Active Network Tomography and the Design of Efficient Probing Experiments

George Michailidis
Department of Statistics
The University of Michigan
Ann Arbor, MI
USA
gmichail@umich.edu

Vijayan N. Nair
Department of Statistics
The University of Michigan
Ann Arbor, MI
USA
vnn@umich.edu

Abstract

Estimation of quality of service parameters, such as packet loss rates and delay distributions, is of considerable importance to network administrators and service providers. We consider the active network tomography problem where the goal is to estimate packet loss rates and delay distributions of all internal links in a network from measurements obtained from nodes located on its periphery. This is an example of a large-scale statistical inverse problem. We provide a brief review of the literature, including unicast and multicast probing schemes and introduce a new family of such schemes called $k-$cast probing, which is shown to be flexible and computationally efficient for collecting the necessary data in global network monitoring problems. Several methods of estimation for loss rates and delay distributions will be described, including the use of EM-algorithms for the Maximum Likelihood estimators and several classes of least-squares estimates for loss rates. Some practical issues regarding the design of bicast probing experiments on different network tree topologies and results from simulation studies will also be discussed.

1 Introduction

There is an increasing demand for sophisticated tools for monitoring network utilization and performance from ISPs, due to the offer of more complex services that require quality of service (QoS) guarantees (e.g. video-conferencing, IP telephony, etc). Furthermore, such tools are also important for numerous network management tasks, such as fault and congestion detection, ensuring service level agreement compliance, and dynamic replica management of Web services, just to name a few.

Existing network monitoring tools can be divided into two categories: (i) node-oriented tools that collect packet and network flow information from various types of network devices, such as routers, switches and hosts and (ii) path-oriented tools that can obtain information about connectivity and latency in a network through active probing schemes. The main shortcomings of the first category of tools are that they require monitoring agents to be deployed at every device and that it proves fairly complicated to provide estimates for path performance measures, such as delay. The second category of tools obtains measurements of network performance measures that indirectly relate to the parameters of interest. At a subsequent stage, the collected information is appropriately processed (through the solution of a statistical inverse problem) to obtain the information of interest (Coates et al., 2002).

In the second category belong the techniques developed in a sequence of papers (see literature review section) that infer link-level loss rates and delays from end-to-end multicast or closely spaced in time unicast transmissions obtained from monitoring a network whose topology is given by a tree.

In this paper, we investigate the packet loss rate and delay problem under the novel end-to-end measurement scheme introduced in Xi et al. (2003b) and briefly assess the theoretical properties and efficiency of several estimation methods such as maximum likelihood based on the EM algorithm and a class of least squares estimates for loss rates.

2 Problem Description and Literature Review

The problem of interest at hand can be briefly described as follows: a physical network can be logically represented by a graph $G = (\mathcal{V}, \mathcal{E})$ consisting of nodes $v \in \mathcal{V}$ connected by edges/links $e \in \mathcal{E}$. We are interested in monitoring the quality of service (e.g. average number of dropped packets and/or their delay distribution) of certain links $e \in \mathcal{E}$. In active network tomography experiments this objective is achieved by sending probing packets from a source node to a set of destination (receiver) nodes and recording the outcome of interest (e.g. dropped/received). Most of the literature the portion of the network monitored by active network tomography techniques can be represented by a *tree* topology $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ with root (sender) node $\{0\} \in \mathcal{V}$. Let $\mathcal{D}(i) = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ denote the set of *direct descendants* (children) of node i. The set of nodes $\mathcal{R} \subset \mathcal{V}$ such that $\mathcal{D}(i) = \emptyset,\ i \in \mathcal{V}$ (nodes without children) represents the set of *receivers*. An example of a 3-layer symmetric binary tree is given in Figure 1.

