An Introduction to Network Tomography Techniques

Ritabrata Roy and Wade Trappe

April 2004

Abstract

Network tomography deals with the study of estimating the internal characteristics of a network from its end–point measurements. This paper begins with a quick overview of some of the related work in this field and introduces the problem of determining link success probabilities of a network tree from a sample of measurements at the nodes. Both single–source and multiple–source techniques for estimation of link success probabilities are studied and a method is proposed to obtain these link parameters for a general network tree. the method is illustrated by deriving the probability estimates for a simple three–node network. Finally, the problem is reformulated as a general tomography problem to which standard techniques like Maximum Likelihood Estimation can be applied.

1 Introduction

Network monitoring of large–scale networks, like the Internet, cannot rely on measurements from each and every node to obtain performance parameters like link loss and packet delay. Instead, these internal network characteristics are estimated from end-to-end behavior of point-to-point traffic, and this technique is referred to as *Network Tomography* [15] to highlight its similarity with medical tomography.

Network tomography analyses use logical tree–graphs [2] (in contrast to the *physical* tree graph which allows intermediate nodes to have only one child [4]) to represent one–to–many communication between the root (a single source node) and the leaves (a number of destination nodes) via various internal nodes. Studies have been developed for both unicast [7],[8] and multicast networks [1],[4],[3], although most of the recent work focus almost exclusively on multicast networks.

Network tomography techniques, such as multicast inference of network characteristics (MINC) [1], rely on sending probe traffic through the network periodically and making corresponding measurements at the terminal nodes. Internal network parameters are then inferred by exploiting the correlation in performance observed in multicast receivers— for instance, if a receiver observes a probe, its parent must have observed the probe as well. This intuition may be used to estimate the path delay distribution [8], the internal loss characteristics [4], or the network topology itself [3],[11].

By contrast, the model in this paper discusses a passive [13] monitoring method which does not use a separate probe packet to determine the network characteristics. On the other hand, requiring the intermediate nodes to record every packet passing through them would need enormous amounts of storage space. Instead, each of the intermediate nodes in the network sample the packets passing through them and record information about a fraction of them with a certain *a-priori* probability, the aim being to minimize the amount of data recorded by the intermediate nodes and still obtain the network's internal characteristics to a very high degree of accuracy. Although this model might appear to be at odds with standard network tomography techniques, it will be shown later that the formulation may be reduced to a standard network tomography problem.

The rest of this treatise is arranged as follows: Section 2 discusses a few existing papers that are relevant to the current discussion, and the problem itself is formulated in Section 3. This is then analyzed in Section 4 and comparisons drawn with existing results in Section 5. Finally, the paper concludes with areas of future work, and the derivation of maximum likelihood estimation (MLE) of link success probabilities in the Appendix.

2 Related Work

2.1 Packet Loss Estimation

The intuition behind packet loss estimation is that the arrival of a packet at a given internal node in the tree can be inferred from the packet's arrival at one or more receivers descended from that node [1]. Packet loss is modeled as independent across different links of the tree and between different probes. The loss model associates with each link in the tree, a conditional probability that a packet reaches the terminating node of the link, given that it reaches its parent node. The outcome of *each* probe at *each* receiver is recorded. The link probabilities are inferred by estimators obtained by using the actual frequencies of the outcomes, and it has been shown in [5] that this estimator is the maximum likelihood estimator (MLE) when a sufficiently large number of probes are used. The Appendix introduces a widely used method [5][4] to estimate the success/loss probabilities of each link. A more simplistic analysis to estimate the link probabilities for a general k–node network is carried out in Section 3.

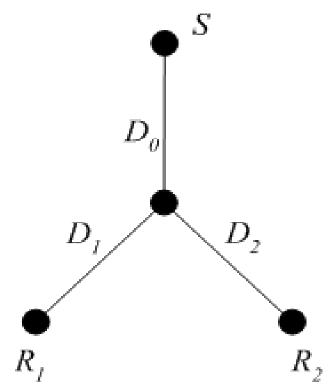

Figure 1: Model of two–leaf tree for calculating variance of internal link delay.

2.2 Internal Delay Estimation

The intuition behind internal delay estimation is that closely time–spaced packets should experience the same delay on each shared link in their path [8], and therefore delay aberrations at different receivers must be caused by delays on the individual links. Thus, associated with each individual link in the network is a probability mass function, PMF, for the delay. The aim of this class of problems is to estimate these probabilities based on the end-to-end measurements.

Assuming that link delays are independent random variables, both spatially (*i.e.*, between different links) and temporally (*i.e.*, between different packets), the delay variance estimation is based on the additive property of time delay across successive links of a route [12]. In the logical multicast topology of Fig. 1, packets are being multicast from the root S to receivers R_1 and R_2. If D_i denotes the random delay on link i, then the source–to–leaf delays in this case would be:

$$x_1 = D_0 + D_1 \tag{1}$$

$$x_2 = D_0 + D_2 \tag{2}$$

$$
\begin{aligned}
\text{Hence,} \quad \operatorname{cov}(x_1, x_2) &= \mathrm{E}[x_1 x_2] - \mathrm{E}[x_1]\mathrm{E}[x_2] \\
&= \mathrm{E}\left[(D_0 + D_1)(D_0 + D_2)\right] - \mathrm{E}[D_0 + D_1]\mathrm{E}[D_0 + D_2] \\
&= \mathrm{E}[D_0^2] + \mathrm{E}[D_0 D_1] + \mathrm{E}[D_0 D_2] + \mathrm{E}[D_1 D_2] - \mathrm{E}^2[D_0] - \mathrm{E}[D_0]\mathrm{E}[D_1] - \mathrm{E}[D_0]\mathrm{E}[D_2] - \mathrm{E}[D_1]\mathrm{E}[D_2] \\
&= \mathrm{E}[D_0^2] - \mathrm{E}^2[D_0] \quad \text{, assuming independence} \\
&= \operatorname{var}(D_0)
\end{aligned}
\tag{3}
$$

Finally, an unbiased estimate of $\operatorname{var}(D_0)$ is obtained by forming an unbiased estimate of $\operatorname{cov}(x_1, x_2)$ directly from end-to-end measurements of source-to-leaf delays.

If the network is approximately stationary over the measurement period, a natural approach is to use the maximum likelihood estimator— the resulting maximization of the joint likelihood function requires numerical optimization and may be successfully solved using the expectation maximization (EM) algorithm [10]. More generally, however, the dynamics of the network could be changing over time and it is suggested in [8] that a sequential Monte–Carlo (SMC) procedure may be employed for tracking the time–varying delay distributions.

