

Multicast-Based Loss Inference with Missing Data

N.G. Duffield, J. Horowitz, D. Towsley, W. Wei, T. Friedman

Abstract

Network tomography using multicast probes enables inference of loss characteristics of internal network links from reports of end-to-end loss seen at multicast receivers. In this paper we develop estimators for internal loss rates when reports are not available on all probes or from all receivers. This problem is motivated by the use of unreliable transport protocols, such as RTP, to transmit loss reports to a collector for inference. We use a maximum likelihood (ML) approach in which we apply the Expectation Maximization (EM) algorithm to provide an approximating solution to the the ML estimator for the incomplete data problem. We present a concrete realization of the algorithm that can be applied to measured data. For classes of models we establish identifiability of the probe and report loss parameters, and convergence of the EM sequence to the MLE. Numerical results suggest that these properties hold more generally. We derive convergence rates for the EM iterates, and the estimation error of the MLE. Last, we evaluate the accuracy and convergence rate through extensive simulations.

Keywords: End-to-end Measurement, Network Tomography, Missing Data, Maximum Likelihood Estimation, EM Algorithm, Multicast, RTP, RTCP.

I. INTRODUCTION

A. Motivation

As the Internet grows in size and diversity, its internal performance becomes ever more difficult to measure. Any one organization has administrative access to only a small fraction of the network's internal nodes, whereas commercial factors often prevent organizations from sharing internal performance data. One promising technique that avoids these problems, *Multicast Inference of Network Characteristics* (MINC), uses end-to-end multicast measurements to infer link-level loss rates and delay statistics by exploiting the inherent correlation in performance observed by multicast receivers. These measurements do not rely on administrative access to internal nodes since the inference can be calculated using only information recorded at the end hosts. Efficient inferencing algorithms are given in [3] for loss, [16] for delay distributions, [10] for delay variances, and [4] for inferring the logical multicast tree topology itself. Extensions of these ideas to unicast (where multicast is replaced by a packet pair [6] or a packet stripe [11]) have also been proposed.

All of the algorithms based on the MINC methodology rely on the availability of complete information from the receivers. Providing such measurements from a dedicated infrastructure using reliable transmission of measurements to an inference engine raises issues of cost and scalability. An alternative lightweight approach the extension of RTCP, the RTP [20] control protocol, to provide extended loss reports [12], [2]. By piggybacking MINC loss reports on a standard transport protocol, one can effectively co-opt regular applications and their traffic to form a lightweight impromptu measurement infrastructure that encompasses many host end-points. However, loss reports are typically transmitted unreliably. Furthermore, the RTP standard imposes a constraint on the bandwidth that can be used by RTCP packets. Thus, this deployment will result in the availability of only *incomplete data sets* for the purpose of network inference. This motivates the need to modify the inferencing techniques to be able to handle incomplete data sets.

This work was supported in part by DARPA and the AFL under agreement F30602-98-2-0238

N.G. Duffield is with AT&T Labs–Research, 180 Park Avenue, Florham Park, NJ 07932, USA. E-mail: duffield@research.att.com. J. Horowitz is with the Dept. of Math. and Statistics, University of Massachusetts, Amherst, MA 01003, USA. E-mail: joe@math.umass.edu. D. Towsley, W. Wei, T. Friedman are with the Dept. of Computer Science, University of Massachusetts, Amherst, MA 01003, USA. E-mail: {towsley,weiwei,friedman}@cs.umass.edu

B. Contribution

In this paper we adapt the multicast inference techniques of [3] to perform inference of internal link loss characteristics when data is missing from some of the receivers. The data for inference comprises measured end-to-end loss of multicast probes sent from a source to a number of destinations but where only a subset of the destinations report observations for each multicast probe. As in [3], we use a Maximum Likelihood Estimator (MLE) arising from the corresponding likelihood function for the link probabilities. In contrast [3], it is not generally possible to determine the MLE by simple root finding when data is missing. Instead, we use the Expectation Maximization (EM) algorithm [8] to generate an approximating sequence to the corresponding MLE. We now outline the remainder of the paper and the detailed contributions.

