FRANTIC: A Fast Reference-based Algorithm for Network Tomography via Compressive Sensing

Sheng Cai   Mayank Bakshi   Sidharth Jaggi   Minghua Chen
The Chinese University of Hong Kong

***Abstract*—We study the problem of link and node delay estimation in undirected networks when at most k out of n links or nodes in the network are congested. Our approach relies on end-to-end measurements of path delays across pre-specified paths in the network. We present a class of algorithms that we call FRANTIC [1]. The FRANTIC algorithms are motivated by compressive sensing; however, unlike traditional compressive sensing, the measurement design here is constrained by the network topology and the matrix entries are constrained to be positive integers. A key component of our design is a new compressive sensing algorithm SHO-FA-INT that is related to the SHO-FA algorithm [1] for compressive sensing, but unlike SHO-FA, the matrix entries here are drawn from the set of integers $\{0, 1, \ldots, M\}$. We show that $\mathcal{O}(k \log n / \log M)$ measurements suffice both for SHO-FA-INT and FRANTIC . Further, we show that the computational complexity of decoding is also $\mathcal{O}(k \log n / \log M)$ for each of these algorithms.**

I. Introduction

Monitoring performance characteristics of individual links is important for diagnosing and optimizing network performance. Making direct measurements for each link, however, is impractical in large-scale networks because (i) nodes inside the networks may not be available to carry out measurements due to physical or protocol constraints, and (ii) measuring each link *separately* incurs excessive control-traffic overhead and is not scalable.

A viable alternative approach is network tomography [2]. It aims to infer the performance characteristics of internal links by *path measurements* between controllable nodes, where a path measurement is function of the characteristics of the links on the path. It does not require access to all the nodes and, more importantly, it allows clever solutions to leverage the network structure (*e.g.*, topology and graph properties) to *jointly* infer the performance characteristics of multiple links via path measurements. Many existing work have explored such insight to design excellent solutions that are able to infer the congested links with much less measurements than the direct link measurement approach [3]–[7]. See [8] for a survey.

Recently, Xu *et al.* [9] further argue that usually only a small fraction of network links, say k out of total $|E|$ links ($k \ll |E|$), are congested (*i.e.*, experiencing large congestion delay or high packet loss rate). They interpret each path measurement as a linear combination of the delays or loss rates of the k congested links. With these understanding, the problem of network tomography can be viewed as recovering a k-sparse link vector from a set of linear measurements.

Exploiting the "sparse congestion structure" insight, Xu *et al.* [9] propose a compressive sensing[2] based scheme that can identify any k congested links using $\mathcal{O}(T(\mathcal{N})k \log |E|)$ path measurements over any sufficiently-connected graph. Here, each of the path measurement is a random walk on the graph, and $T(\mathcal{N})$ is the mixing time of the random walk. Further, they show that one can actually *recover* the performance characteristics of any k congested links with again $\mathcal{O}(T(\mathcal{N})k \log |E|)$ path measurements by using an l_1 minimization technique. Similar results are also obtained by [12]–[14]. Given all these exciting results, a natural question is that can we do better and how?

In this paper, we build upon our recently developed compressive sensing algorithm named SHO-FA [1] to design a new network tomography solution that we call FRANTIC . FRANTIC achieves the following performance:

- FRANTIC can identify any k congested links (or nodes) out of n and recover the corresponding link (or node) performance characteristics using $\mathcal{O}(k \log n / \log M)$ path measurements with a high probability. Here, $M \in \mathbb{N}$ is a design parameter that is related to the number of hops allowed in each measurement.
- Furthermore, the FRANTIC decoding algorithm can recover the link (or node) performance characteristics in $\mathcal{O}(k \log n / \log M)$ steps.

As compared to the solution in [9], our solution improves both the number of measurements and the number of recovery steps from $\mathcal{O}(T(\mathcal{N})k \log n)$ to $\mathcal{O}(k)$ (obtainable by setting $M = \mathcal{O}(n)$.

The main techniques that lead to these improvements are as follows. First, in Section V, we develop an efficient compressive sensing algorithm SHO-FA-INT when the entries of the measurement matrix are constrained to be positive integers. Our algorithm borrows key ideas from a prior work [1] that studies compressive sensing in the unconstrained setting. A key technique here is to group together measurements and choose the "weights" of the measurement matrix as co-prime vectors. This ensures that each network link has a distinct signature in the measurement output, which allows us to decode the delay values for congested links in an iterative

[1]FRANTIC stands for **F**ast **R**eference-based **A**lgorithm for **N**etwork **T**omography v**I**a **C**ompressive sensing.