2.3 Network Topology Inference

In the previous methodologies, it was assumed that the logical multicast tree was known in advance. However, the underlying multicast topology itself may be inferred using the loss measurements by the following three techniques [11]:

1. A grouping estimator that exploits the monotonicity of loss rates with increasing path length,
2. A maximum–likelihood estimator (MLE), and
3. A Bayesian estimator.

Details of topology inference from end-to-end measurements is well beyond the scope of this paper.

2.4 Solution Space of the Inference

Tomography techniques aim at finding certain characteristics of each link in a network using statistical inference methods like maximum likelihood estimation. However, unless the solution space for the estimate is studied, an iterative algorithm may get trapped in a local maximum. It is shown in [16] that if the loss on each link is

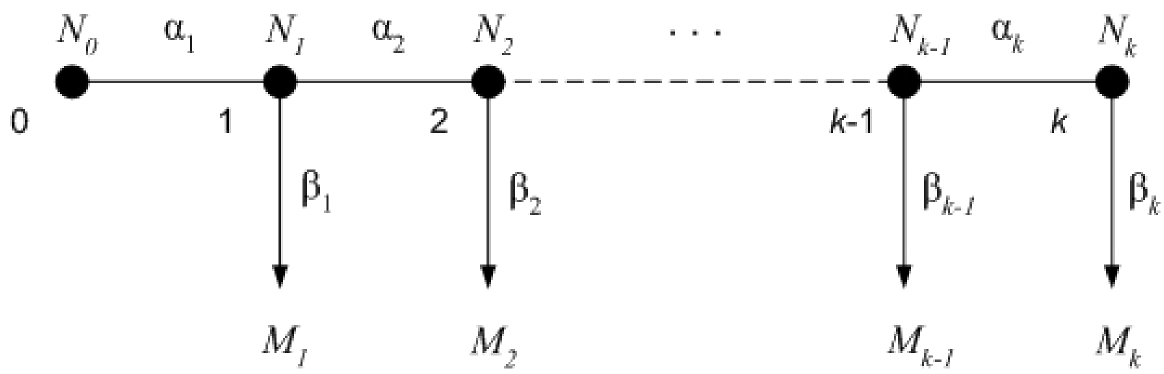

Figure 2: Model of simple network comprising a source and k nodes.

modeled as an independent Bernoulli process, the solution space is concave, ensuring that the global maximum would be correctly identified.

The proof hinges on the assumption that maximizing the joint probability of network with an independent Bernoulli loss on each link is equivalent to maximizing a log-likelihood function, as given in [3]. Again, since the logarithmic function is concave, and any linear combination of continuous concave functions is continuous and concave, the solution space of the objective function is concave and therefore the maximum likelihood estimator provides the global maximum solution.

3 Problem Definition

The general network model, along with the labeling notation being used, is introduced in this section.

Consider a network comprising a source and k nodes, as illustrated in Fig. 2. In terms of graph theory [2], this is equivalent to a $(k+1)$-vertex simple linear graph. Data originates at the source node 0 and is multicasted through to the other end terminated by node k. Packets are lost in each of the links at an unknown rate, and the packets reaching each of the intermediate nodes are sampled with a certain known probability.

The various notations used in the analysis are as follows:

N_i : number of packets at node i
α_i : probability of successful transmission on link $(i-1, i)$
β_i : sampling rate at node i
M_i : number of packets sampled at node i

L_i : list of nodes sampled at node i, its cardinality being M_i

It may be pointed out that a *one-to-many* multicast tree is assumed, and therefore each link may be uniquely identified by the index of its destination node.

The interrelationships between the various parameters introduced above are as follows:

$$\alpha_i = \frac{N_i}{N_{i-1}} \tag{4}$$

$$\beta_i = \frac{M_i}{N_i} \tag{5}$$

The problem aims at estimating the link success probabilities α_i as well as the number of packets at each of the intermediate nodes $\{i \neq 0 \neq k\}$. A multiple source network will be introduced in Section 4.2.

4 Analysis

4.1 Single Source Network

4.1.1 General k–node Network

Equation (4) gives the link success probabilities in terms of the number of packets at each node, N_i, which are unknown except for $i = 0$. This leads to a formulation for the number of packets in the ith node, which again is unknown:

$$N_i = \alpha_i \alpha_{i-1} \dots \alpha_1 N_0 \tag{6}$$

Equation (6) may be applied to equation (5) to obtain the general sampling probability.

$$\begin{aligned} \beta_i &= \frac{M_i}{N_i} \\ &= \frac{M_i}{\alpha_i \alpha_{i-1} \dots \alpha_1 N_0} \end{aligned} \tag{7}$$

Dividing β_i by β_{i-1}, the link success probabilities may be obtained to the first degree of approximation.

$$\begin{aligned} \alpha_i &= \frac{M_i \beta_{i-1}}{M_{i-1} \beta_i}, \quad i \in [1, k] \\ \text{with} \quad \beta_0 &= 1 \\ \text{and} \quad M_0 &= N_0 \end{aligned} \tag{8}$$

As has already been pointed out, the intuition behind packet loss inference is that the event that a packet has reached a given internal node in the tree can be inferred from the packet's arrival at one or more receivers descended from that node [6]. Thus, advantage is taken of this "information spill" to obtain better estimates for the link probabilities. For instance, L_i, consisting of the packets sampled by the ith node, can be enhanced by appending all packets sampled by nodes $i+1, i+2, \ldots, k$ since they must have been received by node i as well.

Combining these ideas with the standard notation of set theory, the enhanced set M'_{k-1} for node $k-1$ may be obtained as follows:

$$\begin{aligned} M'_{k-1} &= |L_{k-1} \cup L_k| \\ &= |L_{k-1} + L_k - L_{k-1} \cap L_k| \\ &= M_{k-1} + M_k - M_{k-1}\alpha_k\beta_k \end{aligned} \tag{9}$$

where $|L_{k-1} \cap L_k|$ is the set of common packets in lists L_{k-1} and L_k and is approximated by the number of packets from node $k-1$ that reach node k (with probability α_k) and get recorded (with probability β_k).

