_k(u)$) is the set of all points in $SE_k(u)$ (resp. $NW_k(u)$) with coordinates (i', j) or (i, j'). Expressions $\partial SE(u)$ and $\partial NW(u)$ are defined analogously. Also, we say that a direction X (north, south, west, east) is *W-saturated on u* all neighbours of u in direction X are in W.

Definition 1. *(W-unreachability) Let $u = (i, j)$ be a node in $\mathcal{H}_{n,\omega}$ and W be a set of nodes in $\mathcal{H}_{n,\omega}$. A node $u' = (i', j')$ for $i' \geq i$ and $j' \geq j$ is W*-unreachable from u *if either $\partial SE(u) \subseteq W$ or $\partial NW(u') \subseteq W$. Otherwise we say that u' is W-reachable from u.*

A *canonical monitor placement* for $\mathcal{H}_{n,\omega}$ is a pair (S, T), such that S is formed by the node $(1, 1)$ and its neighbours, and T formed by (n, n) and it neighbours. Hence $|S| = |T| = 2\omega - 1$. We are now ready to state the main result in this section.

Theorem 4. *Let $n, \omega \in \mathbb{N}$, $\omega > 2$ and $n > 3(\omega - 1)$. Let (S, T) be a canonical monitor placement for $\mathcal{H}_{n,\omega}$. Then $\mu(\mathcal{H}_{n,\omega}) \geq 2(\omega - 1) - 1$.*

Proof. We have to prove that for any pair of node sets U, and W of cardinality at most $2(\omega - 1) - 1$, with $U \triangle W \neq \emptyset$ we can build an $S - T$ path touching exactly one of them. Assume without loss of generality that $u \in U \setminus W$. Since $|S| = |T| = 2(\omega - 1)$ and $|W| < 2(\omega - 1)$, there is a node in $s \in S \setminus W$ and a node $t \in T \setminus W$. Assume without loss of generality that $s = (1, 1)$ (the case $s \neq (1, 1)$ is similar, and give even better results). Similarly for T, assume that $t = (n, n) \notin W$.

We build two *disjoint* paths i_u and o_u such that their concatenation is an $S - T$ path passing through u and not touching nodes in W. We show how to build i_u (o_u is analogous).

If $u = (i, j)$ and $\min(i, j) > \omega - 1$ and u is W-reachable from s we proceed by a careful induction on $|NW(u)|$. In the inductive step, two things can happen. If u is far from the north and west borders there is at a least a direction X between North and West which is not W-saturated. Hence there is a node $u' \in NW(u) \setminus W$ on direction X from u at distance less than ω. Hence there is an edge $\{u', u\} \in \mathcal{H}_{n,\omega}$. Since $NW(u') \subset NW(u)$ the inductive hypothesis applied on u', give us a path $i_{u'}$ and the path $i_u = i_{u'}, u$. Alternatively if u is close to s (i.e. $\min(i, j) \leq \omega - 1$) we know that u is W-reachable and this guarantees the existence of a neighbour u' of u in $NW(u)$ that is NOT in W and the inductive hypothesis can be applied to u' again to complete i_u.

The induction reaches a base case in one of two possible ways. If $|NW(u)| = 1$, then $u = s$ and we have done: i_u is s. Otherwise $|NW(u)| > 1$ but u is W-unreachable from s. In such case we proceed as follows (see also Fig. 2 for an example). Notice that in this case it must be that u has less than $\omega - 1$ neighbours either in direction North or West, for otherwise it would not be possible for W, which is of size at most $2(\omega - 1) - 1$, to cover $\partial NW(u)$ or $\partial SE(s)$. Let $u = (i, j)$, hence, by unreachability property, it must be that in $NW(u)$ there are at least $t = (i - 1) + (j - 1) \geq 2$ nodes in W. Let us look at the neighbours of u in $SE(u)$ which are South of u at distance at most $\omega - 1 - i$ from u and East of u at distance at most $\omega - 1 - j$. (Notice that these nodes are at distance at most $\omega - 1$ from the North and West borders). First we claim that either in direction South or direction East, there is a neighbour (say wlog direction South) $u' = (i', j)$ of u at distance at most $\omega - 1 - i$ from u such that both u' and $u'' = (i', 1)$ are not in W. This is because the sum of nodes at distance at most $\omega - i - 1$ from u in direction South and at distance at most $\omega - j - 1$ in direction East is at most $2\omega - i - j$. Hence there are at most $2\omega - i - j$ pairs of nodes of the type (u', u''), but only $2\omega - i - j - 1$ nodes in W (the latter is because $|W| \leq 2\omega - 3$ and $t = (i - 1) + (j - 1)$ nodes of W are already used in $NW(u)$). Then there is at least a pair (u', u'') such that neither u' nor u'' are in W. Hence the path $i_u = s, u'', u', u$ connecting s to u without touching W. This path i_u is ok, unless u' is already on the path we have built by induction so far. In that case we can cut i_u at u' and link it to the inductive path.

