Multicast vs. Unicast for Loss Tomography on Tree Topologies

Chang Liu†, Ting He‡, Ananthram Swami§, Don Towsley†, Theodoros Salonidis†, Andrei Iu. Bejan*, and Paul Yu§
†University of Massachusetts, Amherst, MA, USA. {cliu, towsley}@cs.umass.edu,
‡IBM T. J. Watson Research Center, Yorktown Heights, NY, USA. {the, tsaloni}@us.ibm.com,
§Army Research Laboratory, Adelphi, MD, USA. {ananthram.swami, paul.l.yu}.civ@mail.mil,
*University of Cambridge, Cambridge, UK. aib29@cam.ac.uk

Abstract—Loss tomography using multicast and unicast measurements have been investigated separately. In this paper we compared the performance of loss tomography using multicast and unicast on tree structures. We proved identifiability of unicast measurements on tree structures with no 2-degree nodes. To theoretically compare multicast and unicast, we built an observation model for multicast on trees and developed expressions for calculating the Cramér-Rao bound. We applied measurement design for unicast in trees and developed a simpler solution than our earlier work. Using a packet level simulator, we evaluated and compared the per link MSE of multicast and unicast under varying parameter settings include link weights, link loss rate distribution and size of the tree. The results show that in contrast to general belief that multicast outperforms unicast, unicast can outperform multicast under tight constraint on probing overhead, especially in terms of a weighted average of per-link MSEs. Meanwhile, multicast achieves more consistent performance wrt varying link success distribution or tree size.

I. INTRODUCTION

In complex networks (e.g., MANET-cellular hybrid networks, coalition networks), global state of all the networked components (e.g., links) are important for network management. The methodology to infer the internal state such as link loss rate by measuring external performance of the network, is called network tomography [5]. The challenge of network tomography is that since the measurements are generally functions of the states of multiple components, one has to “invert” these functions. By modeling the performance of each link as a random variable with a (partially) unknown probability distribution, one can apply statistical techniques to estimate the parameter of this distribution from path measurements [4], [7], [14], [11]. The most adopted and investigated measurement methods in existing works are unicast measurement [14] and multicast measurement [3], [2], [7], [11]. Tomography with unicast gathers independent measurements on multiple end-to-end paths in a network via unicast probes and inverts path performance metrics to estimate corresponding link performance metrics. Tomography with multicast, on the other hand, gathers correlated measurements along multicast trees between each source and its corresponding receivers via multicast probes. In networks that do not directly support multicast communications, multicast-like measurements can be obtained by sending batches of back-to-back unicast probes, referred to as correlated unicast, so that probes in the same batch experience similar performance on the same link [6].

For each of the above probing methods, there have been studies on how to allocate probes across different paths/trees so that the overall information about the link parameters of interest can be maximized [8], [15], [9]. There is, however, a lack of understanding in when a certain method is preferable against others.

In this paper, we aims at giving initial understanding to the strengths and weaknesses of each probing method when applied to infer link loss rates in networks with tree topologies. Tree topology represents a case of special interest in network tomography. Besides its simplicity, tree topology is shown to approximate latency and bandwidth in the Internet [13], and most tomography-based topology discovery methods generate logical topologies that are trees [10]. Given a network spanned by a single multicast tree, we ask the following questions: (i) Can unicast consistently estimate link loss rates in the tree? (ii) If so, how should we allocate probes among different paths? (iii) How do different probing methods compare in terms of the accuracy of estimated link loss rates, and how does the comparison depend on parameters such as the amount of probing, the size of the tree, and the values of link loss rates?