Moving on to the next node on the left, $k-2$, the analysis is repeated with the modified value of M_{k-1}:

$$\begin{aligned} M'_{k-2} &= M_{k-2} + M'_{k-1} - \left|L_{k-2} \cap L'_{k-1}\right| \\ &= M_{k-2} + M_{k-1} + M_k - M_{k-1}\alpha_k\beta_k - |L_{k-2} \cap L_{k-1}| - |L_{k-2} \cap L_k| + |L_{k-2} \cap L_{k-1} \cap L_k| \\ &= M_{k-2} + M_{k-1} + M_k - M_{k-1}\alpha_k\beta_k - M_{k-2}\alpha_{k-1}\beta_{k-1} - M_{k-2}\alpha_{k-1}\alpha_k\beta_k + M_{k-2}\alpha_{k-1}\beta_{k-1}\alpha_k\beta_k \end{aligned} \tag{10}$$

In this manner, the revised sample list may be obtained by moving from the right to the left (Fig. 2), and using the previously obtained set at each step. This formulation is now generalized for the ith node using the Inclusion–Exclusion Principle of Set Theory [9]. Finally, the number of packets at each node, and hence the link success probabilities, can be approximated from these quantities.

$$
\begin{aligned}
M_i' &= \left|L_i \cup L_{i+1}'\right| \\
&= \left|L_i \cup L_{i+1} \cup L_{i+2}'\right| \\
&= \left|L_i \cup L_{i+1} \cup L_{i+2} \cup \ldots \cup L_k\right| \\
&= \sum_{i \le j \le k} |L_j| - \sum_{i \le j1 < j2 \le k} |L_{j1} \cap L_{j2}| + \sum_{i \le j1 < j2 < j3 \le k} |L_{j1} \cap L_{j2} \cap L_{j3}| - \ldots + (-1)^{k-1} |L_i \cap L_{i+1} \cap \ldots \cap L_k| \\
&= \sum_{i \le j \le k} M_j - \sum_{i \le j1 < j2 \le k} M_{j1} \alpha_{j1+1} \alpha_{j1+2} \cdots \alpha_{j2} \beta_{j2} \\
&\quad + \sum_{i \le j_1 < j_2 < j_3 \le k} M_{j1} \alpha_{j1+1} \alpha_{j1+2} \cdots \alpha_{j2} \beta_{j2} \alpha_{j2+1} \alpha_{j2+2} \cdots \alpha_{j3} \beta_{j3} \\
&\quad - \ldots + (-1)^{k-i} M_i \alpha_{i+1} \beta_{i+1} \alpha_{i+2} \beta_{i+2} \cdots \alpha_k \beta_k
\end{aligned}
\tag{11}
$$

Thus,

$$N_i' = \frac{M_i'}{\beta_i} \quad \text{and} \tag{12}$$

$$\alpha_i' = \frac{N_i'}{N_{i-1}'} = \frac{M_i' \beta_{i-1}}{M_{i-1}' \beta_i} \tag{13}$$

However, β_i is defined for $i \in [1, k]$ and M_i' is defined for $i \in [1, k-1]$. It follows, therefore, that α' in (13) will be valid for $i \in [2, k-1]$.

Thus, the complete solution for α_i' is given by:

$$\alpha_i' = \begin{cases} \frac{M_1'}{\beta_1 N_0}, & i = 1 \\ \frac{M_i' \beta_{i-1}}{M_{i-1}' \beta_i}, & i \in [2, k-1] \\ \frac{M_k \beta_{k-1}}{M_{k-1}' \beta_k}, & i = k \end{cases} \tag{14}$$

A simple three-node network (Fig. 3) is used to illustrate the use of (14) to obtain improved estimates for the link success probabilities. The result is then compared with the estimation of the same link success probabilities from first principles.

Example

For the three–node network illustrated in Fig. 3, the link probability estimates are obtained by applying (14):

$$\alpha_1' = \frac{M_1'}{\beta_1 N_0} = \frac{M_1 + M_2 - M_1 \alpha_2 \beta_2}{\beta_1 N_0} \tag{15}$$

$$\alpha_2' = \frac{M_2 \beta_1}{M_1' \beta_2} = \frac{M_2 \beta_1}{(M_1 + M_2 - M_1 \alpha_2 \beta_2)\, \beta_2} \tag{16}$$

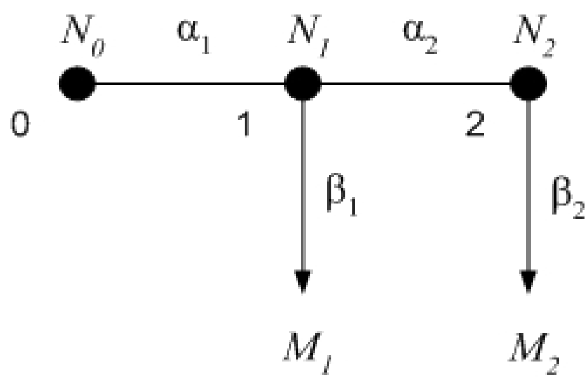

Figure 3: Model of simple network comprising 3 nodes to illustrate the analysis.

where α_2 is defined by (8).

The problem is now repeated by obtaining the original and revised estimates for the link gain probabilities by taking advantage of the correlation of the data recorded in lists L_1 and L_2.

$$\alpha_1 = \frac{M_1}{\beta_1 N_0} \tag{17}$$

$$\alpha_2 = \frac{M_2\beta_1}{M_1\beta_2} \tag{18}$$

$$M_1' = M_1 + M_2 - M_1\alpha_2\beta_2 \tag{19}$$

$$N_1' = \frac{M_1'}{\beta_1} = \frac{M_1 + M_2 - M_1\alpha_2\beta_2}{\beta_1} \tag{20}$$

$$\alpha_2' = \frac{N_2}{N_1'} = \frac{N_2\beta_1}{M_1 + M_2 - M_1\alpha_2\beta_2} \tag{21}$$

$$\alpha_1' = \frac{N_1'}{N_0} = \frac{M_1 + M_2 - M_1\alpha_2\beta_2}{N_0\beta_1} \tag{22}$$

Since $\frac{M_2}{\beta_2} = N_2$, the results obtained by the two methods are the same.

Thus, substituting the expression for α_2 in (15) and (16), the estimates are obtained as:

$$\alpha_1' = \frac{M_1 + M_2 - M_2\beta_1}{\beta_1 N_0} \tag{23}$$

$$\alpha_2' = \frac{M_2\beta_1}{(M_1 + M_2 - M_2\beta_1)\,\beta_2} \tag{24}$$

Simple MATLAB simulations were performed based on equations (11) and (14), and their results were compared. It was seen that the estimation improved with increasing number of "recording" nodes in the network. It is also intuitive that a higher recording probability or sampling rate is required nearer the destination since not much information is available from succeeding nodes.