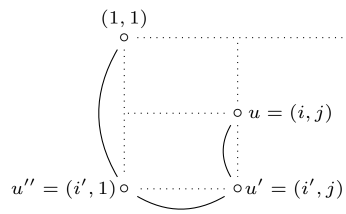


Fig. 2. An example of how to build i_u when u is not W-reachable and in $SE_k((1,1))$ for some $k < \omega - 1$.

The argument presented so far leaves a gap in che case when $u \in U \setminus W$ is close to s and it is W-unreachable. This case is in fact not very different from the last one we have considered. As in that case we consider the neighbours of u in $\partial SE_{\omega-i}(u)$ at distance $\omega - i$ in the South direction (instead of distance $\omega - i - 1$ as in the previous case) and at distance $\omega - j$ in the East direction (instead of $\omega - j - 1$). Exactly the same counting argument now justifies three pairs of nodes (u', u''), (v', v'') and (w', w'') such that none of $u', u'', v', v'', w', w''$ are in W (notice that as before, u'', v'' and w'' are nodes on the border of $\mathcal{H}_{n,\omega}$). If either one between v' and w', say wlog w', is exactly at distance $\omega - 1$ from u, then we can define a path i_u touching u and not touching W as $i_u = (1,1), u'', u', u, w'$ (see also Fig. 3). If both of v' and w'' are at distance $< \omega - 1$, then two among u', v' and w', say u', and v' are on the same direction, say wlog South, and hence one of them, say wlog u' is northern of the other. In this case we define i_u as the path $(1,1), u'', u', u, v'$.

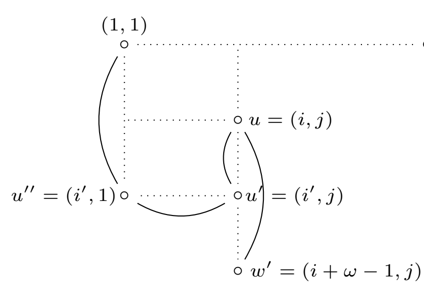


Fig. 3. An example where $u \in SE_k(1,1))$ for some $k < \omega - 1$ and w' is at South distance exactly $\omega - 1$ from u.

□

Theorem 4 generalizes to d-dimensional augmented hypergrids. We leave the details to the full version of this work.

Theorem 5. *Let $d, n, \omega \in \mathbb{N}$, $d, n \geq 2$ and $\omega > 2$. There is a monitor placement for $\mathcal{H}_{n,d,\omega}$ for which $\mu(\mathcal{H}_{n,d,\omega}, \mathbb{P}, S, T) \geq d(\omega - 1) - 1$.*

4 General Topologies

We now look at the maximal identifiability in arbitrary networks. Theorem 2 stated in Section 1 will be a consequence of two independent results. In [15] it was proved that $\mu(G) \leq \delta(G)$, for any monitor placement (S, T). Here we show that $\mu(G)$ can be upper bounded in terms of κ_{ST}, the size of the minimal node set separating S from T.

Theorem 6. *Let $G = (V, E)$ be a graph and (S, T) be a monitor placement. Then $\mu(G) \leq \kappa_{ST}(G)$.*

Proof. If there is no vertex set in G separating S and T, $\kappa_{ST}(G) = \infty$ and the result is trivial. Let K be the set witnessing the minimal separability of S from T in G. Hence $|K| = \kappa_{ST}(G)$. Let $N(K)$ be the set of neighbours of nodes in K and notice this cannot be empty since K is disconnecting G. Pick one $w \in N(K)$ and define $U := K$ and $W := U \cup \{w\}$. Clearly $\mathbb{P}(U) \subseteq \mathbb{P}(W)$. To see the opposite inclusion assume that there exists a path from S to T passing from w but not touching $U = K$. Then K is not separating S from T in G. Contradiction. $\square$

Note that, while in general $\kappa_{ST}(G)$ may be larger than $\delta(G)$, if S and T are separated by a set of $\kappa(G)$ vertices then, by inequality (2), the bound in Theorem 6 is at least as good as the minimum degree bound proved earlier by the first two authors [15]. This implies the upper bound in Theorem 2.