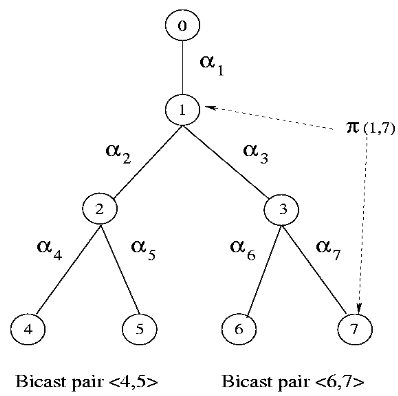

Figure 1: A 3-layer logical topology

The goal of active network tomography becomes to infer the *internal* link loss rates and delay distributions from end-to-end measurement collected from appropriately designed probing experiments. The way measurements are collected depends upon the transmission scheme used. For example under a multicast transmission, the sender sends each packet to a group of designated receivers by employing the following mechanism: at internal nodes of the logical tree where forking occurs (e.g. node 1 in Figure 2) each multicast packet is duplicated and sent along each branching path. Thus, the observed outcome for the t^{th} probe packet has the form $\{Y_i(t) = A\}_{i \in \mathcal{R}}$ with $A \in \{0, 1\}$. For the 3-layer tree depicted in Figure 2 there are 16 possible outcomes; for example the outcome (0101) indicates that the multicast packet was received only by nodes (receivers) 5 and 7, while the outcome (1111) indicates that it was received by all nodes in the receiver set. A unicast transmission scheme involves sending packets from the source to a single receiver node. However, as can be seen from Figure 1, the number of links in $\mathcal{T}$ exceeds the number of receivers and therefore using unicast end-to-end measurements would fail to identify the parameters of interest. In order to overcome this difficulty Tsang et al. (2003) proposed a modification of the unicast transmission mechanism, by sending two closely spaced in time unicast probes to neighboring receivers (e.g. receivers 4 and 5). Such a pair of packets would share a common path (e.g. 0-2) and this information can be used under certain additional assumptions to overcome the identifiability problem. The main difficulty with the multicast scheme is that the number of possible outcomes grows exponentially with the number of receivers in the tree topology, while the modified unicast scheme needs to make strong assumptions to resolve the identifiability issues. Xi et al. (2003b) proposed using a collection of *bicast* probing experiments, where multicast packets are sent to selected *pairs* of receiver nodes (e.g. pair $\langle 4, 5 \rangle$ or pair $\langle 6, 7 \rangle$ in Figure 1). Unlike unicast schemes, the bicast end-to-end scheme contains information on the bivariate loss/delay structure, which allows us to estimate the internal link losses/delays.

Caceres et al. (1999) developed an algorithm for obtaining maximum likelihood estimates of the link-loss rates under an independent model (i.e. packet losses are assumed to be independent Bernoulli random variables on the various links) using a multicast transmission scheme. In Xi et al. (2003) a regression formulation of the problem is introduced and least squares (LS), generalized LS and iterative reweighted

LS estimators obtained. The problem of network delays has been studied in a series of papers. More specifically, Lo Presti et al. (2002) introduced a heuristic algorithm that estimates the probability mass function of discretized end-to-end delay measurements for each internal link of $\mathcal{T}$ using multicast probes. Lawrence et al. (2003) and Liang et al. (2003) obtain maximum likelihood estimates under the same setting, while maximum likelihood estimates using back-to-back unicast probes are studied in Tsang et al. (2003). Shih et al. (2003) study the delay inference problem by modeling the link delays as a mixture of Gaussian distributions and a point mass at zero. Finally, the link loss and delay problems using bicast probing schemes are considered in Michailidis et al. (2003), Xi et al. (2003a) and Lawrence et al. (2003a).

3 Analysis of k-cast Probing Experiments

In this Section we provide an overview of the results we have obtained for the link loss and delay problems using k-cast and in particular bicast ($k = 2$) probing schemes.

Identifiability: In a symmetric binary tree with L layers, there are $2^{L-2}(2^{L-1} - 1)$ possible bicast pairs. It turns out that we need only a subset of these to estimate all the internal loss and delay parameters. The necessary and sufficient conditions for this minimal bicast scheme are given in Xi et al. (2003b) and Lawrence et al. (2003b). Simply stated, the minimal bicast scheme contains one pair that splits at every internal node $s \in \mathcal{V} - \{\mathcal{R} \cup \{0\}\}$. For example, for the 3-layer tree in Figure 1, the pairs $\langle 4, 5\rangle, \langle 5, 6\rangle$ and $\langle 6, 7\rangle$ constitute a minimum bicast scheme. So, the number of required pairs equals $2^{L-1} - 1$, the number of internal nodes in an $L-$layer tree: 3 in a 3-layer tree, and 7 in a 4-layer tree. Since the number of possible outcomes for a bicast pair is always four, this drastically reduces the number of possible outcomes compared to the multicast. This reduces the computational complexity greatly in the delay problem. It turns out that a slightly modified bicast scheme, with minimum + 1 pairs, has better symmetry properties and is recommended.