4.1.2 Modified k–node Network

If it is assumed that there is a single data source in a network located at the root of the tree, then the simple three–node network of Fig. 3 may be converted into a symmetric two–leaf tree using the steps outlined on Fig. 4. Standard network tomography techniques may now be applied to estimate the link characteristics α_1, β_1 and $\alpha_2\beta_2$. This method is perfectly general, however, and can also be applied to the k–node network of Fig. 2 to obtain a logical tree with link characteristics $\beta_1, \beta_2, \ldots, \beta_{k-1}; \alpha_1, \alpha_2, \ldots, \alpha_{k-1}$ and the product $\alpha_k\beta_k$. Maximum likelihood estimation (MLE) techniques may be applied to the resulting tree to obtain the various link parameters (assuming that the sampling probabilities, β_is, are known).

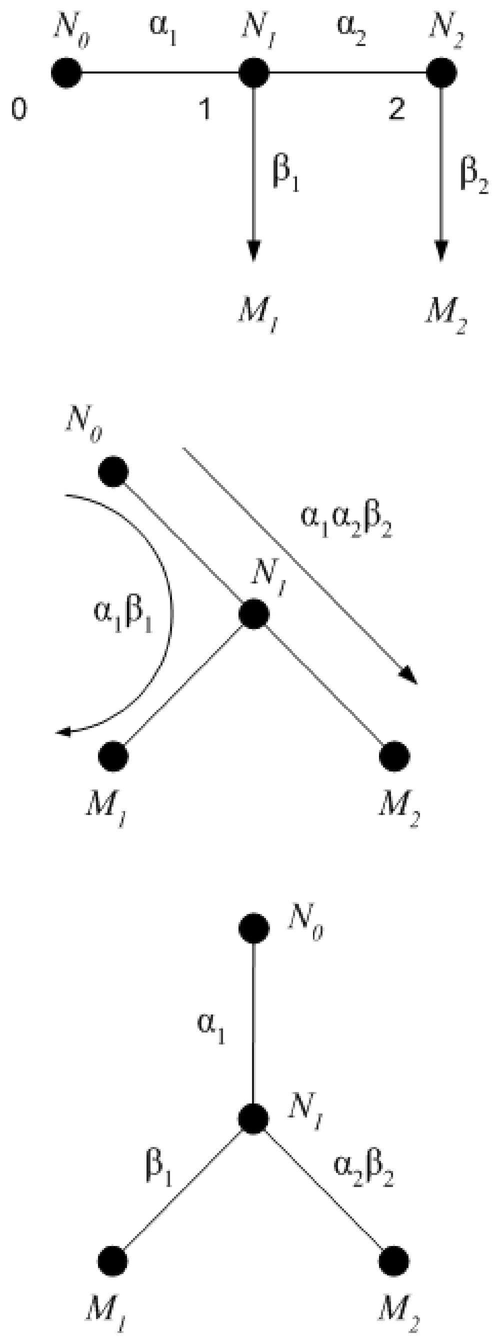

Figure 4: Modifying a simple three–node network to a standard network tomography problem.

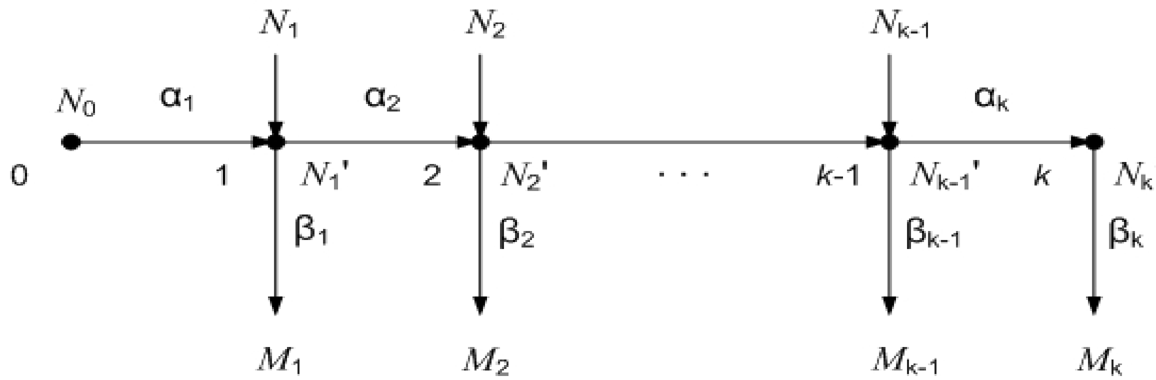

Figure 5: Model of a multiple source network comprising $k-1$ sources and $k+1$ nodes.

4.2 Multiple Source Network

As can be seen from Figure 5, a multiple source network is treated differently since the packets sampled at each internal node include packets from the original source 0, as well as new packets from intermediate sources.

As a result, the link success probabilities now change to:

$$\begin{aligned}
\alpha_1 &= \frac{N_1' - N_1}{N_0} = \frac{\frac{M_1}{\beta_1} - N_1}{N_0} = \frac{M_1 - \beta_1 N_1}{\beta_1 N_0} \\
\alpha_2 &= \frac{N_2' - N_2}{N_1'} = \frac{\frac{M_2}{\beta_2} - N_2}{\frac{M_1}{\beta_1}} = \frac{\beta_1 M_2 - \beta_1 \beta_2 N_2}{\beta_2 M_1} \\
&\vdots \\
\alpha_{k-1} &= \frac{N_{k-1}' - N_{k-1}}{N_{k-2}'} = \frac{\frac{M_{k-1}}{\beta_{k-1}} - N_{k-1}}{\frac{M_{k-2}}{\beta_{k-2}}} = \frac{\beta_{k-2} M_{k-1} - \beta_{k-2}\beta_{k-1} N_{k-1}}{\beta_{k-1} M_{k-2}} \\
\alpha_k &= \frac{N_k'}{N_{k-1}'} = \frac{\frac{M_k}{\beta_k}}{\frac{M_{k-1}}{\beta_{k-1}}} = \frac{\beta_{k-1} M_k}{\beta_k M_{k-1}}
\end{aligned} \tag{25}$$

Thus, the complete solution for α_i may be summarized as:

$$\alpha_i = \begin{cases} \frac{M_1 - \beta_1 N_1}{\beta_1 N_0}, & i = 1 \\ \frac{\beta_{i-1} M_i - \beta_{i-1}\beta_i N_i}{\beta_i M_{i-1}}, & i \in [2, k-1] \\ \frac{\beta_{k-1} M_k}{\beta_k M_{k-1}}, & i = k \end{cases} \tag{26}$$

Proceeding in the same way as Section 4.1, and denoting the list maintained at sampler i by L'_i, the enhanced sampling list may now be created by appending all packets sampled by nodes $i+1$, $i+2$, ..., k since they must have been received by node i as well.