A. Related Work

In existing works, statistical tomography models each link metric as a random variable with a (partially) unknown probability distribution, and applies various estimation techniques to infer the distribution from path measurements. When supported, multicast-based probing has been proposed to estimate link parameters from measurements at multicast receivers [3], [2], [7], [11]. Specifically, [3] derives a maximum likelihood estimator to infer link loss rates from packet losses observed at receivers of a single-source multicast tree, which is later extended to use losses observed from multiple trees in [2]. Analogous results have been obtained for delays, where packet delays observed at multicast receivers are used to infer variances or distributions of delays at internal links [7], [11]. Multicast probing has the benefit of uniquely identifying metrics of link segments between branching points [2], but it also has the limitation of requiring network layer support.
for multicast communications. To relax this limitation, [6] has proposed a technique to emulate multicast using back-to-back unicast probes referred to as correlated unicast in this paper, under the assumption that unicast probes sent sufficiently close to each other (in time) on a given path will experience the same performance on each link of the path.

The theory of experiment design for general statistical inference casts the problem as an optimization of a set of design objectives that capture various aspects of estimation accuracy [1]. The approach has recently been applied to design experiments for network tomography. Under multicast or correlated unicast, [8], [15] have proposed to measure the quality of an experiment design by appropriate functions of the Fisher Information Matrix (FIM) and to design probing experiments such that certain performance criteria based on the FIM can be optimized (A-optimality in [8], D-optimality in [15]). These solutions either rely on numerical solvers [15] or a coarse approximation that ignores off-diagonal elements of the FIM [8]. This approach has recently been extended to unicast probing, where closed-form solutions are derived to optimally allocate probes among unicast paths under the criterion of D-optimality of A-optimality [9].

B. Summary of Contributions

In investigating multicast and unicast on tree structures and comparing their performance, our specific contributions are:

1) We prove the identifiability of all links in trees without degree 2 nodes using unicast between leaves, and proposed a path construction algorithm to achieve identifiability.

2) We derive a closed-form expression for optimal probe allocation for unicast.

3) We define explicit formulas for evaluating fisher information and the Cramér-Rao bound (CRB) for both unicast and multicast.

4) We use packet level simulation to evaluate the performance of unicast, multicast and correlated unicast under varying system parameters including link weights, link success rate distribution and tree size. Besides confirming that multicast always outperforms correlated unicast, our results show that unicast outperforms multicast under tight probing budget in terms of total number of hops traversed by probes, especially when links have heterogeneous weights, meanwhile multicast and correlated unicast are more robust than unicast against different link success distributions and different sizes of tree.

II. LOSS TOMOGRAPHY ON TREES

A. Network Model

Let \( T = (V, L) \) denote a directed tree with nodes \( V \) and links \( L \). Let \( s \in V \) denote the source and \( R \subset V \) the receivers. Following [3], we assume that the other nodes (referred to as branching nodes) have degrees of at least three. Without loss of generality, we label the nodes so that \( s = 0 \) and \( R = \{ |L| - |R| + 1, \ldots, |L| \} \). We label the links so that link \( i \) is the link leading to node \( i \) (from node 0). We refer to node 0 as the root of the tree and \( R \) as the leaves. Let \( f(i) \) denote the parent of node \( i \) in the tree, \( d(i) \) the set of children of node \( i \), and \( s(i) \) the set of siblings of node \( i \) (i.e., \( s(i) = \{ j \in d(f(i)) : j \neq i \} \)). Losses on link \( l \in L \) are decided by a Bernoulli process with an (unknown) loss probability \( 1 - \theta_l \) (success probability \( \theta_l \)). We assume that losses are independently from link to link.

B. Observation model and MLE of unicast

Let \( P \) denote the set of paths the monitoring system can inject probes on and observe the end-to-end performance. Link success rates are then inferred from unicast measurements on paths. Following the definition in [9], the measurement matrix is a \( |P| \times |L| \) matrix \( A := [A_{y,l}] \) defined by \( P \), where \( A_{y,l} = 1 \) if link \( l \) is on path \( p_y \) and \( A_{y,l} = 0 \) otherwise. We use the same probabilistic design model as in [9], where each probe is sent over path \( p_y \) randomly selected from \( P \), with probability \( \phi_y \). Here \( \phi := (\phi_y)_{y=1}^{|P|} \), satisfying \( \phi_y \geq 0 \) and \( \sum_{y=1}^{|P|} \phi_y = 1 \), is a design parameter. Let \( y \) be the selected path for a probe and \( x \) an indicator that the probe successfully reaches its destination. Then the observation model becomes:

\[
f(x, y; \theta, \phi) = \phi_y (\prod_{l \in p_y} \theta_l)^x (1 - \prod_{l \in p_y} \theta_l)^{1-x}.
\]