[2]Compressive sensing is a recently proposed paradigm for efficient sample and recovery of sparse signals [10], [11]. We will give a brief introduction later in Section V.

manner. Theorem 1 states the performance guarantees of our algorithm. Next, we propose a design for measuring the delay on congested links in a network in Section VI-A. An important insight in our design is that by using local loops at individual edges, end-to-end delay measurements can be designed to assign different integer weights to delays for different edges. We start with a compressive sensing matrix given by SHO-FA-INT and emulate the output of the matrix by first designing two correlated network paths, and then cancelling out the contribution of unwanted links by subtracting one from another. Theorems 2-4 state the performance guarantees of the FRANTIC algorithms. We also show that the decoding may be performed by employing SHO-FA-INT on the vector of end-to-end delays.

II. Model and problem formulation

Network and delay model: Let $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ be a undirected network with node set $\mathcal{V}$ and link set $\mathcal{E}$. In this paper, we consider the reference-based tomography problem where the topology of $\mathcal{N}$ is known. We assume that $\mathcal{N}$ is connected. We say that a node $v \in \mathcal{V}$ has delay d_v if every packet that passes through v is delayed by d_v. Similarly, a link $e \in \mathcal{E}$ has delay d_e if every packet passing through e in any direction is delayed by d_e. Let $\boldsymbol{d}_{\mathcal{V}} = (d_v : v \in \mathcal{V})$ and $\boldsymbol{d}_{\mathcal{E}} = (d_e : e \in \mathcal{E})$. Assume that $||\boldsymbol{d}_{\mathcal{V}}||_0 = k_{\mathcal{V}}$ and $||\boldsymbol{d}_{\mathcal{E}}||_0 = k_{\mathcal{E}}$. Both $\boldsymbol{d}_{\mathcal{V}}$ and $\boldsymbol{d}_{\mathcal{E}}$ are unknown but constant.

Measurement model: Each measurement is performed by sending test packets over pre-determined paths and measuring the end-to-end time taken for its transmission. Some nodes (resp. links) may be visited more than once in a given path. As a result, each measurement output y_i, $i = 1, 2, \ldots, \mu$, is a weighted sum of delays involving nodes and links that lie in the measurement path, where, weight of a given node or link is the number of times it is visited by the measurement path. In this paper, we consider the following two classes of measurements.

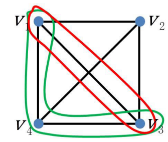
1. *Node measurements:* In the node measurement model, we associate each node with a real valued delay (see [13], for example). Let $\mathcal{S} \subseteq \mathcal{V}$ denote a subset of nodes in $\mathcal{N}$. Let $\mathcal{E}_{\mathcal{S}}$ denote the subset of links with both ends in $\mathcal{S}$, then $\mathcal{N}_{\mathcal{S}} = (\mathcal{S}, \mathcal{E}_{\mathcal{S}})$ is the induced subgraph of $\mathcal{N}$. A set $\mathcal{S}$ of nodes can be measured together in one measurement if and only if $\mathcal{N}_{\mathcal{S}}$ is connected.

2. *Link measurements:* In link measurement setup, we associate each link with a real valued delay. Let $\mathcal{T} \subseteq \mathcal{E}$ denote a subset of links in $\mathcal{N}$. Let $\mathcal{V}_{\mathcal{T}}$ denote the subset of nodes are the ends of $\mathcal{T}$, then $\mathcal{N}_{\mathcal{T}} = (\mathcal{V}_{\mathcal{T}}, \mathcal{T})$ is the induced subgraph of $\mathcal{N}$.

For each of these models, we express the measurement output as a vector $\boldsymbol{y} \in \mathbb{R}^m$ that is related to the delay vector through a measurement matrix $\mathbf{A}$ through matrix multiplication.

III. Key ideas

In this section, we present some key observations and challenges that this paper focuses on. We begin with the observation that there is a high-level connection between the compressive sensing and the network tomography problem. As noted in the previous section, network tomography can be treated as a problem of solving a system of linear equations. With the assumption that the underlying unknown vector is sparse, it becomes a compressive sensing problem [15] [10] [11] in disguise. Then, any compressive sensing algorithm with 0-1 measurement matrix [16] [17] can be applied to recover the vector $\boldsymbol{d}_{\mathcal{V}}$. However, when we go beyond complete



$$\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \boldsymbol{d}_{\mathcal{V}}$$

Figure 1. Node Delay Estimation: For a complete graph with four vertices. We can get any measurements we want since each subgraph of a complete graph is connected. For example, the subgraphs induced by $\{v_1, v_3\}$ (covered by red cycle) and $\{v_1, v_3, v_4\}$ (covered by green cycle) are connected, therefore we get the measurements $[1\ 0\ 1\ 0]\boldsymbol{d}_{\mathcal{V}}$ and $[1\ 0\ 1\ 1]\boldsymbol{d}_{\mathcal{V}}$ respectively.

graphs and node measurements, it is not straightforward to apply compressive sensing directly. Specifically, we face the following challenges:

1. General Networks: If the network is not a complete graph, we cannot get arbitrary measurement matrix. (See Fig. 2.)

2. Edge Delay Estimation: If we consider the network tomography problem for link delay estimation, we cannot freely choose the measurement matrix even if the graph is complete. (See Fig. 5.)