Consequently,

$$\begin{aligned} M'_{k-1} &= \left|L'_{k-1} \cup L'_k\right| \\ &= \left|L'_{k-1} + L'_k - L'_{k-1} \cap L'_k\right| \\ &= M_{k-1} + M_k - M_{k-1}\alpha_k\beta_k \end{aligned} \tag{27}$$

where $\left|L'_{k-1} \cap L'_k\right|$ is the set of common packets in lists L'_{k-1} and L'_k and is approximated by the number of packets from node $k-1$ that reach node k (with probability α_k) and get recorded (with probability β_k).

Moving on to the next node on the left, $k-2$, the analysis is repeated with the modified value of M_{k-1}:

$$\begin{aligned} M'_{k-2} &= \left|L'_{k-2} \cup L'_{k-1} \cup L'_k\right| \\ &= M_{k-2} + M_{k-1} + M_k - \left|L'_{k-2} \cap L'_{k-1}\right| - \left|L'_{k-1} \cap L'_k\right| - \left|L'_{k-2} \cap L'_k\right| + \left|L'_{k-2} \cap L'_{k-1} \cap L'_k\right| \\ &= M_{k-2} + M_{k-1} + M_k - M_{k-1}\alpha_k\beta_k - M_{k-2}\alpha_{k-1}\beta_{k-1} - M_{k-2}\alpha_{k-1}\alpha_k\beta_k + M_{k-2}\alpha_{k-1}\beta_{k-1}\alpha_k\beta_k \end{aligned} \tag{28}$$

In this manner, the revised sample list may be obtained by moving from the right to the left (Fig. 5), and using the previously obtained set at each step. This formulation is now generalized for the ith node using the Inclusion–Exclusion Principle of Set Theory [9]. Finally, the number of packets at each node, and hence the link success probabilities, can be approximated from these quantities.

$$\begin{aligned} M'_i &= \left|L'_i \cup L'_{i+1}\right| \\ &= \left|L'_i \cup L'_{i+1} \cup L'_{i+2}\right| \\ &= \left|L'_i \cup L'_{i+1} \cup L'_{i+2} \cup \ldots \cup L'_k\right| \\ &= \sum_{i \le j \le k} \left|L'_j\right| - \sum_{i \le j1 < j2 \le k} \left|L'_{j1} \cap L'_{j2}\right| + \sum_{i \le j1 < j2 < j3 \le k} \left|L'_{j1} \cap L'_{j2} \cap L'_{j3}\right| - \ldots + (-1)^{k-1} \left|L'_i \cap L'_{i+1} \cap \ldots \cap L'_k\right| \\ &= \sum_{i \le j \le k} M_j - \sum_{i \le j1 < j2 \le k} M_{j1}\alpha_{j1+1}\alpha_{j1+2}\cdots\alpha_{j2}\beta_{j2} \\ &\quad + \sum_{i \le j_1 < j_2 < j_3 \le k} M_{j1}\alpha_{j1+1}\alpha_{j1+2}\cdots\alpha_{j2}\beta_{j2}\alpha_{j2+1}\alpha_{j2+2}\cdots\alpha_{j3}\beta_{j3} \\ &\quad - \ldots + (-1)^{k-i} M_i\alpha_{i+1}\beta_{i+1}\alpha_{i+2}\beta_{i+2}\cdots\alpha_k\beta_k \end{aligned} \tag{29}$$

Substituting this in (26), the improved link success probabilities are obtained.

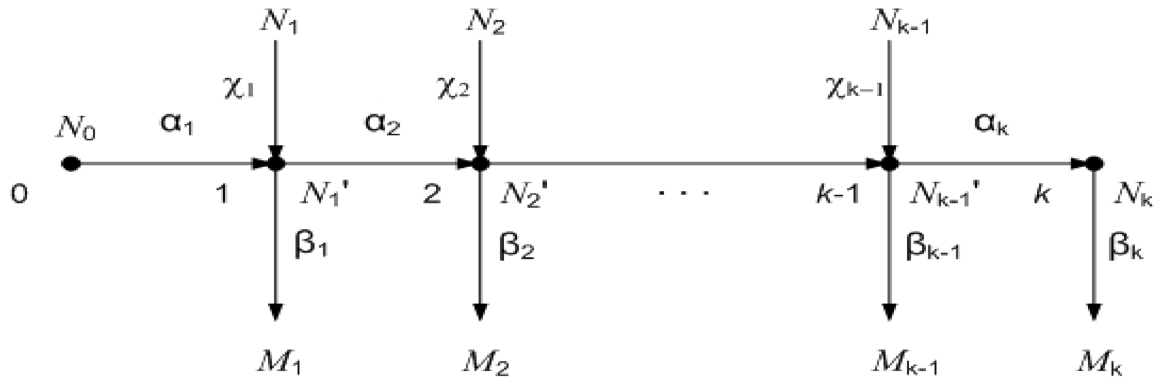

Figure 6: Model of a multiple source network comprising $k-1$ sources and $k+1$ nodes and probabilities γ associated with each source.

Now, if the sources also have a success probability, γ, associated with them (as shown in Fig. 6), the probabilities α and γ are obtained as follows:

$$\begin{aligned}
\alpha_1 &= \frac{N_1' - \gamma_1 N_1}{N_0} = \frac{\frac{M_1}{\beta_1} - \gamma_1 N_1}{N_0} = \frac{M_1 - \beta_1\gamma_1 N_1}{\beta_1 N_0} \\
\alpha_2 &= \frac{N_2' - \gamma_2 N_2}{N_1'} = \frac{\frac{M_2}{\beta_2} - \gamma_2 N_2}{\frac{M_1}{\beta_1}} = \frac{\beta_1 M_2 - \beta_1\beta_2\gamma_2 N_2}{\beta_2 M_1} \\
&\vdots \\
\alpha_{k-1} &= \frac{N_{k-1}' - \gamma_{k-1} N_{k-1}}{N_{k-2}'} = \frac{\frac{M_{k-1}}{\beta_{k-1}} - \gamma_{k-1} N_{k-1}}{\frac{M_{k-2}}{\beta_{k-2}}} = \frac{\beta_{k-2} M_{k-1} - \beta_{k-2}\beta_{k-1}\gamma_{k-1} N_{k-1}}{\beta_{k-1} M_{k-2}} \\
\alpha_k &= \frac{N_k'}{N_{k-1}'} = \frac{\frac{M_k}{\beta_k}}{\frac{M_{k-1}}{\beta_{k-1}}} = \frac{\beta_{k-1} M_k}{\beta_k M_{k-1}}
\end{aligned} \tag{30}$$

