Optimal Network Tomography

Yolanda Tsang
Advisor: Robert Nowak
Department of Electrical and Computer Engineering
Rice University, Houston, TX 77005
E-mail: ytsang@ieee.org, nowak@engr.wisc.edu

I. Motivation

Network tomography is a powerful method for measuring and analyzing link specific characteristics, such as link loss rates and delay statistics, using only end–to–end probes as opposed to making measurements on every internal link in the network. This capability enables the diagnosis of congestion and performance degradation across heterogeneous networks. Routing algorithms, servicing strategies, security problems, and performance verification can all benefit from monitoring techniques that report such information. Most network tomography schemes considered to date explicitly or implicitly distribute the probes uniformly among the subpaths.

In this paper, we formulate network tomography as an optimal probe allocation problem, in which a fixed number of network probes are optimally distributed to minimize the squared estimation error of the tomographic process. Explicit forms for the estimation errors are derived in terms of topology, noise levels, and number and distribution of probes. This analysis reveals the dominant sources of ill-conditioning and issues scalability of network tomography. To our knowledge, this issue has not been addressed in literatures.

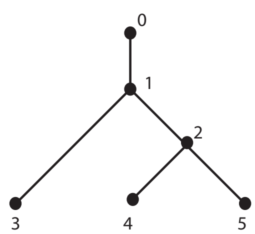

Fig. 1. An example of tree-structured topology with single sender (node 0) and 3 receivers (node 3 through 5).

II. Measurement Framework

Typical network tomography schemes involve probing several receivers from a single sender. Standard routing protocols thus lead to tree-structured topologies, with the sender at the root and the receivers at the leaves. We consider the case where probes are sent using unicast protocols (Internet traffic is unicast in nature). We make the standard assumption for performance characteristics study that the topology remains fixed and the paths are unique between the sender and each receiver. This assumption holds when the measurement period is significantly shorter that routing tables are updated. Furthermore, we assume the topology is known, e.g., using standard diagnostic tools such as `traceroute` [1]. Connections between the sender, routers, and receivers are called *links*. Figure 1 depicts an example of a logical tree-structured topology.

In network tomography studies, the goal is to infer internal link level characteristics from end-to-end measurement; i.e., we are restricted to only edge-based measurement of a tree topology. The measurements are usually the number of packets received successfully at each receiver or the packet traversal time between the sender and each receiver. This raw data must be transformed, through a matrix inversion or analogous process, to obtain link-level characteristics. This inverse problem is ill-posed as there is no unique mapping between the path-level properties and the link-level properties. In this paper, we focus on some dominant sources of ill-conditioning and issues scalability of network tomography. For detailed descriptions and tomographic techniques, we refer readers to [2]–[5].

The edge-based measurements, coupled with the tree-structured topology, leads to a linear system of measurement equations [6] (possibly after a linearizing transformation such as the logarithm) of the form

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\epsilon}$$

where $\boldsymbol{Y}$ is the raw data acquired from end-to-end probing, $\boldsymbol{A}$ is a $n \times n$ routing matrix determined by the topology, $\boldsymbol{X} = [x_1, x_2, \cdots, x_n]^T$ is a $n \times 1$ vector of link characteristics (e.g., link delay variances, link loss rates) for links $1, 2, \cdots, n$ and $\boldsymbol{\epsilon}$ is an error term representing various sources of noise and randomness in the measurement process.

We model the errors $\boldsymbol{\epsilon}$ as a zero-mean Gaussian random vector with covariance matrix Σ. Each element of $\boldsymbol{Y}$ corresponds to the congestion experienced along a certain subpath. Therefore, we will refer to Y_i, the i-th entry of $\boldsymbol{Y}$, as the data for subpath $s(i)$ (the path from node 0 to node i), $i = 1, \ldots, n$.

There are several assumptions in the framework that are worthy of discussion. The underlying measurement framework assumes that link characteristics are stationary during the measurement period. We also assume spatial and temporal independence. This implies the link characteristics on neighboring links are not correlated as a result of cross traffic and successive packets across the same link would have the same link experience. These assumptions can be relaxed, both in theory and practice [3], but such extensions are not central to the main points in this paper.

We assume that the measurements are statistically independent, and therefore Σ is diagonal with entries equal to the variance of the data $Y_1, \ldots, Y_n$, respectively. Specifically, let σ_i^2 denote the variability of a single back-to-back probe measurement associated with Y_i, for $i = 1, \ldots, n$. If k_i probe measurements are averaged to form the final statistics Y_i, then $\Sigma = \text{diag}(\sigma_1^2/k_1, \ldots, \sigma_n^2/k_n)$. The main question addressed in this paper is the following. *Given a total budget of N probes (i.e., $\sum_{i=1}^n k_i = N$), how should these probes be distributed among the different subpath measurements Y_i to minimize the overall error of the network tomography experiment?*