We use the MLE proposed in [9] for unicast as follows,

**Proposition 1.** [9] If the measurement matrix \( A \) has full column rank and there is at least one successful probe per path, then the MLE for loss tomography equals:

\[
\hat{\theta} = \exp \left( (A^T A)^{-1} A^T \log \hat{\alpha} \right),
\]

where \( \hat{\alpha} \) is the vector of empirical path success rates.

C. Observation model and MLE of multicast

For multicast probing, the observation model is more complicated. Let \( X = (X_i)_{i \in V} \) denote the indicators for a multicast probe to reach individual nodes in the tree, \( X_i = 0 \) is the probe doesn’t reach leaf \( i \) and \( X_i = 1 \) if it does. \( X_R = (X_i)_{i \in R} \) denote the subset of indicators for leaf nodes, \( I := V \setminus R \) the set of internal nodes in the tree, and \( X_I = (X_i)_{i \in I} \) the subset of indicators for internal nodes (including the root, for which \( X_0 := 1 \)) ; note that only \( X_R \) is observable.

Since only a fraction of the \( X_i \)’s are observable, the likelihood function is a marginal conditional distribution:

\[
p(X_R|\theta) = \sum_{X_I} p(X|\theta) \tag{3}
\]

where \( p(X|\theta) \) is the joint conditional distribution of all \( X_i \)’s for given link success rates \( \theta \). Each \( X_i, i \in I \) only depends on it’s parent \( X_{f(i)} \), so we can write the joint distribution as:

\[
P(X|\theta) = \prod_{k \in V} P(X_k|X_{f(k)}, \theta) \tag{4}
\]

1For ease of presentation, we use \( g(z) \) to denote the vector obtained by applying a scalar function \( g(\cdot) \) to each element of a vector \( z \).
Note that because of the tree structure, each $X_i, i \in V$ appears only once in (4).

$$P(X_k|X_{f(k)}, \theta)$$ is the observation model at each node $k$ given the observation at its parent node $f(k)$. If $X_{f(k)} = 1$, then node $k$ observes $X_k = 1$ or 0 with probability $\theta_k$ or $\theta_k := 1 - \theta_k$ (recall that $\theta_k$ is the success probability of link $k$ that connects nodes $f(k)$ and $k$); if $X_{f(k)} = 0$, then $X_k = 0$. Therefore, we have

$$p(X_k|X_{f(k)}, \theta) = X_{f(k)}(X_k\theta_k + (1 - X_k)\theta_k) + (1 - X_{f(k)})(1 - X_k).$$

Substituting (5) into (4) and then into (3) gives an explicit expression of the likelihood function of multicast probing.

An indirect form of the MLE is provided in [3] as follows. For each node $k$, define $\gamma_k$ as the probability that any receiver under node $k$ receives a multicast probe, and $a_k$ the probability that a multicast probe reaches node $k$. Then $\theta_k$, $\gamma_k$, and $a_k$ are related as follows:

$$\theta_k = a_k/a_{f(k)}, \quad k \in V \setminus \{0\}$$

$$a_k = \gamma_k, \quad k \in R$$

$$1 - \gamma_k/a_k = \prod_{j \in \delta(k)} (1 - \gamma_j/a_k), \quad k \notin R.$$

Equations (6–8) jointly define a transformation from $\gamma$ to $\theta$. Note that the empirical value of $\gamma_k$, namely $\hat{\gamma}_k := \text{fraction of multicast probes that are received by at least one of the receivers under node } k$, is directly measurable. Moreover, $\hat{\gamma}_k$ is the MLE of $\gamma_k$. If the transformation from $\gamma$ to $\theta$ is one-to-one, then we can easily obtain the MLE of $\theta$ from $\hat{\gamma}$ by applying the invariance property of MLE. Indeed, this has been shown in [3].