The values of γ are similarly obtained:

$$\begin{aligned}
\gamma_1 &= \frac{N_1' - N_0}{N_1} = \frac{\frac{M_1}{\beta_1} - N_0}{N_1} = \frac{M_1 - \beta_1 N_0}{\beta_1 N_1} \\
\gamma_2 &= \frac{N_2' - N_1'}{N_2} = \frac{\frac{M_2}{\beta_2} - \frac{M_1}{\beta_1}}{N_2} = \frac{\beta_1 M_2 - \beta_2 M_1}{\beta_1 \beta_2 N_2} \\
&\vdots \\
\gamma_{k-1} &= \frac{N_{k-1}' - N_{k-2}'}{N_{k-1}} = \frac{\frac{M_{k-1}}{\beta_{k-1}} - \frac{M_{k-2}}{\beta_{k-2}}}{N_{k-1}} = \frac{\beta_{k-2} M_{k-1} - \beta_{k-1} M_{k-2}}{\beta_{k-2}\beta_{k-1} N_{k-1}}
\end{aligned} \tag{31}$$

Thus, the link success probabilities may be summarized as follows:

$$\alpha_i = \begin{cases} \frac{M_1 - \beta_1 \gamma_1 N_1}{\beta_1 N_0}, & i = 1 \\ \frac{\beta_{i-1} M_i - \beta_{i-1} \beta_i \gamma_i N_i}{\beta_i M_{i-1}}, & i \in [2, k-1] \\ \frac{\beta_{k-1} M_k}{\beta_k M_{k-1}}, & i = k \end{cases} \tag{32}$$

and,

$$\gamma_i = \begin{cases} \frac{M_1 - \beta_1 N_0}{\beta_1 N_1}, & i = 1 \\ \frac{\beta_{i-1} M_i - \beta_i M_{i-1}}{\beta_{i-1} \beta_i N_i}, & i \in [2, k-1] \\ 0, & i = k \end{cases} \tag{33}$$

Once again, the values of γ and α can be improved by using the enhanced sample number M' from (29) in the calculations.

5 Results

The link success probabilities of a three–node network (Fig. 4) have been obtained in [5], and derived in the Appendix, as follows:

$$\begin{aligned}
\hat{\alpha}_1 &= \frac{(\hat{p}(01) + \hat{p}(11))\,(\hat{p}(10) + \hat{p}(11))}{\hat{p}(11)} \\
\hat{\alpha}_2 &= \frac{\hat{p}(11)}{\hat{p}(01) + \hat{p}(11)} \\
\hat{\alpha}_3 &= \frac{\hat{p}(11)}{\hat{p}(10) + \hat{p}(11)}
\end{aligned} \tag{34}$$

The probabilities $\{p_{01}, p_{10}, p_{11}\}$ are now expressed in terms of the number of observations of the three combinations at the leaf nodes:

$$\begin{aligned}
\hat{\alpha}_1 &= \frac{(n(01)+n(11))\,(n(10)+n(11))}{n(11)} \\
\hat{\alpha}_2 &= \frac{n(11)}{n(01)+n(11)} \\
\hat{\alpha}_3 &= \frac{n(11)}{n(10)+n(11)}
\end{aligned} \tag{35}$$

Here n denotes the total number of probes sent out, so that $n(00)+n(01)+n(10)+n(11)=n$.

These link success probabilities are now compared with the results obtained in (15) and (16), which are reproduced below.

$$\begin{aligned}
\alpha_1' &= \frac{M_1+M_2-M_1\alpha_2\beta_2}{\beta_1 N_0} \\
\alpha_2' &= \frac{M_2\beta_1}{(M_1+M_2-M_1\alpha_2\beta_2)\,\beta_2} \\
\alpha_2 &= \frac{M_2\beta_1}{M_1\beta_2}
\end{aligned} \tag{36}$$

Comparing with the notation of [5], it follows from Fig. 4 that

$$\begin{aligned}
N_0 &= n \\
N_1 &= n(11)+n(10)+n(01)+n(11) \\
M_1 &= n(11)+n(10) \\
M_2 &= n(11)+n(01)
\end{aligned} \tag{37}$$

and, (38)

$$\begin{aligned}
\hat{\alpha}_1 &= \alpha_1 \\
\hat{\alpha}_2 &= \beta_1 \\
\hat{\alpha_3} &= \alpha_2\beta_2
\end{aligned} \tag{39}$$

Then,

$$\begin{aligned}\hat{\alpha}_1 = \alpha_1 &= \frac{M_1 + M_2 - M_2\beta_1}{\beta_1 N_0} \\ &= \frac{n(11) + n(10) + n(11) + n(01) - \beta_1 \left[n(11) + n(01)\right]}{n\beta_1}\end{aligned} \tag{40}$$

$$\hat{\alpha}_2 = \beta_1 = \text{known} \tag{41}$$

$$\begin{aligned}\hat{\alpha}_3 = \alpha_2\beta_2 &= \frac{M_2\beta_1}{M_1 + M_2 - M_2\beta_1} \\ &= \frac{\beta_1 \left(n(11) + n(01)\right)}{n(01) + n(10) + 2n(11) - \beta_1 \left[n(11) + n(01)\right]}\end{aligned} \tag{42}$$

It may be pointed out, however, that an important point of difference between the two approaches is that [5] uses n active probes to get an estimate of the link losses, whereas the method outlined in this paper relies on passive network tomography with measurements being made in the course of the transportation of data packets. This has the additional advantage of not contributing to network traffic, since a large multicast network may easily get overburdened with active probes. The use of the expectation-maximization (EM) algorithm in passive network tomography is outlined in [14].

6 Conclusion

This paper introduces the subject of network tomography through a variety of existing papers and applications. These ideas are then correlated to the problem at hand, namely, attempting to estimate the packet transfer probability of each link of a linear graph. The technique exploits the fact that packets that pass through a node must also pass through its parent node, and derives estimates for the individual links of a simple three–node network. The problem is then recast as a typical network tomography problem and methods like the MLE are suggested to obtain its solution.