Moving to lower bounds, in this section we prove the following:

Theorem 7. *Let $G = (V, E)$ and (S, T) be a monitor placement for G. Then $\mu(G) \geq \min(\kappa(G) - 1, |S|, |T|) - 1$.*

The lower bound in Theorem 2 can be derived easily from Theorem 7. Let K be a vertex separator in G of size $\kappa(G)$, set S^K to be the first $\lfloor \kappa(G)/2 \rfloor$ elements of K and $T^K = K \setminus S^K$. By Theorem 7 the maximal identifiability of G is at least $|S^K| - 1 = \lfloor \kappa(G)/2 \rfloor - 1$.

The proof of Theorem 7 uses Menger's Theorem, a well-known result in graph theory (see [16, Theorem 5.10, p. 48] for its proof).

Theorem 8. *(Menger's Theorem) Let $G = (V, E)$ be a connected graph. Then $\kappa(G) \geq k$ if and only if each pair of nodes in V is connected by at least k node-disjoint paths in G.*

Menger's Theorem is central to the following Lemma which is used in the proof of Theorem 7.

Lemma 1. *Let $G = (V, E)$. Let $W \subseteq V$ such that $|W| \leq \kappa(G) - 2$. Then any pair of vertices in $V \setminus W$ is connected by at least two vertex-disjoint simple paths not touching W.*

Proof. By Menger's Theorem, for any pair of nodes u and v in $V \setminus W$ there are at least $\kappa(G)$ vertex-disjoint paths from u to v in G. Call $\mathbb{P}$ the set of such paths. Since $|W| \leq \kappa(G) - 2$, then the nodes of W can be in at most $\kappa(G) - 2$ of paths in $\mathbb{P}$. Hence there are at least two paths in $\mathbb{P}$ not touching W. $\square$

Proof of Theorem 7. Let $G = (V, E)$ be an undirected connected graph and (S, T) be a monitor placement in G. Note that without loss of generality $\min(\kappa(G) - 1, |S|, |T|) > 1$ (for otherwise there is nothing to prove).

Assume first that $|S| \geq \kappa(G) - 1$ and $|T| \geq \kappa(G) - 1$. We claim that

$$\mu(G) \geq \kappa(G) - 2.$$

We show that for any distinct non-empty subsets U and W of V of size at most $\kappa(G) - 2$, there is a path in $\mathbb{P}$ touching exactly one between U and W. Given such U and W, fix a node $u \in U \triangle W$ and assume w.l.o.g. that $u \in U$. Since $|W| \leq \kappa(G) - 2$ and $|S| \geq \kappa(G) - 1$ there is at least a node in $s \in S \setminus W$. By the Claim above applied to nodes s and u and to the set W, there are two vertex-disjoint simple paths π_1^s, π_2^s from s to u not touching W. The same reasoning applied to T, guarantees the existence of a node $t \in T \setminus W$ and two vertex-disjoint paths π_1^t, π_2^t from u to t not touching W. If at least one between π_1^s, and π_2^s only intersects one of π_1^t, and π_2^t at u then the concatenation of such paths is a (longer) simple path from s to t passing through u and not touching W. Otherwise the concatenation of one between π_1^s, and π_2^s with one between π_1^t, and π_2^t is a non simple path. In what follows we show that the subgraph of G induced by the four paths does contain a simple path from s to t passing through u and not touching W. In the construction below we exploit the fact that π_1^s, and π_2^s (resp. π_1^t, and π_2^t) are simple and vertex disjoint. Let p be a path from s to u. Define an order on the nodes of p as follows: $v \prec_p w$ if going from v to u we pass though w. From now on we will use $\prec$ instead of $\prec_p$ when the path under consideration will be clear from the context. Let Z_1^j be the nodes in $\pi_1^s \cap \pi_j^t$. Z_1^1 and Z_1^2 are disjoint but there will be a node in those sets, say z, which is minimal according to $\prec$. Without loss of generality let us say that $z \in Z_1^1$. The subpath $\pi_1^s[s \dots z]$ of π_1^s going from s to z, is intersecting neither π_1^t nor π_2^t. Hence the concatenation of the following three disjoint paths defines a simple path from s to t passing through u avoiding W, hence a path in $\mathbb{P}$ with the required properties:

1. $\pi_1^s[s \dots z]$, going form s to z;
2. $\pi_1^t[z \dots u]$ a sub path of π_1^t going from u to z and traversed in the other direction;
3. π_2^t, connecting u to t.

Now assume that at least one between $|S|$ and $|T|$ is less than $\kappa(G) - 1$. Let $r = \min(|S|, |T|) - 1$. As before we prove that for all distinct non-empty U and W subsets of V of size at most r, there is an $S - T$ path in G, hence in $\mathbb{P}$, touching exactly one between U and W. Let $u \in U \triangle W$ and without loss of generality assume $u \in U$. Notice that $r + 1 = \min(|S|, |T|)$, then both $|S| \geq r + 1$

and $|T| \geq r+1$. Since $|W| \leq r$, as before there are $s \in S \setminus W$ and $t \in T \setminus W$. Furthermore, since $\kappa(G) \geq \min(|S|, |T|)$, then by previous observation on $|S|$ and $|T|$, $\kappa(G) \geq r+1$ and, since $|W| \leq r$, then $\kappa(G)-|W| \geq 2$, that is $|W| \leq \kappa(G)-2$. As in the previous case we can apply the Claim above once to s, u and W getting the vertex-disjoint paths π_1^s and π_2^s from s to u, and once to t, u and W getting the vertex-disjoint paths π_1^t and π_2^t from t to u. The proof then follows by the same steps as in the previous case. We then have proved that if $|S|$ or $|T|$ are smaller than $\kappa(G)-1$, then $\mu(G) \geq \min(|S|, |T|)-1$ and the proof of Theorem 7 is complete. □

Proof of Theorem 3. We complete this section investigating a different way to relate the graph vertex connectivity to $\mu(G)$. It is easy to see that, in general, the bounds in Theorem 2 are not very tight, particularly when $\kappa(G)$ is large. However, if $\kappa(G)$ is small, we can do better.

In what follows let K be a minimal vertex separator in G. Let $G_i^K = (V_i^K, E_i^K)$, $i \in \{1, \ldots, r_K\}$ be the $r_K \geq 2$ connected components remaining in G after removing K. Since $\kappa(G) \leq \frac{n}{3}$, then $2\kappa(G) \leq n - \kappa(G)$ and one can define disjoint sets S, and T with $\kappa(G)$ vertices each in such a way that the smallest among the V_i^K's contains only elements of S. This can be done as follows: if the smallest V_i^K's has less than $\kappa(G) - \ell$ nodes, then assign all its nodes to S. Then use the other components G_j^K's to assign ℓ nodes to S and $\kappa(G)$ other nodes to T. If the smallest V_i^K has more than $\kappa(G)$ nodes, choose $\kappa(G)$ among them and put them in S. Choose $\kappa(G)$ nodes in other components and assign them to T.

We now prove that the set of simple paths between S and T defined as above allow a very high identifiability. The lower bound on $\mu(G)$ follows from Theorem 7 noticing that $|S| = |T| > \kappa(G) - 1$. We now prove that $\mu(G) \leq \kappa(G)$. Let G_i^K be the component where all the S-nodes are assigned. Let w be a node in $V_i^K \cap N(K)$. This node has to exists since G was connected and the removal of K is disconnecting G_i^K from K. Fix $U = K$ and $W = K \cup \{w\}$. We will show that $\mathbb{P}(U) = \mathbb{P}(W)$. It suffices to prove that $\mathbb{P}(\{w\}) \subseteq \mathbb{P}(K)$, since clearly $\mathbb{P}(U) \subseteq \mathbb{P}(W)$. Observe that no $S-T$ path p in G can live entirely inside G_i^K, i.e. have all of its nodes in V_i^K. This is because at least one end-point (that in T) it is necessarily missing in any path entirely living only in G_i^K. Hence a path touching w is either entering or leaving G_i^K. But outside of G_i^K w is connected only to K, since otherwise K would not be a minimal vertex separator. Hence it must be $\mathbb{P}(\{w\}) \subseteq \mathbb{P}(K)$. We have found U, W of size $\leq \kappa(G)$ such that $\mathbb{P}(U) = \mathbb{P}(W)$. The upper bound follows.