Theorem 2 ([3]). The transformation from $\gamma$ to $\theta$ defined by (6–8) is a bijection. Therefore, $\theta$ defined by substituting $\hat{\gamma}$ into (6–8) is the MLE of $\theta$ under multicast probing.

III. IDENTIFIABILITY AND PATH CONSTRUCTION

In order to compare the two approaches, one based on unicast and the other based on multicast it is important that both methods identify all of the links. Multicast probing is shown to identify all links in trees without degree 2 nodes (i.e., all nodes are either leaves or branching nodes). Below we will show that with suitably constructed paths, unicast probing can achieve the same identifiability using the same set of measurement nodes (source and all receivers in the multicast tree).

We establish the above by specifying a path construction algorithm that achieves identifiability. Consider the algorithm in Algorithm 1. It constructs a set of $|L|$ paths in two steps:

1. select all paths from root to leave (lines 2–3);
2. for each branching node $v$, select a path between two arbitrary leaves under different children of $v$ (lines 4–7).

Here we use the notation $p_{v \rightarrow w}$ to denote the (unique) path in the tree between nodes $v$ and $w$. For example, for the multicast tree in Fig. 1, Step (1) constructs paths $p_{0 \rightarrow 4}$, $p_{0 \rightarrow 5}$, $p_{0 \rightarrow 6}$, and $p_{0 \rightarrow 7}$, and Step (2) constructs paths $p_{4 \rightarrow 5}$, $p_{4 \rightarrow 7}$, and $p_{6 \rightarrow 7}$. Below, we establish the above by specifying a path construction algorithm that achieves identifiability. Consider the algorithm in Algorithm 1. It constructs a set of $|L|$ paths in two steps:

1. select all paths from root to leaves (lines 2–3);
2. for each branching node $v$, select a path between two arbitrary leaves under different children of $v$ (lines 4–7).

Here we use the notation $p_{v \rightarrow w}$ to denote the (unique) path in the tree between nodes $v$ and $w$. For example, for the multicast tree in Fig. 1, Step (1) constructs paths $p_{0 \rightarrow 4}$, $p_{0 \rightarrow 5}$, $p_{0 \rightarrow 6}$, and $p_{0 \rightarrow 7}$, and Step (2) constructs paths $p_{4 \rightarrow 5}$, $p_{4 \rightarrow 7}$, and $p_{6 \rightarrow 7}$.

**Algorithm 1** Unicast Path Construction for Tree Topology

1. $P \leftarrow \emptyset$
2. for each leaf $v \in R$ do
3. $P \leftarrow P \cup p_{0 \rightarrow v}$
4. for each branching node $v \in V \setminus \{0 \cup R\}$ do
5. select $c_1, c_2 \in \delta(v)$ such that $c_1 \neq c_2$
6. select leaves $l_1$ under $c_1$ and $l_2$ under $c_2$
7. $P \leftarrow P \cup p_{l_1 \rightarrow l_2}$

Fig. 1. Example: unicast paths for identifying links in a multicast tree.

The following theorem establishes that paths constructed by Algorithm 1 identify all links in the tree.

**Theorem 3.** For a tree $T$ with no degree 2 nodes, unicast probing between the monitors identifies all links in the tree if and only if all degree 1 nodes are monitors.

**Proof:** It is easy to see that the condition is necessary, as otherwise the link leading to a non-monitor leaf cannot be measured. The proof of necessity consists of showing that, given the metrics of the paths constructed by Algorithm 1 each equal the sum of the traversed link metrics, allowing us to identify the metrics of all the links in the tree. Since, under the assumption of independent losses for unicast, link/path success rates can be converted to additive metrics by taking logarithm, the result follows.