In Section II we set up the models for the multicast tree, probe propagation, and report loss, and review some results for loss inference from complete data from [3]. We describe model frameworks for missing data and give two examples: inference using unicast stripes [11], and inference using extended RTCP reports, as proposed in [1]. In Section III we set up the incomplete data likelihood function, and describe the EM algorithm and its application to the present model. We establish conditions required for convergence of the EM iterates to the MLE. We translate these into conditions on the measured data. In Section IV we tailor the EM algorithm to our specific problem and present an algorithm for use on measured data. Section V addresses conditions for identifiability of model parameters. Convergence of the MLE as the number of probes grows is investigated in Section VI; in particular we obtain expressions for the asymptotic variance of the MLE for a class of models. A related expression for the convergence rate of the EM iterates is obtained in Section VII. The algorithm from Section IV is evaluated in model-based simulation and using experimentally derived traces in Section VIII. We conclude in Section IX. Details of most proofs are deferred until Section X.

C. Related work

Several tools and methodologies exist for characterizing link-level behavior from end-to-end multicast measurements. However, most of these, including the MINC methodologies referred to earlier, require complete data from all of the receivers in the multicast tree. These methodologies have been adapted to unicast through the transmission of packet pairs [6] or stripes [11] pairs of receivers. The data then consists of observations from pairs of receivers and can be interpreted as observations in which the data is missing from all but these pairs of receivers. The method presented in [11] disjoins inference into separate problems for each pair of receivers, producing link estimates by averaging over all estimates produced from each receiver pair.

In [6], the authors introduce an additional link parameter, the conditional probability that the second packet within a pair is not lost given that the first packet is not lost. The authors then treat the outcomes of the each of the packets in a pair within the tree as unobserved data and use the EM algorithm to infer the link probabilities and conditional link probabilities. Due to the complexity of this task, they propose a heuristic for inferring these parameters. Because we rely on multicast, we only have one set of link parameters to infer. Our solution methodology uses the EM algorithm to obtain a solution to the likelihood equation. Coates and Nowak have extended their EM-based, unicast-based techniques to infer delay statistics in [7].

Last, several approaches infer round trip link behavior, including *pathchar* [9], [13] and the linear algebraic approach of [21]. The former infers loss, delay, and available link bandwidth; the latter infers round trip link delays. The former requires considerable time to converge. Both lose accuracy with asymmetric round trip paths.

II. MODELS FOR PROBE AND REPORT TRANSMISSION

A. Tree model

Let $\mathcal{T} = (V, L)$ denote a logical multicast tree with nodes V and links L . We identify the root ρ as the probe source, and the set of leaves $R \subset V$ with the set of receivers. We assume that the root has single child, denoted by 1. If not, then we can treat separately the trees descended through each child of ρ , each one having this property. Each node k , apart from the root ρ , has a parent $f(k)$ such that $(f(k), k)$ is a link in L . We sometimes refer to the link $(f(k), k)$ that terminates at k simply as link k . Define recursively the ancestors of k by $f^n(k) = f(f^{n-1}(k))$ with $f^0(k) = k$. We say j is descended from k , and write $j \prec k$, if $k = f^n(j)$ for some $n \in \mathbb{N}$. The set of children of k , namely $\{j \in V : f(j) = k\}$ is denoted by $d(k)$. $\mathcal{T}(k) = (V(k), L(k))$ denotes the subtree rooted at k ; $R(k) = R \cap V(k)$ is the set of receivers in $\mathcal{T}(k)$. Define $U = V \setminus \{\rho\}$.

B. Packet loss model

We assume a Bernoulli loss model in which probes are independent and each probe is successfully transmitted across link k with probability α_k . Thus the progress of each probe down the tree is described by an independent copy of a stochastic process $X = (X_k)_{k \in V}$. $X_\rho = 1$. $X_k = 1$ if the probe reaches node $k \in V$ and 0 otherwise. If $X_k = 0$, then $X_j = 0, \forall j \prec k$. Otherwise, $P[X_j = 1 | X_{f(j)} = 1] = \alpha_j$ and $P[X_j = 0 | X_{f(j)} = 1] = 1 - \alpha_j$. By convention $\alpha_\rho = 1$. Denote $\alpha = (\alpha_i)_{i \in V}$. \mathbf{P}_α is the distribution of X .