3. Inaccessible Nodes: If there are some constraints on the start and end-points of paths in the network, some measurements may be unavailable. (See Fig. 3.)

4. Other Ensembles: Going beyond 0-1 matrix might give us more efficient compressive sensing algorithm *e.g.*, SHO-FA algorithm [18].

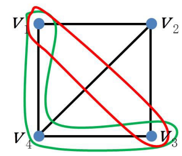


Figure 2. General Network: If the link (v_1, v_3) is removed from the original complete graph, we cannot get the measurement $[1\ 0\ 1\ 0]\boldsymbol{d}_{\mathcal{V}}$ any longer.

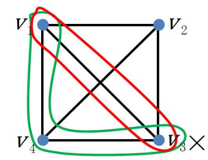


Figure 3. Inaccessible Nodes: If there is some constraint that we can not access to v_3 and v_3 is the destination of the paths for us to get the measurement, then any measurement in Fig. 1 is not available.

Xu *et al.* [9] get around some of these problems by using random walks . One drawback of their approach is that it involves a factor of mixing time $T(\mathcal{N})$ for both the number of measurements and path length. For networks without sufficient connectivity, mixing time may be very large, *e.g.*, cycle graph, $T(\mathcal{N}) = \mathcal{O}(|\mathcal{E}|^2)$. In the following, we propose two news ideas that enable us to circumvent the above problem.

Idea 1: *Cancellation enables selecting disconnected subsets of links and nodes.* The idea here is similar to that used in

[13] where they use the structure called hub to get arbitrary measurement matrix. However, they only consider the node delay model and special graphs which have r-partition. In this paper, we expand this approach to both link delay and node delay models. By considering correlated measurements, we can cancel out the contribution of the undesired links and nodes in a given measurement. Using this approach, we can mimic arbitrary measurements on general graphs. See Fig. 4 as an illustration. One drawback of the cancellation based approach is that if the selected measurement has too many disjoint components, then the number of measurements required is very large. In the example in Fig. 4, the number of cancellations is 2.

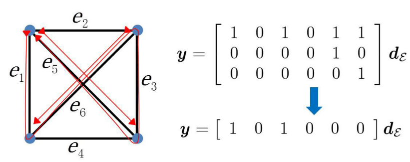


Figure 4. Cancellation: There are three paths in this graph:$\{e_1e_6e_3e_5\}$, $\{e_5\}$ and $\{e_6\}$. Triangles indicate the source and destination of a path. Correspondingly, we can derive three measurements $[1\ 0\ 1\ 0\ 1\ 1]\boldsymbol{d}_{\mathcal{E}}$, $[0\ 0\ 0\ 0\ 1\ 0]\boldsymbol{d}_{\mathcal{E}}$, and $[0\ 0\ 0\ 0\ 1]\boldsymbol{d}_{\mathcal{E}}$. Subtracting the second and the third measurements from the first measurement, we get the measurement $[1\ 0\ 1\ 0\ 0\ 0]\boldsymbol{d}_{\mathcal{E}}$ which cannot be got by just one path.

Idea 2: *Weighted measurements reduce the number of cancellations required and allow us to implement arbitrary integer valued matrices.* The insight here is that if we have two paths along the same set of links, we can assign different weights for each link (or node) on these paths by performing local loops. Specifically, for a given set of weights on a subset of links (or nodes), we construct two measurements - a spanning measurement, and a weighted measurement. The spanning measurement is constructed by finding any path that visits through all the links (or nodes) in the desired subset at least once. The weighted measurement, then follows the same set of edges as the spanning measurement, but visits each link (or node) an additional number of times in accordance with the desired weight for that link (or node). Finally, we subtract the end-to-end delay for the weighted path from that of the spanning path to get an output that is proportional to the output of the corresponding compressive sensing problem (Fig. 7).

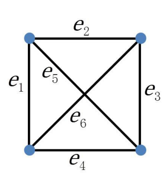


Figure 5. Edge Delay Estimation: We know that we can not get arbitrary measurement by one path even if the graph is complete. (*e.g.*, the measurement $[0\ 0\ 0\ 0\ 1\ 1]\boldsymbol{d}_{\mathcal{E}}$ cannot be got since there is no path just visiting e_5 and e_6.)