Let $w_{i,j}$ denote the metric of the path between nodes $i$ and $j$; Given the path construction of Algorithm 1, $w_{i,j}$ can be directly estimated from path performance (e.g., losses) if and only if both $i$ and $j$ are degree-1 nodes (i.e., root or leaves). Suppose we construct paths according to Algorithm 1. The constructed paths have a property that they can identify $w_{0,v}$ for all $v$. In particular, if $v$ is a branching node (e.g., node 1 in Fig. 1), and the path constructed for $v$ in Step (2) is between leaves $v_1$ and $v_2$ (nodes 4 and 7 in Fig. 1), then $w_{0,v} = (w_{0,v_1} + w_{0,v_2} - w_{v_1,v_2})/2$. The metric of link $(i, j)$ is then determined by $w_{i,j} = w_{0,i} - w_{0,i,j}$, assuming node $i$ is closer to node 0 than node $j$.

IV. PERFORMANCE BOUND AND EXPERIMENT DESIGN

Given an observation model $f(O: \theta)$ where $O$ represents the observations, the (per-measurement) FIM wrt $\theta$ is an $|L| \times |L|$ matrix, whose $(i, j)$-th entry is defined by

$$E \left[ \frac{\partial}{\partial \theta_i} \log f(O: \theta) \left( \frac{\partial}{\partial \theta_j} \log f(O: \theta) \right) \right].$$

For unicast $O = \{x, y\}$ in (1), and for multicast $O = X_R$ in (3).

\[ \text{Algorithm 1} \] Unicast Path Construction for Tree Topology

1. $P \leftarrow \emptyset$
2. for each leaf $v \in R$ do
3. $P \leftarrow P \cup p_{0 \rightarrow v}$
4. for each branching node $v \in V \setminus \{0 \cup R\}$ do
5. select $c_1, c_2 \in \delta(v)$ such that $c_1 \neq c_2$
6. select leaves $l_1$ under $c_1$ and $l_2$ under $c_2$
7. $P \leftarrow P \cup p_{l_1 \rightarrow l_2}$

Fig. 1. Example: unicast paths for identifying links in a multicast tree.

The following theorem establishes that paths constructed by Algorithm 1 identify all links in the tree.

**Theorem 3.** For a tree $T$ with no degree 2 nodes, unicast probing between the monitors identifies all links in the tree if and only if all degree 1 nodes are monitors.

**Proof:** It is easy to see that the condition is necessary, as otherwise the link leading to a non-monitor leaf cannot be measured. The proof of necessity consists of showing that, given the metrics of the paths constructed by Algorithm 1 each equal the sum of the traversed link metrics, allowing us to identify the metrics of all the links in the tree. Since, under the assumption of independent losses for unicast, link/path success rates can be converted to additive metrics by taking logarithm, the result follows.

Let $w_{i,j}$ denote the metric of the path between nodes $i$ and $j$; Given the path construction of Algorithm 1, $w_{i,j}$ can be directly estimated from path performance (e.g., losses) if and only if both $i$ and $j$ are degree-1 nodes (i.e., root or leaves). Suppose we construct paths according to Algorithm 1. The constructed paths have a property that they can identify $w_{0,v}$ for all $v$. In particular, if $v$ is a branching node (e.g., node 1 in Fig. 1), and the path constructed for $v$ in Step (2) is between leaves $v_1$ and $v_2$ (nodes 4 and 7 in Fig. 1), then $w_{0,v} = (w_{0,v_1} + w_{0,v_2} - w_{v_1,v_2})/2$. The metric of link $(i, j)$ is then determined by $w_{i,j} = w_{0,i} - w_{0,i,j}$, assuming node $i$ is closer to node 0 than node $j$.