C. Inference of link loss from complete data

When a probe is sent down the tree from the root ρ , we cannot observe the whole process X . We assume that, at most, we know only the outcome $(X_k)_{k \in R} \in \Omega = \{0, 1\}^R$ that indicates whether or not the probe reached each receiver. When the entire outcome for a probe is known (i.e. X_k for all receivers k), we say that we have complete data from that probe. In [3] it was shown how the link probabilities can be determined from the the distribution of (complete) outcomes. We briefly review this.

Consider an experiment in which n probes are dispatched from the root ρ . Each probe $i = 1, \dots, n$ gives rise to an independent realization $X^{(i)}$ of the probe process X . We call

$$X_{\text{cplt}} = (X_k^{(i)})_{k \in R}^{i=1, \dots, n} \quad (1)$$

the complete data for the experiment. For each outcome $x \in \Omega$, let $n(x)$ denote the number of probes $i = 1, \dots, n$ for which $X_k^{(i)} = x_k$ for all $k \in R$. Let

$$p_\alpha(x) = \mathbf{P}_\alpha[X_k = x_k, \forall k \in R] \quad (2)$$

denote the probability of an outcome $x \in \Omega$. The *complete* data log-likelihood to obtain the data $X_{\text{cplt}} = (X^{(1)}, \dots, X^{(n)})$ can be written in terms of the $n(x)$ as

$$\mathcal{L}_c(\alpha) = \log \mathbf{P}_\alpha[X_{\text{cplt}}] = \sum_{x \in \Omega} n(x) \log p_\alpha(x). \quad (3)$$

We characterize the Maximum Likelihood Estimator (MLE) of α , namely, $\arg \max_\alpha \mathcal{L}(\alpha)$, as follows. For $k \in V$, let A_k be the probability that the probe reaches k . Thus $A_k = \prod_{j \succeq k} \alpha_k$, the product of the probabilities of successful transmission on each link between k and the root ρ . For each $k \in U$ set

$$\gamma_k = \mathbf{E}_\alpha[\mathbb{1}_{\{j \in R(k)\}} X_j] \quad (4)$$

i.e., γ_k is the probability that a probe reaches at least one receiver descended from node k . Denote by $\hat{\gamma}_k$ the corresponding empirical quantity, i.e., the proportion of the n probes that reach at least one leaf descended from k :

$$\hat{\gamma}_k = n^{-1} \sum_{i=1}^n \mathbb{V}_{j \in R(k)} X_j^{(i)}. \quad (5)$$

In what follows we consider α to lie in the open parameter set $\mathcal{A} = \{\alpha \mid \alpha_k \in (0, 1), k \in U\}$. Some of the results of the following theorem also hold on subsets of the boundary of \mathcal{A} .

Theorem 1 ([3]) Assume $\alpha \in \mathcal{A}$.

- (i) For each $k \in U$, $(1 - \gamma_k/A_k) = \prod_{j \in d(k)} (1 - \gamma_j/A_j)$, where empty product for $k \in R$ is taken as 0.
- (ii) Let $\mathcal{G} = \{(\gamma_k)_{k \in U} : \gamma_k > 0 \forall k; \gamma_k < \sum_{j \in d(k)} \gamma_j \forall k \in U \setminus R\}$. For each $\gamma \in \mathcal{G}$ and $k \in U$, the equation in (i) has a unique solution, denoted $\mathcal{H}_k(\gamma)$, in the interval $(\gamma_k, 1)$.
- (iii) If $\hat{\gamma} \in \mathcal{G}$, the likelihood equation $\partial \mathcal{L}_c / \partial \alpha_k = 0$, $k \in U$ has unique solution $\hat{\alpha}_k = \mathcal{K}_k(\hat{\gamma}) := \mathcal{H}_k(\hat{\gamma}) / \mathcal{H}_{f(k)}(\hat{\gamma})$, $k \in U$.
- (iv) With probability one, for sufficiently large n , both $\hat{\alpha}$ and the MLE of α lie in \mathcal{A} , and are hence equal.
- (v) The parameters α are identifiable, i.e., $P_\alpha = P_{\alpha'}$ for $\alpha, \alpha' \in \mathcal{A}$ implies $\alpha = \alpha'$.