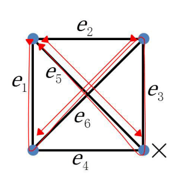


Figure 6. Inaccessible Nodes: The second measurement in Fig. 4 cannot be got since the v_3 which is the destination of the second path is not accessible (the same node identifier in Fig. 1).

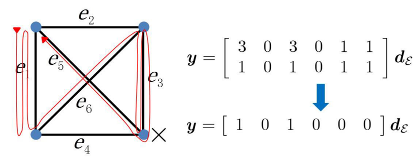


Figure 7. Cancellation using weighted measurements: To get the measurement $[1\ 0\ 1\ 0\ 0\ 0]\boldsymbol{d}_{\mathcal{E}}$, we design the paths as follows. First, we just follow the path $\{e_1e_6e_3e_5\}$, we get the measurement $[1\ 0\ 1\ 0\ 1\ 1]\boldsymbol{d}_{\mathcal{E}}$. Second, when visiting e_1 and e_3 for the first time, the probe does one more local loop for both links to get the measurement $[3\ 0\ 3\ 0\ 1\ 1]\boldsymbol{d}_{\mathcal{E}}$. Finally, after getting the difference of these two measurements and divide the result by 2, we achieve the objective. Note that 1) Only one cancellation is required. 2) Even if v_3 is inaccessible, we still can achieve the two target measurements. 3) If we following the same process except that we do additional local loop for e_1 in the second step (e_1 is visited 5 times), we can get the measurement $[2\ 0\ 1\ 0\ 0\ 0]\boldsymbol{d}_{\mathcal{E}}$. Controlling the number of local loops allows us to implement other ensembles of measurement matrix for compressive sensing.

IV. Main Results

Theorem 1 (Compressive sensing via integer matrices)**.** *Let $M \in \mathbb{Z}^+$. For some $m = \mathcal{O}(k\lceil \log n / \log M \rceil)$, the ensemble of $\mathbb{Z}_M$-valued matrices $\{\mathbf{A}_{m\times n}\}$ designed in Section V and the* SHO-FA-INT *reconstruction algorithm has the following properties:*

1) *Given $(\mathbf{A}_{m\times n}, \mathbf{A}_{m\times n}\boldsymbol{d})$ as input, where $\boldsymbol{d}$ is an arbitrary k-sparse vector in $\mathbb{R}^n$,* SHO-FA-INT *outputs a vector $\hat{\boldsymbol{d}} \in \mathbb{R}^n$ that equals $\boldsymbol{d}$ with probability at least $1 - \mathcal{O}(1/k)$. Here, the error probability is derived from the probability distribution of $\mathbf{A}_{m\times n}$ over the ensemble $\{\mathbf{A}_{m\times n}\}$.*
2) SHO-FA-INT *requires $\mathcal{O}(k\lceil \log n / \log M \rceil)$ arithmetic operations.*
3) *The expected number of non-zero entries in each row of $\mathbf{A}_{m\times n}$ is $\mathcal{O}(n/k)$.*

Theorem 2 (Network tomography for link congestion)**.** *Let $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ be an undirected network of diameter D such that at most k have unknown non-zero link delays. Let $M \in \mathbb{Z}^+$ Then, the* FRANTIC *algorithm described above has the following properties:*

1) FRANTIC *requires $\mathcal{O}(k\lceil \log |\mathcal{E}| / \log M \rceil)$ measurements.*
2) *For every edge delay vector $\boldsymbol{d}_{\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|}$,* FRANTIC *outputs $\hat{\boldsymbol{d}}_{\mathcal{E}}$ that equals $\boldsymbol{d}_{\mathcal{E}}$ with probability $1 - \mathcal{O}(1/k)$.*
3) *The* FRANTIC *reconstruction algorithm requires $\mathcal{O}(k\lceil \log |\mathcal{E}| / \log M \rceil)$ arithmetic operations.*
4) *The number of links of $\mathcal{N}$ traversed by each test measurement packet in* FRANTIC *is $\mathcal{O}(D|\mathcal{E}|/k)$ and the total number of hops for each packet is $\mathcal{O}(DM|\mathcal{E}|/k)$.*

Theorem 3 (Network tomography for node congestion)**.** *Let $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ be an undirected network of diameter D such that at most k have unknown non-zero node delays. Let $M \in \mathbb{Z}^+$ Then, the* FRANTIC *algorithm described above has the following properties:*

1) FRANTIC *requires $\mathcal{O}(k\lceil \log |\mathcal{V}| / \log M \rceil)$ measurements.*
2) *For every edge delay vector $\boldsymbol{d}_{\mathcal{V}} \in \mathbb{R}^{|\mathcal{V}|}$,* FRANTIC *outputs $\hat{\boldsymbol{d}}_{\mathcal{V}}$ that equals $\boldsymbol{d}_{\mathcal{V}}$ with probability $1 - \mathcal{O}(1/k)$.*

3) *The* FRANTIC *reconstruction algorithm requires* $\mathcal{O}(k\lceil \log|\mathcal{V}|/\log M\rceil)$ *arithmetic operations.*
4) *The number of links of* $\mathcal{N}$ *traversed by each test measurement packet in* FRANTIC *is* $\mathcal{O}(D|\mathcal{V}|/k)$ *and the total number of hops for each packet is* $\mathcal{O}(DM|\mathcal{V}|/k)$.