IV. PERFORMANCE BOUND AND EXPERIMENT DESIGN

Given an observation model $f(O: \theta)$ where $O$ represents the observations, the (per-measurement) FIM wrt $\theta$ is an $|L| \times |L|$ matrix, whose $(i, j)$-th entry is defined by

$$E \left[ \frac{\partial}{\partial \theta_i} \log f(O: \theta) \left( \frac{\partial}{\partial \theta_j} \log f(O: \theta) \right) \right].$$

For unicast $O = \{x, y\}$ in (1), and for multicast $O = X_R$ in (3).
The significance of the FIM is that it provides a fundamental bound on the error of unbiased estimators. Specifically, if \( \theta \) is an estimator of \( \theta \) using \( N \) i.i.d. measurements, then the covariance matrix of \( \theta \) satisfies\(^2\) \[ \text{cov}(\theta) \geq \frac{1}{N} I^{-1}(\theta; \phi), \] known as the Cramér-Rao bound (CRB) [12]. In particular if \( \hat{\theta} \) is unbiased, then the MSE in estimating \( \theta_i \), given by \( \text{cov}(\theta)|_{\theta=\hat{\theta}} \), is lower bounded by \( I_{\phi}^{-1}(\theta; \phi)/N \).

A. FIM Based Experiment Design for Unicast

Based on the observation model (1), as shown in [9], the \((i, j)\)-th entry of the FIM for unicast loss tomography is:

\[
I_{i,j}(\theta; \phi) = \sum_{y=1}^{[P]} \phi_y \frac{\alpha_y(\theta)}{\theta_i \theta_j (1 - \alpha_y(\theta))} \mathbb{1}\{i, j \in p_y\}. \tag{10}
\]

where \( \mathbb{1}\{ \} \) is the indicator function.

Based on the FIM, the goal of experiment design is to optimize some function of FIM, which is related to bounding estimation errors, with the design parameter \( \phi \). We leverage the previous results on optimal experiment design [9], where we consider weighted A-optimality, which is to minimize the weighted trace \( \text{Trace}(I^{-1} \cdot \omega) \) with \( \omega = (\omega_{ij})_{i \leq j} \) denoting the link weights, as it directly corresponds to weighted link MSE minimization for unbiased estimators. The previous result states that given a set of paths forming a basis of the link space, the optimal probe allocation is given by \( \phi_y = \frac{\sqrt{A_i(\theta; \omega)}}{\sum_{i=1}^{[L]} \sqrt{A_i(\theta; \omega)}} \), where \( A_i(\theta; \omega) \) for link \( i \) is a function of \( \theta \) and \( \omega \). Our focus is therefore on selecting the optimal basis that optimizes the overall design objective, i.e., the trace of the inverse FIM. Under the optimal probe allocation, this design objective equals \( \sum_{i=1}^{[L]} \omega_k \theta_k^2 \), where the coefficients \( A_i \) implicitly depend on the probing paths. To optimize path construction, we first derive a closed-form expression of \( A_i \), that explicitly depends on decision variables in path construction.

Consider the path construction algorithm Algorithm 1. Let \( p_v \) denote the path associated with node \( v \): if \( v \) is a leaf, \( p_v = p_{0 \rightarrow v} \), if \( v \) is a branching node, \( p_v = p_{i \rightarrow j} \) for the selected leaves \( i \) and \( j \) under different children of \( v \). Note that for a given tree with a given root, the decision variables for this algorithm are \( \{i, j : \forall \text{ branching node } v\} \).

The key in deriving an explicit expression of \( A_i(\theta) \) is to derive an explicit expression of the inverse matrix of \( \pi^{-1} = [b_{k,i}]_{k,i=1}^{[L]} \). Let \( w_{i,j} \) denote the metric of the path segment between nodes \( i \) and \( j \), and \( m_v \) denote the end-to-end metric of path \( p_v \). We have that the metric of link \( k \) equals inner product between the \( k \)-th row of \( \pi^{-1} \) and the vector of path metrics, i.e., \( w_{f(k),k} = \sum_{i=1}^{[L]} b_{k,i} m_i \). Therefore, we can obtain \( b_{k,i} \) by expressing \( w_{f(k),k} \) as a function of \( m_i \)’s. Specifically, we can express the metric of each 0-to-\( v \) path segment as

\[
w_{0,v} = \begin{cases} \frac{m_v}{m_i + m_j - m_v} & \text{if } v \in R, \\ \frac{m_i + m_j - m_v}{2} & \text{if } v \not\in R. \end{cases} \tag{11}
\]