It turns out that Theorem 1(iv) is weaker than required for the present paper. We now establish a stronger version that provides a test as to whether or not $\hat{\alpha}_k$ is the MLE for finite n .

Theorem 2: Assume $\hat{\gamma} \in \mathcal{G}$. If $\hat{\alpha} \in \mathcal{A}$, it is the MLE for α .

Theorem 1(iv) establishes that for n sufficiently large, the MLE lies in \mathcal{A} and hence must be $\hat{\alpha}$, the solution of the likelihood equation. Theorem 2 is more useful from the computational point of view; it says that provided $\hat{\alpha}$ lies in \mathcal{A} , a condition that can be checked by inspection, it is the MLE, for any n . As a consequence of the MLE property, $\hat{\alpha}$ is consistent ($\hat{\alpha} \xrightarrow{n \rightarrow \infty} \alpha$ with probability 1), and asymptotically normal ($\sqrt{n}(\hat{\alpha} - \alpha)$ converges in distribution to a multivariate Gaussian random variable as $n \rightarrow \infty$); see e.g. [19].

D. Missing data model

We now want to generalize the problem by admitting the possibility that some outcomes may not be completely known because the receiver variables are missing. Let $T = (T_k^{(i)})_{k \in R}^{i=1, \dots, n}$ denote the $n \times \#R$ matrix of missing data indicators, with $T_k^{(i)}$ taking the value 0 if the variable $X_k^{(i)}$ is missing, and $T_k^{(i)} = 1$ if it is present. The set of observed data and missing data are thus, respectively,

$$X_{\text{obs}} = \{X_k^{(i)} \mid T_k^{(i)} = 1\} \quad \text{and} \quad X_{\text{mis}} = \{X_k^{(i)} \mid T_k^{(i)} = 0\}. \quad (6)$$

In this paper we assume that the missing data mechanism is ignorable in a sense we now describe; see [15] for further details. We treat T as a random variable whose distribution is parameterized by some quantity θ . P_θ will denote the distribution of T under θ , and $P_{\alpha, \theta}$ the joint distribution of X_{cpld} and T . We henceforth assume that the missing data is *missing at random* (MAR). This is the property that the distribution of the missing-data mechanism T does not depend on the missing values X_{mis} . More formally, we can write the MAR property as $P_\theta[T \mid X_{\text{obs}}, X_{\text{mis}}] = P_\theta[T \mid X_{\text{obs}}]$. As a consequence of MAR, the joint distribution of the observed data and the missing-data mechanism enjoys the following factorization property:

$$P_{\theta, \alpha}[X_{\text{obs}}, T] = P_\theta[T \mid X_{\text{obs}}] P_\alpha[X_{\text{obs}}]. \quad (7)$$

Assuming the parameters (α, θ) to be distinct with product parameter space $\mathcal{A} \times \Theta$, (7) says that the missing data mechanism is ignorable in that likelihood-based inference for α based on the joint likelihood $\mathbb{P}_{\theta, \alpha}[X_{\text{obs}}, T]$ are the same as those based upon $\mathbb{P}_{\alpha}[X_{\text{obs}}]$. Thus for purposes of inferring α , we may ignore the parameters θ of the missing data mechanism. A special case of MAR is data *missing completely at random* (MCAR). With MCAR the missingness probabilities do not depend on any data: $\mathbb{P}_{\theta}[T | X_{\text{obs}}] = \mathbb{P}_{\theta}[T]$.

E. Example: Inference using unicast data.

In [11], the authors describe an approach to *unicast* based inference in which n sets of packets, known as stripes, are transmitted by a source to all receiver pairs. The motivation is that within each stripe, packets are transmitted back-to-back, and so their loss behavior on common links should be highly correlated. With perfect correlations (i.e. both packets being either transmitted or lost on a common link) the stripe has the same behavior as a notional multicast packet that follows the same route and is subject to the same loss.