Arbitrary LoS networks have minimum degree, and hence also vertex connectivity at most $2d(\omega-1)$. The next corollary follows directly from Theorem 3.

Corollary 1. *Let G be an arbitrary LoS network over n nodes and with fixed range parameter ω, independent of n, such that $n \geq \omega$. Then $\mu(G) \geq \kappa(G) - 2$.*

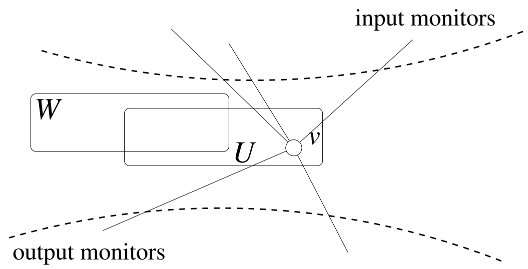


Fig. 4. A node $v \in U \Delta W$ and a possible way to connect it to S and T.

5 Random Networks and Tradeoffs

The main aim of this work is to characterize the identifiability in terms of the vertex connectivity. In this section we prove that tight results are possible in random graphs. Also we show an interesting trade-off between the success probability of the various random processes and the size of the sets S and T. Finally, random graphs give us constructions of networks with large identifiability.

5.1 Sub-Linear Separability in Erdős-Rényi Graphs

We start our investigation of the identifiability of node failures in random graphs by looking at the binomial model $G(n,p)$, for fixed $p \leq 1/2$ (in this section only we follow the traditional random graph jargon and use p to denote the graph edge probability rather than a generic path). The following equalities, which hold with probability approaching one as n tends to infinity (that is *with high probability* (w.h.p.)), are folklore:

$$\kappa(G(n,p)) = \delta(G(n,p)) = np - o(n). \tag{3}$$

(see [5]). Here we describe a simple method which can be used to separate sets of vertices of sublinear size.

We assume, for now, that S and T are each formed by $\gamma = \gamma(n)$ nodes with $\kappa(G(n,p)) \leq \gamma < n/2$. Let $M = S \cup T$. Let U and W be two arbitrary subsets of $V \setminus M$ of size k. The probability that U and W are separable is at least the probability that an element v of $U \Delta W$ (w.l.o.g. assume $v \in U \setminus W$) is directly connected to a node in S and to a node in T. This event has probability $(1-(1-p)^\gamma)^2$. Hence the probability that U and W cannot be separated is at most $1-(1-(1-p)^\gamma)^2 = 2(1-p)^\gamma - (1-p)^{2\gamma}$ and therefore the probability that some pair of sets U and W of size k (not intersecting M) fail is at most $2\binom{n-2\gamma}{k}\binom{2k}{k}(1-p)^\gamma$.

Theorem 9. *For fixed p with $p \leq 1/2$, under the assumptions above about the way monitors are placed in $G(n,p)$, the probability that $G(n,p)$ is not k-vertex separable is at most $2k\binom{n}{k}^2 \mathrm{e}^{(2k-\gamma)p}$.*

Proof. The argument above works if both U and W contain no vertex in M. The presence of elements of vertices in M in U or W may affect the analysis in two ways. First v could be in M (say $v \in S$). In this case U and W are separable if v is directly connected to a vertex in T. This happens with probability $(1-(1-p)^{\gamma}) > (1-(1-p)^{\gamma})^2$. Second, M might contain some elements of U and W different from v. In the worst case when v is trying to connect to M, it must avoid at most $2k$ element of such set. There is at most $\sum_{h \leq k} \binom{n}{h}^2 \leq k\binom{n}{k}^2$ pairs of U and W of size at most k. Thus the probability that $G(n,p)$ fails to be k-vertex separable is at most $2k\binom{n}{k}^2(1-p)^{\gamma-2k}$. and the result follows as $1-p \leq \mathrm{e}^{-p}$. $\square$

Note that the bound in Theorem 9 can only be small if $k = o(n)$ for otherwise the factor $\mathrm{e}^{(2k-\gamma)p}$ is large. In fact it has to be $k = O(n^{\epsilon})$ for sufficiently small positive ϵ otherwise the large factor $\binom{n}{k}^2$ is not "killed off" by the magnitude of the small exponential.