Definition 1 (Steiner Tree)**.** *Let* $\mathcal{S} \subseteq \mathcal{V}$. *We say that* $\mathcal{T} \subseteq \mathcal{E}$ *is a Steiner Tree for* $\mathcal{S}$ *if* $\mathcal{T}$ *has the least number of edges among all subsets of* $\mathcal{E}$ *that form a connected graph that is incident on every* $v \in \mathcal{S}$. *Let* $L(\mathcal{S})$ *be the length of a Steiner Tree for* $\mathcal{S}$.

For every $s \in \mathbb{Z}^+$, let $L^*(s) \triangleq \max\{L(\mathcal{S}) : \mathcal{S} \subseteq \mathcal{V}, |\mathcal{S}| \leq s\}$. Note that, in general, $L^*(s) \leq Ds$. Further, in many graphs of practical interest, $L^*(s) \ll Ds$. For example, in a line graph with n vertices, $L^*(s)$ is at most n, while Ds maybe as large as $\mathcal{O}(ns)$. Using this observation, we may further improve the performance guarantee of our algorithm. The following theorem makes this more precise.

Theorem 4 (Network tomography using Steiner Trees)**.** *Let* $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ *be an undirected network such that at most* k *have unknown non-zero node delays. Let* $M \in \mathbb{Z}^+$ *Then, the number of measurements, probability of correct reconstruction, and number of arithmetic operations required for the* FRANTIC *algorithm described in Section VI-C have the same guarantees as in Theorem 3. Further, the number of links of* $\mathcal{N}$ *traversed by each measurement is at most* $L^*(|\mathcal{V}|/k)$ *and the total number of hops is* $\mathcal{O}(ML^*(|\mathcal{V}|/k))$.

V. SHO-FA-INT Algorithm for Compressive Sensing

In this section, we describe a compressive sensing algorithm SHO-FA-INT that differs from the original SHO-FA algorithm in the following important way – the non-zero entries of the sensing matrix $\mathbf{A}$ are positive integers less than or equal to some $M \in \mathbb{N}$. [3] We first describe an ensemble of bipartite graphs $\{\mathcal{G}_{n,\mu}\}_{n,\mu\in\mathbb{N}}$. Next, for each graph $\mathcal{G}_{n,\mu}$ drawn from this ensemble, we construct a measurement matrix $\mathbf{A}_{R\mu\times n}$, where $R = \lceil \log n/\log M\rceil$ is. Next, we briefly outline the reconstruction algorithm for the above measurement procedure. We refer the reader to [19] for the a detailed description and proofs.

Measurement Graph: Let $\{\mathcal{G}_{n,\mu}\}_{n,\mu\in\mathbb{N}}$ be an ensemble of left-regular bipartite graphs drawn according to the following process. Let $\mathcal{G}_{n,\mu}$ be a bipartite graph with left vertex set $\{1, 2, \ldots, n\}$ and right vertex set $\{1, 2, \ldots, \mu\}$. For each left vertex $j \in \{1, 2, \ldots, n\}$, we pick three right uniformly at random without replacement from the set of right vertices $\{1, 2, \ldots, \mu\}$.

Measurement Design: Let $\zeta(\cdot)$ be the Riemann zeta function. Let $R \in \mathbb{N}^+$ such that $M^R/\zeta(R) \geq 3n$ and let $[M]$ denote the set $\{1, 2, \ldots, M\}$. Given the graph $\mathcal{G}_{n,\mu}$, we design a $R\mu \times n$ measurement matrix $\mathbf{A}(= \mathbf{A}_{R\mu\times n})$ as follows. First, we partition the rows of $\mathbf{A}$ into μ groups of rows, each consisting of R consecutive rows as follows. Let $a_{ij}^{(r)}$ be the r-th row entry in the j-th column of the i-th group and let $\boldsymbol{a}_{ij} = [a_{ij}^{(1)} a_{ij}^{(2)} \ldots a_{ij}^{(R)}]^T$. First, for each (i, j) not in $\mathcal{G}_{n,\mu}$, we set $\boldsymbol{a}_{ij} = 0^R$. Next, for each edge (i, j) in the graph $\mathcal{G}_{n,\mu}$, we set $\boldsymbol{a}_{ij}$ by uniform sampling without replacement from the set $\mathcal{C} \triangleq \{[c_1, c_2, \ldots, c_R]^T \in (\mathbb{Z}_M)^R : \gcd(c_1, c_2, \ldots, c_R) = 1\}$. [1, Lemma 6] shows that, $M^R/\zeta(R) \leq |\mathcal{C}| \leq M^R$. Since the total number of edges in $\mathcal{G}_{n,\mu}$ is $3n$, the assumption that $M^R/\zeta(R) \geq 3n$ ensures that such a sampling is possible.