We can put each receiver pair in correspondence with a missing data indicator as follows. $T^{(i)} = (T_k^{(i)})_{k \in R}$ identifies the pair of receivers corresponding to the i -th stripe, i.e., $T_j^{(i)} = T_k^{(i)} = 1, T_l^{(i)} = 0, l \in R, l \neq j, k$ if the pair of receivers is $j, k \in R, j \neq k$. Thus missingness of data from probe i at receiver j occurs because j is not a member of the pair of receiver nodes selected for the probe.

If the receiver pairs are chosen independently from stripe to stripe using the same distribution, then $T = (T^{(i)})_{i=1}^n$ is a sequence of IID random variables. Thus T has the following distribution,

$$\mathbb{P}[T = t] = \prod_{i=1}^n \mathbb{P}[T^{(i)} = t^{(i)}], \quad \forall t \in \{0, 1\}^{\#R} \quad (8)$$

where

$$\mathbb{P}[T^{(i)} = t] = \sum_{j, k \in R; j \neq k} \mathbf{1}\{t_j = 1\} \mathbf{1}\{t_k = 1\} \times \prod_{l \in R \setminus \{j, k\}} \mathbf{1}\{t_l = 0\} p_{j, k}, \quad \forall t \in \{0, 1\}^{\#R} \quad (9)$$

Here $p_{j, k}$ is the probability that the receiver pair j, k is chosen. If we further assume that T is independent of X , the data is MCAR. One can formulate variations, e.g., the sender cycles round robin through the pairs.

F. Example: Inference using RTP/RTCP.

The Reliable Transport Protocol (RTP) [20] enables transfer of data from a single sender to one or more receivers. Associated with it is a control protocol RTCP that allows receivers to multicast loss behavior to each other and to a third party. Typically, the observations are batched and each batch is sent as a single report. The third party can collect the observations and apply inference methods to them. However, these reports are not transmitted reliably. Consequently, the data collector must deal with missing information.

In the current implementation of RTCP, receivers broadcast only average loss rates. Extensions to the protocol, proposed in [1], enable receivers to report on the reception of individual packets. However, due to the constraints imposed on reporting volumes by RTCP, it may not be possible to report on every packet. The omission of certain reports to fulfill this constraint is thus an additional source of missingness.

We propose a simple model for this scenario. Consider receiver $j \in R$ that collects loss observations and sends them to a data collector. Let $\{\{A_{i, j}\}_{i=1}^{\infty}\}_{j \in R}$ be a set of random variables where $A_{i, j}$ is the number of observations placed in the i -th report by the j -th receiver. Let $\{\{C_j^{(k)}\}_{k=1}^{\infty}\}_{j \in R}$ be indicator random variables

that represent the outcome of the transmission of the k -th loss report by receiver j to the data collector; it is received by the receiver if $C_j^{(k)} = 1$ and lost otherwise. Define $\pi(i, j) = \min\{l : \sum_{k=1}^{l-1} A_{k,j} < i \leq \sum_{k=1}^l A_{k,j}\}$, i.e., $\pi(i, j)$ identifies which report the i -th observation from receiver j is placed in. Let $\{\{S_j^{(i)}\}_{i=1}^\infty\}_{j \in R}$ be a set of indicator random variables that represent whether probe i was actually selected for reporting from receiver j ; it is selected if $S_j^{(i)} = 1$. Then the missing data indicator $T_j^{(i)}$, $i = 1, \dots, n$, $j \in R$ can be expressed as $T_j^{(i)} = S_j^{(i)} C_j^{(\pi(i,j))}$.

Under strong simplifying assumptions, namely that the random variables A, S and C are independent of X , the model is MCAR. However we can posit a situation in which independence may not hold in practice. Suppose the collector lies at a node k in the multicast tree. Then the path for reports from receivers in $R \setminus R(k)$ to k intersects with the paths of probe packets from ρ to receivers in $R(k)$. Thus we may expect the missingness variables $\{T_j^{(i)} : j \in R \setminus R(k)\}$ to be correlated with the receiver state $\{X_j : j \in R(k)\}$. This is precisely the type of model allowed when data is MAR.