5.2 Linear Separability in Erdős-Rényi Graphs

The argument above cannot be pushed all the way up to $\kappa(G(n,p))$. When trying to separate vertex sets containing $\Omega(n)$ vertices the problem is that these sets can form a large part of M and the existence of direct links from v to $S \setminus W$ and $T \setminus W$ is not guaranteed with sufficiently high probability. However a different argument allow us to prove the following:

Theorem 10. *For fixed $p \leq 1/3$, $\kappa(G(n,p)) - 1 \leq \mu(G(n,p)) \leq \kappa(G(n,p))$ w.h.p.*

Full details of the proof are left to the final version of this paper, but here is an informal explanation. The upper bound follows immediately from (3) and Theorem 6. For the lower bound we claim that the chance that two sets of size at most $\kappa(G(n,p)) - 1$ are not vertex separable is small. First note that w.h.p. $G(n,p)$ has a single vertex of minimum degree. Choose S of size at least $n/3$ so that it contains such vertex. Choose T of size at least $n/3$ in $V \setminus (S \cup N(S))$ arbitrarily. To believe our claim pick two sets U and W, assume without loss of generality that $U \setminus W \neq \emptyset$ and remove, W from the graph. $G(n,p) \setminus W$ is still a random graph on at least $n - np = \Omega(n)$ vertices and constant edge probability. Results in [4] imply that $G(n,p) \setminus W$ has a Hamilton path starting at any $s \in S$ with probability at least $1 - o(2^{-n})$ (and in fact one can use well-known algorithmic techniques [1] to find one such path in polynomial time, w.h.p.). Such Hamiltonian path, by definition, contains a path from S to T passing through $v \notin W$, for every possible choice of v. This proves, w.h.p., the separability of sets of size up to $\kappa(G(n,p)) - 1$ (if $|W| = \kappa(G(n,p)) - 1$, v is the unique vertex of minimum degree and $W \subseteq N(v)$ then one needs to use a Hamiltonian path starting at v).

5.3 Random Regular Graphs

A standard way to model random graphs with fixed vertex degrees is Bollobas' configuration model [3]. There's n buckets, each with r free points. A random pairing of these free points has a constant probability of not containing any pair containing two points from the same bucket or two pairs containing points from just two buckets. These configurations are in one-to-one correspondence with r-regular n-vertex simple graphs. Denote by $\mathcal{C}_{n,r}$ the set of all configurations $C(n,r)$ on n buckets each containing r points, and let $G(r\text{-reg})$ be a random r-regular graph.

As before assume $|S| = |T| = \gamma$. The main result of this section is the following:

Theorem 11. *Let $r \geq 3$ be a fixed integer. $r-2-o(1) \leq \mu(G(r\text{-reg})) \leq r$ w.h.p.*

The result resembles Theorem 3 but its proof uses different techniques. The upper bound is true of any r-regular graph G as $\mu(G) \leq \delta(G) = r$. The lower bound is a consequence of the following:

Lemma 2. *Let $r \geq 3$ be a fixed integer. Two sets U and W with $U, W \subseteq V(G(r\text{-reg}))$ and $\max(|U|, |W|) \leq k$ are separable w.h.p. if $k = r - 2 - o(1)$.*

Proof. In what follows we often use graph-theoretic terms, but we actually work with a random configuration $C(n,r)$. Let U and W be two sets of k buckets. For simplicity assume that (the vertices corresponding to the elements of) both U and W are subsets of $V \setminus M$. The probability that U and W can be separated is at least the probability that a (say) random element v of $U \triangle W$ (w.l.o.g. $v \in U \setminus W$) is connected to S by a path of length at most ℓ_s and to T by a path of length at most ℓ_t, neither of which "touch" W. Fig. 5 provides a simple example of the event under consideration. The desired paths can be found using algorithm PathFinder below that builds the paths and $C(n,r)$ at the same time.