III. Approach

The least squares solution to the linearized system of equations $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\epsilon}$ is

$$\hat{\boldsymbol{X}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y}$$

$\hat{\boldsymbol{X}}$ is an unique minimizer of $||\boldsymbol{A}\boldsymbol{X} - \boldsymbol{Y}||_2^2$. In network tomography problem, the goal is to infer the internal link level characteristics, thus, minimizing the variance of the estimates, $\text{cov}(\hat{\boldsymbol{X}})$. The A–optimality criterion is commonly used in optimal experimental design [7]. It is based on minimizing the sum of the variances of the estimated parameters, which is the trace of the $\text{cov}(\hat{\boldsymbol{X}})$ in this case. By computing $\text{cov}(\hat{\boldsymbol{X}})$, we show that the A–optimality criterion relates the topology, noise levels and probes in the squared error computation of the network tomography problem. The corresponding optimization problem is

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \frac{d(i)\sigma_i^2}{k_i} \\ \text{subject to} & \sum_{i=1}^n k_i = N \end{array} \tag{1}$$

where N is the probe budget that we wish to optimally allocate among the n subpath measurements, $Y_i, i = 1, \cdots, n$, and $d(i)$ is the outdegree, defined to be the number of directly connected links from node i. We want to distribute the total budget N to each links depending on the noise level as well as the outdegree of the node.

We note that the optimization problem given in Equation 1 is a discrete resource allocation problem. Moreover, the objective function is a sum of separable, convex functions $f_i(k_i) = \frac{\sigma_i^2 d(i)}{k_i}$ in k_i. The objective function can be optimized using the "marginal allocation" algorithm. We independently "discovered" the algorithm and later become aware of the fact that the same algorithm has been rediscovered by others from time to time [8]. The marginal allocation algorithm is a simple and efficient greedy algorithm. At the beginning of the algorithm, we first initialize $k_i = 0, \forall i$. At each iteration, we assign one unit of the resource to the most favorable node k_i, i.e., the smallest $f_i(k_i)$ in which by increasing k_i, the objective function is minimized. We repeat the algorithm for N iterations, or until all of the resources are allocated. For any given N, the algorithm provides an optimal and feasible solution. Refer to [8] for proofs and further details.

In an ideal situation, the probe will be allocated according to the optimal solution given the outdegree $d(i)$ as well as the noise power σ_i^2. However, in most of the cases, σ_i^2 is not known *a priori*. One possibility is to apply a two-step procedure: (1) find the estimate of the noise power, $\hat{\sigma}_i^2$ using a small number of probes distributed evenly to all measurements by assuming $\sigma_i^2 =$ constant; (2) apply the marginal allocation algorithm to the objective function with $\sigma_i^2 = \hat{\sigma}_i^2$ to obtain the optimal probe allocation k^*. For heterogeneous noise power σ_i^2, the algorithm remains unchanged and the solution remains optimal and feasible for each given budget of N measurements. Due to the lack of space, the simulation results on the performance of the algorithm are omitted. Experimentations over randomly generated network demonstrate the potential gains achievable through optimized probe allocation. Online and distributed versions of the optimized probing algorithm are potentially fruitful avenues for future work.

IV. Implications

Results from optimal probe allocation problem also provide insight into the scalability of network tomography estimation. We assume that the noise power is constant, and we define $d_{max} = \max(\{d(i)\})$ and $d_{min} = \min(\{d(i)\})$ for $i = 1, 2, \cdots, n$. The minimum squared error per receiver of the network tomography estimation can be bounded as follows:

$$\frac{d_{min}}{N} r \leq \min_{k_i : \sum k_i = N} \left\{ \frac{1}{r} \sum_{i=1}^n \frac{d(i)\sigma_i^2}{k_i} \right\} \leq 4 \frac{d_{max}}{N} r$$

The number of probes required to achieve a given level of accuracy per receiver grows linearly with the number of receivers involved in the experiment. In conclusion, in this paper, we employ resource allocation optimization techniques to design optimal probing schemes, and develop algorithms for optimal probe allocation based on topological considerations and noise characteristics.

References

[1] V. Jacobson, "traceroute," 1989, ftp://ftp.ee.lbl.gov/traceroute.tar.Z.

[2] Y. Tsang, M. Coates, and R. Nowak, "Network delay tomography," *IEEE Tran. on Signal Processing*, vol. 51, no. 8, pp. 2125–2136, August 2003.

[3] M. Coates, A. Hero, R. Nowak, and B. Yu, "Internet tomography," *IEEE Signal Processing Magazine*, May 2002.

[4] Y. Tsang, M. Coates, and R. Nowak, "Passive network tomography using EM algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, May 2001.

[5] "Multicast-based inference of network-internal characteristics (MINC)," URL: http://www.research.att.com/projects/minc.

[6] Y. Vardi, "Network tomography: estimating source-destination traffic intensities from link data," *J. Amer. Stat. Assoc.*, pp. 365–377, 1996.

[7] F. Pukelsheim, *Optimal Design of Experiments*, John Wiley & Sons, Inc., 1993.

[8] T. Ibaraki and N. Katoh, *Resource Allocation Problems: Algorithmic Approaches*, The MIT Press, 1988.