G. Approaches to the problem of missing data

A simple approach to manage missing data would be to restrict inference to subsets of probes and receivers for which complete data is available, then patch together such estimators to draw inference on the complete tree. A minimal way would be to use only probes for which reports were received from all receivers. A more sophisticated variant would be to perform inference on subtrees using probes whose reports were received from each receiver on that subtree. However, such ‘‘patchwork’’ approaches have three pitfalls:

- (i) It will not be possible to perform inference on subtrees for which no probes have reports from all receivers.
- (ii) Such estimators are not consistent unless data is MCAR. Furthermore, checking whether data is consistent with MCAR is complex; the number of consistency conditions to be checked grows exponentially with the number of leaves in the tree.
- (iii) Such estimators are not generally efficient, even if data is MCAR.

For these reasons we instead extend the previous ML approach to cover the missing data case directly: under general conditions ML-estimators are consistent and efficient. This is the subject of the next section.

III. ESTIMATION OF LINK RATES FROM INCOMPLETE DATA

In this section we present the likelihood function \mathcal{L} for the incomplete data. Determination of the corresponding ML estimator for the link probabilities turns out to be significantly more complex than in the complete data case. We turn to a standard iterative method, the EM algorithm, to derive an approximating sequence to the incomplete data MLE.

A. Description of incomplete data and the likelihood function

The corresponding *incomplete* data likelihood function is the marginal distribution function of the observed data; formally we write this as $\int P_\alpha[X_{\text{obs}}, X_{\text{mis}}] dX_{\text{mis}}$. We now obtain an explicit expression. In order to represent both missing and observed data in a compact form, we extend the set of outcomes to the set $\Omega^* = \{0, 1, u\}^R$, where u is used to denote that a given receiver datum is missing. $u^* = (u, \dots, u) \in \Omega^*$ will denote the outcome in which data is missing from all receivers. Let $t \in \{0, 1\}^R$ denote the generic vector of

missing data indicator variables. With each such t and $x \in \Omega$ we then associate an element $x(t)$ of Ω^* through

$$x_k(t) = \begin{cases} \mathbf{u} & \text{if } t_k = 0 \\ x_k & \text{otherwise} \end{cases}, \quad k \in R. \quad (10)$$

An inverse of the above map associates with $x^* \in \Omega^*$ its missing data indicator $t(x^*)$ by

$$t_k(x^*) = \begin{cases} 0 & \text{if } x_k^* = \mathbf{u} \\ 1 & \text{otherwise} \end{cases}, \quad k \in R. \quad (11)$$

The set of complete outcomes x that can give rise to an incomplete outcome x^* is the set

$$\Omega(x^*) = \{x \in \Omega \mid x_k^* = x_k \Leftrightarrow t_k(x^*) = 1\}, \text{ and conversely } \Omega^*(x) = \{x^* \in \Omega^* \mid x \in \Omega(x^*)\} \quad (12)$$

is the set of complete outcomes x^* that can be obtained from a complete outcome x . The equivalent conditions $x \in \Omega(x^*)$ and $x^* \in \Omega^*(x)$ can be rewritten as $x(t(x^*)) = x^*$.

The probability to record an incomplete outcome $X^{(i)}(T^{(i)}) = x^*$ is denoted

$$q_{\alpha, \theta}(x^*) = \mathbf{P}_{\alpha, \theta}[X^{(i)}(T^{(i)}) = x^*]. \quad (13)$$

Now $\{X^{(i)}(T^{(i)}) = x^*\} = \{X^{(i)} \in \Omega(x^*)\} \cap \{T^{(i)} = t(x^*)\}$. Using the MAR property (7) we factorize

$$q_{\alpha, \theta}(x^*) = p_{\alpha}^*(x^*)\theta(x^*) \quad (14)$$

where

$$p_{\alpha}^*(x^*) = \mathbf{P}_{\alpha}[X^{(i)} \in \Omega(x^*)] = \sum_{x \in \Omega(x^*)} p_{\alpha}(x) \text{ and } \theta(x^*) = \mathbf{P}_{\theta}[T^{(i)} = t(x^*) \mid X^{(i)} \in \Omega(x^*)]. \quad (15)$$