PathFinder(v, ℓ_s, ℓ_t, W) SimplePaths(v, ℓ_s, ℓ_t, W)**.** Starting from v, build a simple path p^s of length ℓ_s that avoids W. Similarly, starting from v, build a simple path p^t of length ℓ_t that avoids W. RandomShooting(p^s, p^t)**.** Pair up all un-matched points in p^s and p^t. Complete the configuration $C(n,r)$ by pairing up all remaining points.

Sub-algorithm SimplePaths can complete its constructions by pairing points starting from elements of the bucket v then choosing a random un-matched point in a bucket u, then picking any other point u and then again a random un-matched point and so on, essentially simulating two random walks RW_s and RW_t on the set of buckets. Note that the process may fail if at any point we re-visit a previously visited bucket or if we hit W or even M. However the following can be proved easily.

Claim. RW_s and RW_t succeed w.h.p. provided $\ell_s, \ell_t \in o(n)$.

As to RandomShooting, the process succeeds if we manage to hit an element of S from p^s and an element of T from p^t.

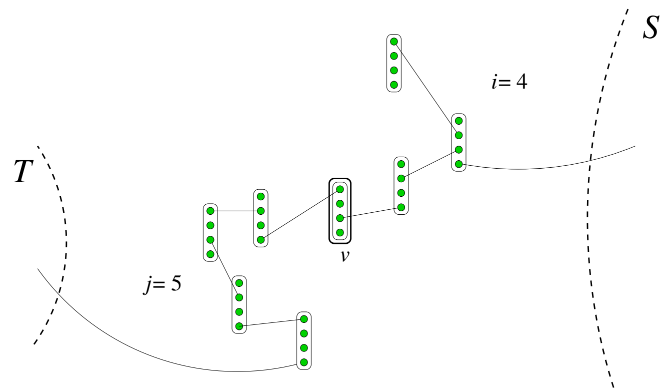


Fig. 5. Assume $r = 4$. The picture represents a bucket (i.e. vertex) $v \in U \triangle W$ and two possible "paths" (sequences of independent edges such that consecutive elements involve points from the same bucket) of length 3 and 5, respectively connecting it to S and T.

Claim. RandomShooting(q_s, q_t, S, T) succeeds w.h.p. if $\ell_s, \ell_t \in \omega(1)$.

Any un-matched point in p^s or p^t after SimplePaths is complete is called *useful*. Path p^s (resp. p^t) contains $q_s = (r-2)\ell_s + 1$ (resp $q_t = (r-2)\ell_t + 1$) useful points. During the execution of RandomShooting a single useful point "hits" its target set, say S, with probability proportional to the cardinality of S. Hence the probability that none of the q_s useful points hits S is $(1 - \frac{\gamma}{n})^{q_s}$ and the overall success probability is $(1-(1-\frac{\gamma}{n})^{q_s})(1-(1-\frac{\gamma}{n})^{q_t})$.

Back to the proof of Lemma 2 Set $\ell_s = \ell_t = \ell$ and q the common value of q_s an q_t. The argument above implies that the success probability for U and W is asymptotically approximately $(1-(1-\frac{\gamma}{n})^q)^2$ and the rest of the argument (and its conclusion) is very similar to the $G(n,p)$ case (the final bound is slightly weaker, though). The chance that a random r-regular graph is not k-vertex separable is at most

$$O(n^{2k}) \times (1-(1-(1-\frac{\gamma}{n})^q)^2) \leq O(n^{2k}) \times 2(1-\frac{\gamma}{n})^q \leq O(n^{2k}) \times 2\mathrm{e}^{-\frac{\gamma}{n}q},$$

which goes to zero as n^{-C} provided ℓ is at least logarithmic in n. The constraints on ℓ from the claims above imply that the parameter can be traded-off agains γ to achieve high identifiability. $\square$