The output of the measurement is a $R\mu$-length vector $\boldsymbol{y} = \mathbf{A}\boldsymbol{d}$. Again, we partition $\boldsymbol{y}$ into μ groups of R consecutive rows each, and denote the i-th sub-vector as $\boldsymbol{y}_i$. Note each $\boldsymbol{y}_i \in \mathbb{R}^R$ follows the relation $\boldsymbol{y}_i = [a_{i1} a_{i2} \ldots a_{in}]\boldsymbol{d}$.

Reconstruction algorithm: The decoding process is essentially equivalent to the "peeling process" to find 2-core in uniform hypergraph [20], [21]. The decoding takes place over $\mathcal{O}(k)$ iterations. In each iteration, we find one non-zero undecoded d_j with a constant probability after locating a right node that is connected to exactly one non-zero left node. After decoding the non-zero d_j for the current iteration, we cancel out the contribution of d_j from all measurements and proceed iteratively. Since the decoding algorithm parallels the algorithm presented in [1], we omit the details here and refer the reader to [19] for a description and complexity analysis.

VI. The FRANTIC Algorithm for Network Tomography

A. Link Delay Estimation

We define a *path* $\boldsymbol{P}$ of length T over the network $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ as a sequence $(e_1, e_2, \ldots, e_T) = ((v_1, v_2), (v_2, v_3), \ldots (v_T, v_{T+1}))$ such that $e_t = (v_t, v_{t+1}) \in \mathcal{E}$ for $t = 1, 2, \ldots, T$. For a given path $\boldsymbol{P}$, we define the multiplicity $W(\boldsymbol{P}, e)$ of a link $e \in \mathcal{E}$ as the number of times $\boldsymbol{P}$ visits e. Let $\Delta(\boldsymbol{P})$ be the end-to-end delay for a path $\boldsymbol{P}$

Definition 2 ($\boldsymbol{w}$-spanning measurement)**.** *Given a measurement weight vector* $\boldsymbol{w} = [w_1 w_2 \ldots w_{|\mathcal{E}|}]$, *and a* $\boldsymbol{w}$-spanning measurement *is a path* $\boldsymbol{P} = (e_1, e_2., \ldots e_T)$ *in the network* $\mathcal{N}$ *such that* $\boldsymbol{P}$ *visits each* e *in* $\{e : w_e \neq 0\}$ *at least once.*

Definition 3 ($(\boldsymbol{w}, \boldsymbol{P})$-weighted measurement)**.** *Given a measurement weight vector* $\boldsymbol{w}$ *and a* $\boldsymbol{w}$*-spanning measurement* $\boldsymbol{P} = (e_1, e_2, \ldots e_T)$, *a* $(\boldsymbol{w}, \boldsymbol{P})$-weighted measurement *is a path* $\boldsymbol{Q} = (e'_1, e'_2, \ldots e'_H)$ *in the network* $\mathcal{N}$ *such that* $W(\boldsymbol{Q}, e) = W(\boldsymbol{P}, e) + 2w_e$ *for each link e.*

Observe that the end-to-end delay for a $\boldsymbol{w}$-spanning measurement $\boldsymbol{P}$ is equal to $\Delta(\boldsymbol{P}) = \sum_{e\in\mathcal{E}} W(\boldsymbol{P}, e)d_e$, and that for a $(\boldsymbol{w}, \boldsymbol{P})$-weighted measurement is equal to

$$\Delta(\boldsymbol{Q}) = \sum_{e\in\mathcal{E}} W(\boldsymbol{Q}, e)d_e = \Delta(\boldsymbol{P}) + 2\sum_{e\in\mathcal{E}} w_e d_e. \tag{1}$$

Proof of Theorem 2: Let $n = |\mathcal{E}|$. Let $\mathbf{A}$ be a $R\mu \times n$ matrix drawn from the ensemble of Section V, where $R = \mathcal{O}(\lceil \log n/\log M\rceil)$.

Measurement Design: Let $\boldsymbol{a}(i) = [a_{i1} a_{i2} \ldots a_{in}]$ be the i-th row of $\mathbf{A}$. We design the following two network measurements that are related to the row $\boldsymbol{a}(i)$.