Without loss of generality we have taken the missingness probabilities themselves as parameters θ . Note that by the MAR property, for any $x \in \Omega(x^*)$, $\mathbf{P}_{\theta}[T^{(i)} = t(x^*) \mid X^{(i)} \in \Omega(x^*)] = \mathbf{P}_{\theta}[T^{(i)} = t(x^*) \mid X^{(i)} = x]$. Since $\{t(x^*) \mid x^* \in \Omega^*(x)\} = \{0, 1\}^R$, the conditional probabilities θ satisfy

$$\sum_{x^* \in \Omega^*(x)} \theta(x^*) = 1, \quad \forall x \in \Omega. \quad (16)$$

Now let $m(x^*)$ denote the number of probes $i = 1, \dots, n$ for which $X^{(i)}(T^{(i)}) = x^*$. Due to the factorization property (7), the log-likelihood function $\log \prod_{i=1}^n q_{\alpha, \theta}(X^{(i)}(T^{(i)}))$ can be written as a sum of

$$\mathcal{L}(\alpha) = \sum_{x^* \in \Omega^*} m(x^*) \log p_{\alpha}^*(x^*), \quad (17)$$

with a term that is independent of α . Thus, for the purposes of obtaining an ML estimate of α , we need only consider $\mathcal{L}(\alpha)$. We refer to \mathcal{L} as the incomplete data likelihood function. Note that the term in $m(\mathbf{u}^*)$ makes no contribution to \mathcal{L} since $\Omega(\mathbf{u}^*) = \Omega$ and hence $p_{\alpha}^*(\mathbf{u}^*) = 1$. Hence the sum in (17) can be restricted to $\Omega_0^* = \Omega^* \setminus \{\mathbf{u}^*\}$.

B. Application of the EM algorithm

We can in principle estimate the link probabilities by the incomplete data MLE $\check{\alpha} = \arg \max_{\alpha} \mathcal{L}(\alpha)$. However, we have been unable to obtain a direct solution to the incomplete-data likelihood equation. Instead, we employ a standard statistical method, the Expectation-Maximization (EM) algorithm [8], to obtain an iterative approximation $\hat{\alpha}^{(\ell)}$, $\ell = 0, 1, \dots$ to a stationary value of the incomplete data likelihood. The algorithm comprises the following steps:

(i) *Initialization.* Pick some initial link probabilities $\alpha^{(0)}$. This could be done, e.g., by setting $\hat{\alpha}^{(0)} = \hat{\alpha}$, the complete data MLE determined from the counts of complete outcomes \mathbf{m} if these are non-zero. Without complete data, we can use the fact (see the proof of Theorem 5 in [3]) that $\gamma_k = A_k + O(\|\bar{\alpha}\|^2)$ to approximate $\alpha_k = A_k/A_{f(k)} \approx \gamma_k/\gamma_{f(k)} = (1 - \bar{\gamma}_k)/(1 - \bar{\gamma}_{f(k)}) \approx 1 + \gamma_k - \gamma_{f(k)}$. This suggests the initial value

$$\hat{\alpha}_k^{(0)} = 1 + \hat{\gamma}_k - \hat{\gamma}_{f(k)}. \quad (18)$$

(ii) *Expectation.* For each $\hat{\alpha}^{(\ell)}$ find the conditional expectation of the complete log-likelihood given the incomplete data $Q(\alpha', \hat{\alpha}^{(\ell)}) = \mathbb{E}_{\hat{\alpha}^{(\ell)}}[\mathcal{L}_c(\alpha') \mid \mathbf{m}]$.

(iii) *Maximization.* Find the maximizer of the condition expectation: $\hat{\alpha}^{(\ell+1)} = \arg \max_{\alpha'} Q(\alpha', \hat{\alpha}^{(\ell)})$

(iv) *Iteration.* Iterate steps (ii) and (iii) until some termination criterion is satisfied.