This paper considers only one–to–two tree topologies; the analysis can be extended to two–to–one trees so that a general multiple source network may be analyzed. The use of the EM algorithm for passive network tomography is also a prospective area of research.

Appendix

This section derives the Maximum Likelihood (ML) estimate for link success probabilities using the approach outlined by Caceres *et al.* in [5],[4]. Before proceeding with the analysis, the following notation is defined:

$\tau = (V, L)$	: logical multicast tree consisting of set of nodes V and set of links L
$d(j)$	: set of children of node j, *i.e.*, $d(j) = \{k \in V : (j,k) \in L\}$
$f(j)$	: parents of node j, *i.e.*, $d(j) = \{k \in V : (k,j) \in L\}$
$f^n(j) = f\left(f^{n-1}(k)\right)$	: j is a descendant of k if $k = f^n(j)$ for some integer $n > 0$
$R \subset V$	: set of leaf nodes representing the children
$U = V \backslash \{0\}$	: vertex set V excluding node 0 (root)
$0 \in V$	: root node, *i.e.*, source of the probes
$\alpha_k \in [0,1]$	: probability that a given probe is not lost on the link terminating at $k \in V$
$X = (X_k)_{k \in V}$	: passage of probes down the tree, where each X_k takes a value in $\{0,1\}$
P_α	: distribution of outcomes $(X_k)_{k\in V}$ for a given set of link probabilities $\alpha = (\alpha_k)_{k \in V}$
$\Omega = \{0,1\}^R$	: space of all outcomes
n	: number of probes sent out, each with outcome $x \in \Omega$
$n(x)$	: number of probes for which outcome x is obtained
$p(x;\alpha) = \mathrm{P}_\alpha\left(X_{(R)} = x\right)$	: probability mass function for a single outcome $x \in \Omega$
$k \prec k'$	: k is descended from k', but $k \neq k'$ for $k, k' \in V$
$l = l(k)$	: link k is at level $l(k)$ if there is a chain of l ancestors leading to the root of τ, *i.e.*, $k = f^0(k) \prec f^1(k) \prec \cdots \prec f^l(k)$
$\tau(k) = (V(k), L(k))$	: sub-tree within τ rooted at node k
$R(k) = R \cap V(k)$	: set of receivers descended from k
$\Omega(k)$	: set of outcomes x in which at least one receiver in $R(k)$ receives a packet
$\gamma_k = \mathrm{P}_\alpha\left[\Omega(k)\right]$	: probability that probe reaches at least one receiver
$\hat{p}(x) = \frac{n(x)}{n}$	: observed proportion of trials with outcome x
$\gamma'(k) = \sum_{x \in \gamma(k)} \hat{p}(x)$	: estimate of γ_k
$\beta_k = \mathrm{P}[\gamma(k) \mid X_{f(k)} = 1]$	: probability that probe reaches at least one receiver, given that it reaches the parent $f(k)$

$$= \begin{cases} 1 - \left(\bar{\alpha}_k + \alpha_k \prod_{j \in d(k)} \bar{\beta}_j\right), & k \in V \backslash R \\ \alpha_k, & k \in R \end{cases}$$

Using this notation, the probability of n independent observations $x^1, \ldots, x^n$, given the link success probability α is:

$$\begin{aligned} p\left(x^1, \ldots, x^n; \alpha\right) &= \prod_{m=1}^{n} p\left(x^m; \alpha\right) \\ &= \prod_{x \in \Omega} p(x; \alpha)^{n(x)} \end{aligned}$$

The aim is to estimate the value of α from a set of experimental data $(n(x))_{x \in \Omega}$ so that the maximum likelihood estimator MLE, α', will maximize $p\left(x^1, \ldots, x^n; \alpha\right)$ for the data $x^1, \ldots, x^n$.

Since the logarithmic function is monotonically increasing, maximizing $p\left(x^1, \ldots, x^n; \alpha\right)$ is equivalent to maximizing the log-likelihood function below:

$$\begin{aligned} \mathcal{L}(\alpha) &= \log p\left(x^1, \ldots, x^n; \alpha\right) \\ &= \sum_{x \in \Omega} n(x) \log p(x; \alpha) \end{aligned}$$

Furthermore, a map $\Gamma : \alpha \mapsto \gamma$ will be defined in Theorem 1, and the strategy would be to obtain the MLE

$$\alpha' = \arg \max_{\alpha \in [0,1]^R} \mathcal{L}(\alpha)$$

from the estimates γ' instead by using the inverse mapping Γ^{-1}.

Using the notation defined, the relationship between α and γ is as follows:

$$\begin{aligned} \gamma_k &= \beta_k \prod_{i=1}^{l(k)} \alpha_{f^i(k)} \\ &= \alpha_k \prod_{i=1}^{l(k)} \alpha_{f^i(k)}, \quad k \in R \\ &= \prod_{i=0}^{l(k)} \alpha_{f^i(k)} \\ &\equiv A_k \end{aligned}$$

Before proceeding any further, the following Lemma is necessary.

Lemma 1 *Let C be the set of $c = (c_i)_{i=1,2,\ldots,i_{max}}$ with $c_i \in (0,1)$ and $\sum_i c_i > 1$. Then the equation*

$$(1 - x) = \prod_i (1 - c_i x)$$

has a unique solution $x(c) \in (0,1)$.

Proof

Let

$$
\begin{aligned}
h_1(x) &= (1-x) \\
h_2(x,c) = h_2(x) &= \prod_i (1-c_i x) \\
\text{and} \quad q_i &= \frac{c_i}{1-c_i x} \\
\text{Then, } \log h_2(x) &= \sum_i \log\left(1-c_i(x)\right) \\
\Rightarrow h_2'(x) &= -h_2(x) \sum_i q_i(x) \\
\Rightarrow h_2''(x) &= h_2(x) \left[\left(\sum_i q_i \right)^2 - \sum_i q_i^2 \right] \\
&> 0
\end{aligned}
$$

Thus, $h(x) = h_1(x) - h_2(x)$ is strictly concave and continuous on [0,1], and therefore there is exactly one solution to $h(x) = 0$ for $x \in (0,1)$.