Inference and Related Issues: Xi et al. (2003a) develops two different EM algorithms for computing the MLEs of loss rates using bicast experiments and describes the computational complexity and convergence properties of these algorithms. Efficiency comparisons with multicast schemes using asymptotic and simulation results are also described. Design of optimal bicast experiments are also studied. For example, for a 3-layer tree with equal loss rates, the optimal allocation for a minimum + bicast scheme, is approximately 37% for pair $\langle 4, 5\rangle$ and $\langle 6, 7\rangle$ and 17% for pairs $\langle 4, 7\rangle$ and $\langle 5, 6\rangle$ each.

Fast algorithms for estimating loss rates using least-squares methods are developed in Michailidis et al. (2003). The iteratively reweighted least squares are asymptotically efficient and provide a fast way to compute the information matrix and make related inference.

Lawrence et al. (2003b) develops the EM-algorithm for computing the nonparametric MLEs of the (discrete) delay distributions and studies convergence properties as well as small and large-sample properties of the MLEs.

The performance of all the methods are also studied in a realistic environment with background traffic using the ns-simulator package.

In summary, the bicast experiments have several advantages. While there is some loss in efficiency over the multicast scheme, this is more than offset by the computational advantages. Furthermore, the bicast scheme can be designed to be adaptive so that the number of probes sent to different pairs can be adjusted in real time to efficiently detect and localize network problems.

All the issues mentioned above have been extended and studied in more complex logical topologies (Xi (2004)), such as the one depicted in Figure 2. Such topologies allow more flexibility in designing efficient probing experiments for monitoring purposes.

Acknowledgements

This is joint work with Earl Lawrence and Bowei Xi. The work of the first author was supported in part by NSF under grants IIS-9988095 and DMS 0214171 and of the second author by NSF under grant DMS-0204247.

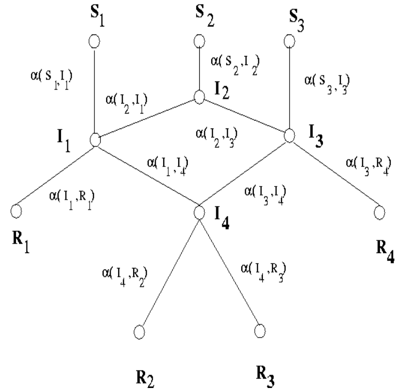

Figure 2: A multi-rooted logical network topology

References

Coates, M.J, Hero, A., Nowak, R.M. and Yu, B. (2002), Internet tomography. *IEEE Signal Processing Magazine*, 19, 47-65

Lawrence, E., Michailidis, G. and Nair, V.N. (2003a), Maximum likelihood estimation of internal network link delay distributions using multicast measurements *Proceedings of the 37th Conference on Information Sciences and Systems*, Baltimore, MD

Lawrence, E., Michailidis, G. and Nair, V.N. (2003b), Estimation of network link delay distributions using bicast probing experiments, Technical Report # 398, Department of Statistics, The University of Michigan

Liang, G. and Yu, B. (2003), Maximum pseudo-likelihood estimation in network tomography, *IEEE Transactions on Signal Processing*, 51, 243-253

Lo Presti, F. and Duffield, N. G. and Horowitz, J. and Towsley, D. (2002), Multicast-based inference of network-internal delay distributions, *ACM/IEEE Transactions on Networking*, 10, 761-775

Michailidis, G., Xi, B. and Nair, V.N. (2003), Fast least squares estimates of network internal losses using k-cast end-to-end measurement schemes, Technical Report # 399, Department of Statistics, The University of Michigan

Shih, C. and Hero, A. (2003), Unicast-based inference of network link delay distributions using mixed finite mixture models, *IEEE Transactions on Signal Processing*, 10, 2219-2228

Tsang, Y., Coates, M. and Nowak, R. (2003), Network delay tomography, *IEEE Transactions on Signal Processing*, 10, 2125-2136

Xi, B., Michailidis, G. and Nair, V.N. (2003a), Least squares estimates of network link loss probabilities using end-to-end multicast measurements, *Proceedings of the 37th Conference in Information Sciences and Systems*, Baltimore, MD

Xi, B., Michailidis, G. and Nair, V.N. (2003b), Estimating network internal losses using bicast end-to-end measurement schemes, Technical Report #397, Department of Statistics, The University of Michigan

Xi, B. (2004), Estimating network internal losses through end-to-end efficient measurement schemes, Doctoral Dissertation, Department of Statistics, The University of Michigan