[3] In [18], the non-zero entries of $\mathbf{A}$ are chosen to be unit norm complex numbers.

$\boldsymbol{a}(i)$-spanning measurement: We first construct an $\boldsymbol{a}(i)$-spanning measurement $\boldsymbol{P}(i)$ by picking the links in $\{e : \boldsymbol{a}(i) \neq 0\}$ one-by-one and finding a path from one link to another. By the definition of the diameter of the graph, there exists a path of length at most D between any pair of links. Therefore, there exists a path of length $\mathcal{O}(Dn/k)$ that covers all the $\mathcal{O}(n/k)$ vertices that have non-zero components in $\boldsymbol{a}(i)$.

$(\boldsymbol{P}(i), \boldsymbol{a}(i))$-weighted measurement: Given the T-length path $\boldsymbol{P}(i) = ((v_1, v_2), (v_2, v_3), \ldots, (v_T, v_{T+1}))$, we construct a path $\boldsymbol{Q}(i) = (e'_1, e'_2 \ldots, e'_{T'})$ of length $T' = T + 2\sum_{e\in\mathcal{E}} a_e(i)$ as follows. Let $e'_1 = (v_1, v_2)$. If $a_{(v_1,v_2)}(i) \neq 0$, we traverse the edge (v_1, v_2) an additional $2a_{(v_1,v_2)}(i)$ times by going in the forward direction, *i.e.* on (v_1, v_2), and the reverse direction, *i.e.* on (v_2, v_1), an additional $a_{(v_1,v_2)}(i)$ times each. Thus, for $\tau = 1, 3, 5, \ldots, 2a_{(v_1,v_2)}(i) + 1$, we set $e'_\tau = (v_1, v_2)$ and for $\tau = 2, 4, \ldots, 2a_{(v_1,v_2)}(i)$, we set $e'_\tau = (v_2, v_1)$. Next, if $v_3 = v_1$, *i.e.,* we have already visited e_2, we traverse the link we traverse the link e_2 once more, else we traverse it $a_{(v_2,v_3)}(i) + 1$ times in the forward direction and $a_{(v_2,v_3)}(i)$ times in the reverse direction, *i.e.*, for $\tau = 2a_{(v_1,v_2)}(i) + 2, 2a_{(v_1,v_2)}(i) + 4, \ldots, 2a_{(v_1,v_2)}(i) + 2a_{(v_2,v_3)}(i) + 2$, we set $e'_\tau = (v_2, v_3)$ and for $\tau = 2a_{(v_1,v_2)}(i) + 3, 2a_{(v_1,v_2)}(i) + 5, \ldots, 2a_{(v_1,v_2)}(i) + 2a_{(v_2,v_3)}(i) + 1$, we set $e'_\tau = (v_3, v_2)$. We continue this process for each link (v_t, v_{t+1}) in the path $\boldsymbol{P}(i)$, , *i.e.*, if (v_t, v_{t+1}) has been visited already in either the forward or reverse direction by $\boldsymbol{Q}(i)$, we add it to $\boldsymbol{P}(i)$ only once, else, we traverse it an additional $a_{(v_t,v_{t+1})}(i)$ times in each direction. Therefore, $\boldsymbol{Q}(i)$ visits every edge $e \in \mathcal{E}$ a total of $2a_e(i)$ times more than $\boldsymbol{P}(i)$ does.

Reconstructing $\boldsymbol{d}_{\mathcal{E}}$: Next, we measure the end-to-end delays for the paths $\boldsymbol{P}(i)$ and $\boldsymbol{Q}(i)$ for each $i = 1, 2, \ldots, R\mu$ and let $y_i = (\Delta(\boldsymbol{Q}(i)) - \Delta(\boldsymbol{P}(i)))/2$. From equation (1), it follows that $y_i = \sum_{e\in\mathcal{E}} a_{ie} d_e$. Note that this exactly equals the output of a compressive sensing measurement with $\boldsymbol{d}$ as the input vector, $\mathbf{A}$ as the measurement matrix, and $\boldsymbol{y}$ and the measurement output vector. Using this observation, we input the vector $\boldsymbol{y}$ to the SHO-FA-INT algorithm to correctly reconstruct $\boldsymbol{d}$ with probability $1 - \mathcal{O}(1/k)$. The guarantees on the decoding complexity follow from the decoding complexity of the SHO-FA-INT algorithm and that on the total number of hops follows by noting that each link in a measurement path may be visited at most $2M$ times. ■

B. Node Delay Estimation

The measurement design and the decoding algorithm for node delay estimation proceeds in a similar way to the link delay estimation algorithm of Section VI-A. The difference here is that instead of assigning weights to links in a path, our design assigns weights to nodes in a path by visiting each node repeatedly. We skip the proof of Theorem 3 here as it essentially follows from the technique used in the proof of Theorem 2.