Based on these 0-to-\( v \) path metrics, we can identify link metrics as

\[
w_{f(k),k} = \begin{cases} \frac{m_i + m_j - m_v}{2} & \text{if } k = 1, \\ \frac{m_i + m_j - m_k}{2} - \frac{m_i + m_j - m_k}{2} & \text{if } k > 1, k \not\in R, \\ - \frac{m_i + m_j - m_k}{2} & \text{if } k \in R. \end{cases} \tag{12}
\]

Comparing (12) with the generic formula of \( w_{f(k),k} = \sum_{i=1}^{[L]} b_{k,i} m_i \) gives the value of \( b_{k,i} \). For \( i \leq |L| - |R| \), we have that

\[
b_{k,i} = \begin{cases} -\frac{1}{2} & \text{if } k = i, \\ \frac{1}{2} & \text{if } k \in d(i), \\ 0 & \text{otherwise}. \end{cases} \tag{13}
\]

For \( i > |L| - |R| \), we have that

\[
b_{k,i} = \begin{cases} 1 & \text{if } i \not\in p_{f(i)}, \\ \frac{1}{2} & \text{if } i \in p_{f(i)}, \\ -\frac{1}{2} & \text{if } i \not\in p_k, i \in p_{f(k)}, \\ 0 & \text{otherwise}. \end{cases} \tag{14}
\]

where \( p_v \) means that node \( i \) is on path \( p_v \). For \( i > |L| - |R| \), we have that

\[
A_i(\theta) = \frac{1 - \alpha_i}{\alpha_i} \left( \omega_i \theta_i^2 \left[ \frac{3}{4} \mathbb{1}\{i \not\in p_{f(i)}\} + \frac{1}{4} \sum_{k \in \Phi_i} \omega_k \theta_k^2 \right] \right), \tag{15}
\]

where \( \Phi_i = \{k \neq i : i \in p_k, i \not\in p_{f(k)}\} \cup \{k \neq i : i \not\in p_k, i \not\in p_{f(k)}\} \).

B. FIM and Performance Bound for Multicast

Based on likelihood function 3, we are ready to derive an explicit expression for the FIM for multicast. By definition, the \((i, j)\)-th entry in the FIM equals

\[
I_{i,j}(\theta) = \sum_{X_R} \frac{1}{p(X_R|\theta)} \left[ p(X_R|\theta) \sum_{X_j} \frac{\partial^2}{\partial \theta_i \partial \theta_j} p(X|\theta) \right] - \left( \sum_{X_j} \frac{\partial}{\partial \theta_j} p(X|\theta) \right) \left( \sum_{X_j} \frac{\partial}{\partial \theta_j} p(X|\theta) \right). \tag{18}
\]

Based on the explicit expression of \( p(X|\theta) \) given by (4)
\[ \frac{\partial}{\partial \theta_i} p(X|\theta) = X_{f(i)}(2X_i - 1) \prod_{k \neq i} p(X_k|X_{f(k)}, \theta), \tag{19} \]

and

\[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} p(X|\theta) = \begin{cases} 0, & i = j, \\ X_{f(i)}(2X_i - 1)X_{f(j)}(2X_j - 1) \\ \times \prod_{k \neq i,j} p(X_k|X_{f(k)}, \theta), & i \neq j. \end{cases} \tag{20} \]

Substituting (19, 20) into (18) gives the FIM for multicast probing, and \( I_{ii}^{-1} \) gives the lower bound on the MSE of any unbiased estimator of \( \theta_i \).

\section{V. Performance Evaluation}

We compare the performance of loss tomography based on multicast or unicast probing by packet-level simulations on binary tree topologies. We build a simulator for multicast measurement and inference, using the inference algorithm from [3]. For unicast, we use the same simulator as in [9] and among its three proposed measurement allocation methods we choose to use the one without assuming known link success rates, referred to as ‘Iterative A-optimal design’. To simulate correlated unicast, we send batches of \(|R|\) back-to-back unicast probes per batch, each following the path from the source to a receiver in the multicast tree, assuming the packet loss realizations on each link are the same for all probes in each batch.