Theorem 1 *Let*

$$
\begin{aligned}
\mathcal{A} &= \left\{ (\alpha_k)_{k \in U} : \alpha_k > 0 \right\} \\
\mathcal{A}^{(1)} &= \left\{ (\alpha_k)_{k \in U} : 1 \geq \alpha_k > 0 \right\} \\
\mathcal{G} &= \left\{ (\gamma_k)_{k \in U} : \gamma_k > 0 \forall k, \gamma_k < \sum_{j \in d(k)} \gamma_j \forall k \in U \backslash R \right\}
\end{aligned}
$$

The map $\alpha \mapsto \gamma$ defined on $\mathcal{A}^{(1)}$ extends to a bijection Γ from a subset of $\mathcal{A}$ onto $\mathcal{G}$.

Proof

It has already been demonstrated that $\gamma_k = A_k, \quad k \in R$. Now define the following:

$$
\begin{aligned}
H_k(A_k, \gamma) &= \left(1 - \frac{\gamma_k}{A_k}\right) - \prod_{j \in d(k)} \left(1 - \frac{\gamma_j}{A_k}\right) \\
&= 0, \quad k \in U \backslash R
\end{aligned}
$$

This is of the same form as Lemma 1, with $x = \frac{\gamma_k}{A_k}$ and $c_j = \frac{\gamma_j}{\gamma_k}$, and therefore, for each $\gamma \in \mathcal{G}$, there is a unique $A_k > \gamma_k$ that solves $H_k(A_k, \gamma) = 0$. α_k is uniquely recovered from A_k by taking the appropriate quotient

$$\alpha_k = \frac{A_k}{A_{f(k)}}, \quad k \in U$$

and setting $A_0 = \alpha_0 = 1$. This construction specifies the inverse map Γ^{-1}.

For the equivalent three-node network shown in Fig. 4 (bottom), the outcomes corresponding to $\Omega = \{0,1\}^2$ can be represented by $\{00, 01, 10, 11\}$. Then,

$$\gamma_1' = \hat{p}(11) + \hat{p}(10) + \hat{p}(01)$$
$$\gamma_2' = \hat{p}(11) + \hat{p}(10)$$
$$\gamma_3' = \hat{p}(11) + \hat{p}(01)$$

Using the result of Lemma 1, the following estimates are obtained:

$$\hat{\alpha}_1 = \frac{\gamma_2'\gamma_3'}{\gamma_2 + \gamma_3 - \gamma_1} = \frac{(\hat{p}(01) + \hat{p}(11))\,(\hat{p}(10) + \hat{p}(11))}{\hat{p}(11)}$$
$$\hat{\alpha}_2 = \frac{\gamma_2' + \gamma_3' - \gamma_1}{\gamma_3'} = \frac{\hat{p}(11)}{\hat{p}(01) + \hat{p}(11)}$$
$$\hat{\alpha}_3 = \frac{\gamma_2' + \gamma_3' - \gamma_1}{\gamma_2'} = \frac{\hat{p}(11)}{\hat{p}(10) + \hat{p}(11)}$$

It can be shown that the estimates $\hat{\alpha}_i$ satisfy the following two properties [1]:

- *Consistency*: $\hat{\alpha}_i$ converges to the true value α_i almost surely as the number of probes, n grows to infinity.
- *Asymptotic normality*: The distribution of the quantity $\sqrt{n}\,(\hat{\alpha}_i - \alpha_i)$ converges to a normal distribution as n grows to infinity.

References

[1] Andrew Adams, Tian Bu, Timur Friedman, Joseph Horowitz, Don Towsley, Ramon Caceres, Nick Duffield, Francesco Lo Presti, Sue B. Moon, and Vern Paxson. The Use of End-to-End Multicast Measurements for Characterizing Internal Network Behavior. *IEEE Comm. Mag.*, pages 152–158, May 2000. 1, 2.1, 6

[2] J.A. Bondy and U.S.R. Murty. *Graph Theory with Applications*. North–Holland, New York, NY, 1976. 1, 3

[3] R. Caceres, N.G. Duffield, J. Horowitz, F. Lo Presti, and D. Towsley. Loss–Based Inference of Multicast Network Topology. In *Proc. IEEE Conf. Dec. and Cont.*, volume 3, pages 3065–3070, December 1999. 1, 2.4

[4] R. Caceres, N.G. Duffield, J. Horowitz, D. Towsley, and T. Bu. Multicast–Based Inference of Network–Internal Characteristics: Accuracy of Packet Loss Estimation. In *Proc. IEEE INFOCOM 1999*, pages 371–379, March 1999. 1, 2.1, 6

[5] R. Caceres, N.G. Duffield, J. Horowitz, and D.F. Towsley. Multicast–Based Inference of Network Internal Loss Characteristics. *IEEE Trans. Info. Theory*, pages 2462–2480, November 1999. 2.1, 5, 5, 5, 6

[6] Ramon Caceres, N.G. Duffield, Sue B. Moon, and Don Towsley. Inferring Link–Level Performance from End-to-End Multicast Measurements. In *Proc. Global Int. Symp.*, December 1999. 4.1.1

[7] Mark Coates and Robert Nowak. Unicast Network Tomography Using EM Algorithms. Technical Report TR-0004, Rice University, September 2000. 1

[8] Mark J. Coates and Robert D. Nowak. Network Tomography for Internal Delay Estimation. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, volume 6, pages 3409–3412, May 2001. 1, 2.2, 2.2

[9] Louis Comtet. *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. D. Reidel Publishing, Dordrecht, 1974. 4.1.1, 4.2

[10] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Roy. Stat. Soc., Series B (Meth.)*, 39(1):1–38, 1977. 2.2

[11] N.G. Duffield, Joseph Horowitz, Francesco Lo Presti, and Don Towsley. Multicast Topology Inference From Measured End-to-End Loss. *IEEE Trans. Info. Theory*, 48(1):26–45, January 2002. 1, 2.3

[12] N.G. Duffield and F. Lo Presti. Multicast Inference of Packet Delay Variance at Interior Network Links. In *Proc. IEEE INFOCOM 2000*, pages 1351–1360, March 2000. 2.2

[13] Meng-Fu Shih and Alfred Hero. Unicast Inference of Network Link Delay Distributions from Edge Measurements. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, pages 3421–3424, May 2001. 1

[14] Yolanda Tsang, Mark Coates, and Robert Nowak. Passive Network Tomography Using EM Algorithms. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, volume 2, pages 1469–1472, May 2001. 5

[15] Y. Vardi. Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data. *J. Amer. Stat. Assoc.*, 91:365–377, 1996. 1

[16] Weiping Zhu. Solution Space of Loss Tomography. In *Proc. IEEE ICC 2003*, volume 1, pages 286–290, 2003. 2.4