C. Reducing path lengths using Steiner Trees

To give a bound on the length of a spanning path $\boldsymbol{P}(\cdot)$ in Section VI-A, we argued that given an arbitrary subset of links, we can for a connected path by traversing from one link to another by adding at most D intermediate links. However, we note that it suffices to find a Steiner Tree that passes through all links specified by a given row of the measurement matrix $\mathbf{A}$. Thus, in general, we can upper bound the number of nodes traversed by each node delay measurement to be $\mathcal{O}(L^*(|\mathcal{V}|/k))$. This proves the assertion of Theorem 4.

References

[1] M. Bakshi, S. Jaggi, S. Cai, and M. Chen, "SHO-FA: Robust compressive sensing with order-optimal complexity, measurements, and bits," *ArXiv.org eprint archive*, 2012, arXiv:1207.2335v1 [cs.IT].

[2] Y. Vardi, "Network tomography: Estimating source-destination traffic intensities from link data," *Journal of the American Statistical Association*, pp. 365–377, 1996.

[3] T. Bu, N. Duffield, F. Presti, and D. Towsley, "Network tomography on general topologies," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 30, no. 1, 2002, pp. 21–30.

[4] Y. Chen, D. Bindel, H. Song, and R. Katz, "Algebra-based scalable overlay network monitoring: algorithms, evaluation, and applications," *Networking, IEEE/ACM Transactions on*, vol. 15, no. 5, pp. 1084–1097, 2007.

[5] J. Kleinberg, "Detecting a network failure," in *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, 2000, pp. 231–239.

[6] H. Nguyen and P. Thiran, "Using end-to-end data to infer lossy links in sensor networks," in *Proc. IEEE Infocom*, 2006, pp. 1–12.

[7] Y. Zhao, Y. Chen, and D. Bindel, "Towards unbiased end-to-end network diagnosis," *Networking, IEEE/ACM Transactions on*, vol. 17, no. 6, pp. 1724–1737, 2009.

[8] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, "Network tomography: recent developments," *Statistical science*, pp. 499–517, 2004.

[9] W. Xu, E. Mallada, and A. Tang, "Compressive sensing over graphs," in *INFOCOM, 2011 Proceedings IEEE*, april 2011, pp. 2087 –2095.

[10] E. J. Candès, J. K. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information." *IEEE Transactions on Information Theory*, pp. 489–509, 2006.

[11] D. L. Donoho, "Compressed sensing." *IEEE Transactions on Information Theory*, pp. 1289–1306, 2006.

[12] M. Firooz and S. Roy, "Network tomography via compressed sensing," in *GLOBECOM 2010, 2010 IEEE Global Telecommunications Conference*, 2010, pp. 1–5.

[13] M. Wang, W. Xu, E. Mallada, and A. Tang, "Sparse recovery with graph constraints: Fundamental limits and measurement construction," in *INFOCOM, 2012 Proceedings IEEE*, march 2012, pp. 1871 –1879.

[14] M. Cheraghchi, A. Karbasi, S. Mohajer, and V. Saligrama, "Graph-constrained group testing," *Information Theory, IEEE Transactions on*, vol. 58, no. 1, pp. 248 –262, jan. 2012.

[15] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5406 –5425, dec. 2006.

[16] R. Berinde, A. Gilbert, P. Indyk, H. Karloff, and M. Strauss, "Combining geometry and combinatorics: A unified approach to sparse signal recovery," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, sept. 2008, pp. 798 –805.

[17] W. Xu and B. Hassibi, "Efficient compressive sensing with deterministic guarantees using expander graphs," in *Information Theory Workshop, 2007. ITW '07. IEEE*, sept. 2007, pp. 414 –419.

[18] M. Bakshi, S. Jaggi, S. Cai, and M. Chen, "Sho-fa: Robust compressive sensing with order-optimal complexity, measurements, and bits," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, oct., 2012.

[19] S. Cai, M. Bakshi, S. Jaggi and M. Chen, "Frantic: A fast reference-based algorithm for network tomography via compressive sensing," *Technical Report*, 2013, available at http://personal.ie.cuhk.edu.hk/~cs010/files/ISIT13.pdf.

[20] M. T. Goodrich and M. Mitzenmacher, "Invertible bloom lookup tables," *ArXiv.org e-Print archive, arXiv:1101.2245 [cs.DB]*, 2011.

[21] M. Molloy, "Cores in random hypergraphs and boolean formulas," *Random Struct. Algorithms*, vol. 27, no. 1, pp. 124–135, Aug. 2005.