We evaluate performance of multicast and unicast on tree structures extended from binary trees. To avoid degree-2 nodes we add a node to the root node of a standard binary tree and consider the added node as root node. Link success rates are randomly generated with a uniform distribution between 0.1 and 1 unless otherwise specified. For unicast, we simulate the iterative A-optimal design by estimating link success rates for every 100 probes, and updating the design parameter \( \phi \) as in Algorithm 2 in [9]. For all topology and link settings, the experiment results are based on simulations of 30 Monte Carlo runs, each consist of 100 of iterations of 100 measurement per iteration.

Though CRB is calculated on the number of probes, it’s not fair to compare the performance of multicast and unicast for the same number of probes because a unicast probe only traverses a path between one pair of degree 1 nodes while a multicast probe traverses the entire tree and thus generates much larger overhead. Thus we use the number of traversed hops as the “cost” of measurement. For multicast, number of hops traversed by each probe is the number of links in the whole tree. For unicast, it’s the length of the probed end-to-end path. For correlated unicast, it’s the sum of path lengths of all root-to-leaf paths in the tree.

\section{A. Convergence Rate}

CRB gives a lower bound on MSE based on information of 1 probe. It decreases with speed \( 1/#\text{probes} \) as number of probes grows large. Figure 2 shows CRB and MSE as number of probes for both multicast and unicast tomography for a 2-leaf tree, where MSE is calculated with 30 Monte Carlo runs of simulation. The link average CRB is 0.24 for multicast and 1.28 for unicast. Multicast has lower CRB and MSE here because each multicast probe gives information of all 3 links in the tree, while unicast only gives information of links in the path that’s probed (each unicast probe traverses 2 links in this case).

\section{B. Impact of link weights}

Unicast measurements has the flexibility to allocate probes unevenly across the network, which intuitively favor the case that only a part of the network are of interest, while multicast measures all the links evenly. We compare performance of all three measurement methods in such case by introducing link weights. Figure 3 shows weighted average MSE agains number of hops a 16-leaf full binary tree with 32 nodes and 31 links. For homogeneous link weights, all links have unit weight. For heterogeneous link weights, we randomly select a link and set its weight as 500 and all the other links have unit weight. When links have homogeneous weights as in Figure 3(a), unicast achieves a lower MSE when the number of hops is small due to its flexibility in sending more probes on paths providing more “information” as measured by FIM. But unicast is outrun by multicast and correlated unicast as number of hops grows larger. When consider heterogeneous weights, the advantage of unicast becomes more significant as its probe allocation takes into account link weights.

\section{C. Impact of link success rate distribution}

Link success rates by default in our simulation is uniformly distributed between 0.1 and 1. We change a fraction of the links to be more reliable and compare the performance difference of the three measurement methods. On a 16-leaf full binary tree we randomly select a fraction of links to be “reliable links”, and set the success rates to be Uniform(0.9, 1). Given a fraction of reliable links, we generate 10 instances of link success rates for all links. As shown in Figure 4, when
the fraction of ‘reliable links’ gets larger, for a fixed and large enough (65000+) number of hops, the MSE of unicast probing becomes smaller, similar effect was also observed in [9]. For multicast and correlated unicast probing, the MSE is largely invariant under changes in the link success rate distribution.

D. Impact of tree size

We increase the size of full binary tree from 3 links (2 leaf nodes) to 63 links (32 leaf nodes) and evaluate the average link MSE. Figure 5 shows MSE over size of tree. For a meaningful comparison among trees of different sizes, we fix the ratio of total hop count and number of links to 2000, so that the total number of hops grows proportionally to the size of tree. With fixed per link overhead, multicast and correlated unicast performs rather robustly as the size of tree scales, while both the median and range of MSE by unicast probing increase as the tree grows larger.

REFERENCES