References

1. D. Angluin, and L. G. Valiant Fast Probabilistic Algorithms for Hamiltonian Circuits and Matchings *Journal of Computers and Systems Sciences*, 18:155–193, 1979.
2. N. Bartolini, T. He, and H. Khamfroush Fundamental limits of failure identifiability by boolean network tomography. In *INFOCOM 2017*. IEEE, 2017.
3. B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1:311–316, 1980.
4. B. Bollobás, T. I. Fenner, and A. M. Frieze. An algorithm for finding hamilton paths and cycles in random graphs. *Combinatorica*, 7:327–341, 1987.
5. B. Bollobás. *Random Graphs*. Number 73 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2nd edition, 2001.
6. R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu. Network tomography: Recent developments. *Statistical Science*, 19(3):499–517, 2004.
7. M. Cheraghchi, A. Karbasi, S. Mohajer, and V. Saligrama. Graph-constrained group testing *IEEE Transactions on Information Theory*, 58(1):248–262, 2012
8. M. Coates, A. O. Hero, R. Nowak, and B. Yu. Internet tomography. *IEEE Signal Processing Magazine*, 19:47–65, 2002.
9. A. Czumaj and X. Wang. Communication problems in random line-of-sight ad-hoc radio networks. In *Stochastic Algorithms: Foundations and Applications, 4th International Symposium, SAGA 2007, Murich, Switzerland, September 13-14, 2007, Proceedings*, pages 70–81, 2007.
10. L. Devroye and L. Farczadi. Connectivity for line-of-sight networks in higher dimensions. *Discrete Mathematics and Theoretical Computer Science*, 15(2):71–86, 2013.
11. D.-M. Du, and F. K. Hwang. Combinatorial Group Testing and Its Applications. World Scientific, 2000.
12. N. G. Duffield. Simple network performance tomography. In *Proceedings of the 3rd ACM SIGCOMM Internet Measurement Conference, IMC 2003, Miami Beach, FL, USA, October 27-29, 2003*, pages 210–215. ACM, 2003.
13. A. M. Frieze, J. M. Kleinberg, R. Ravi, and W. Debany. Line-of-sight networks. In Nikhil Bansal, Kirk Pruhs, and Clifford Stein, editors, *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7-9, 2007*, pages 968–977. SIAM, 2007.
14. D. Ghita, C. Karakus, K. Argyraki, and P. Thiran. Shifting network tomography toward a practical goal. In *Proceedings of the Seventh COnference on Emerging Networking EXperiments and Technologies*, CoNEXT '11, pages 24:1–24:12, New York, NY, USA, 2011. ACM.
15. N. Galesi and F. Ranjbar. Tight bounds for maximal identifiability of failure nodes in boolean network tomography. In IEEE, editor, *2018 IEEE 38th International Conference on Distributed Computing Systems*, pages 212–222, 2018.
16. F. Harary. *Graph Theory*. Addison-Wesley, Reading, MA, 1969.

17. T. Kloks and D. Kratsch Finding All Minimal Separators of a Graph STACS 94, 11th Annual Symposium on Theoretical Aspects of Computer Science, *Lecture Notes in Computer Science* Vol. 775, pp. 759–768, 1994
18. L. Ma, T. He, K. K. Leung, A. Swami, and D. Towsley. Inferring link metrics from end-to-end path measurements: Identifiability and monitor placement. *IEEE/ACM Trans. Netw.*, 22(4):1351–1368, 2014.
19. L. Ma, T. He, A. Swami, D. Towsley, K. K. Leung, and J. Lowe. Node failure localization via network tomography. In Carey Williamson, Aditya Akella, and Nina Taft, editors, *Proceedings of the 2014 Internet Measurement Conference, IMC 2014, Vancouver, BC, Canada, November 5-7, 2014*, pages 195–208. ACM, 2014.
20. 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.*, 25(1):434–450, 2017.
21. W. Ren and W. Dong. Robust network tomography: K-identifiability and monitor assignment. In *35th Annual IEEE International Conference on Computer Communications, INFOCOM 2016, San Francisco, CA, USA, April 10-14, 2016*, pages 1–9. IEEE, 2016.
22. P. Sangha, P.W.H. Wong, and M. Zito. Independent Sets in Restricted Line of Sight Networks In A. Fernández Anta et al., editors, *ALGOSENSORS 2017*, volume 10718 of *Lecture Notes in Computer Science*, pages 1–12. Springer Verlag, 2017.
23. P. Sangha and M. Zito. Finding large independent sets in line of sight networks. In D. Gaur and N. S. Narayanaswamy, editors, *Algorithms and Discrete Applied Mathematics; Third International Conference CALDAM 2017*, volume 10156 of *Lecture Notes in Computer Science*, pages 332–343. Springer Verlag, 2017.
24. Y. Vardi. Network tomography: Estimating source-destination traffic intensities from link data. *Journal of the American Statistical Association*, 91(433):365–377, 1996.