For $k \in V$, define the conditional probabilities for a probe to reach $R(k)$ as

$$\hat{\gamma}_{k,\alpha} = \mathbb{E}_{\alpha}[\mathbb{V}_{j \in R(k)} X_k \mid \mathbf{m}]. \quad (19)$$

For notational convenience we write the conditional probability $\hat{\gamma}_{k,\hat{\alpha}^{(\ell)}}$ derived from the iterate $\hat{\alpha}^{(\ell)}$ as $\hat{\gamma}_k^{(\ell)}$.

Theorem 3: Assume $\hat{\gamma}^{(\ell)} \in \mathcal{G}$. Then

$$\alpha_k^{(\ell+1)} = \mathcal{K}_k(\hat{\gamma}^{(\ell)}), \quad k \in U \quad (20)$$

provided that $\mathcal{K}(\hat{\gamma}^{(\ell)})$ lies in \mathcal{A} .

We now investigate convergence of the iterates $\hat{\alpha}^{(\ell)}$. Whereas the complete data likelihood function can be shown to derive from a standard exponential family (see the proof of Theorem 2), the incomplete data likelihood function derives only from a curved exponential family. Thus we cannot use results based on standard exponential families (see e.g. [22]) alone to conclude convergence of $\hat{\alpha}^{(\ell)}$ to $\check{\alpha}$. We now establish conditions under which the sequence exists in \mathcal{A} and converges to the MLE for the incomplete data problem.

Theorem 4: Assume $\hat{\gamma}^{(\ell)} \in \mathcal{G}$ and $\mathcal{K}(\hat{\gamma}^{(\ell)}) \in \mathcal{A}$ for all ℓ .

(i) $\mathcal{L}(\hat{\alpha}^{(\ell)})$ converges to some limit L .

(ii) If $\{\alpha \in \mathcal{A} \mid \mathcal{L}(\alpha) = L\}$ is discrete, $\hat{\alpha}^{(\ell)}$ converges to some α^* that is a stationary point \mathcal{L} , i.e. $\frac{\partial \mathcal{L}}{\partial \alpha}(\alpha^*) = 0$.

(iii) If \mathcal{L} is unimodal, $\hat{\alpha}^{(\ell)}$ converges to $\check{\alpha}$.

C. Calculation of the EM iterates

An algorithm for the calculating $\mathcal{K}_k(\gamma)$ for a given $\gamma \in \mathcal{G}$ has been detailed in [3]. It remains then to provide an algorithm for the calculation of the $\hat{\gamma}_{k,\alpha}$. Let $n_0 = n - m^*(\mathbf{u}^*)$ denote the number of probes for which the data is not entirely missing. Observe that

$$\hat{\gamma}_{k,\alpha} = \sum_{x^* \in \Omega_0^*} \frac{m(x^*)}{n_0} \hat{\gamma}_{k,\alpha}(x^*) \quad \text{where} \quad \hat{\gamma}_{k,\alpha}(x^*) = \mathbb{E}_{\alpha}[\mathbb{V}_{j \in R(k)} X_k \mid X^* = x^*]. \quad (21)$$

Let $R(k, x^*) = \{j \in R(k) \mid t_j(x^*) = 1\}$ denote the receivers descended from k from which data is observable. Order the elements of the set $\{0, 1, u\}$ as $u < 0 < 1$ and extend the usual maximum operator \vee on $\{0, 1\}$ to an operation \vee^* on $\{0, 1, u\}$, respecting the order in an obvious manner. Let $h(k, x^*)$ denote the closest ancestor h of k for which a packet has been observed to reach at least one descendant leaf, i.e.,

$$h(k, x^*) = \inf_{j \succeq k} \{j : y_j^* = 1\}, \quad (22)$$

where $y_k^* = \vee_{j \in R(k)}^* x_j^*$. (Note $y_k^* = u$ if x indicates all data from $R(k)$ is missing; $y_k^* = 1$ if x indicates a probe observed to reach at least one receiver in $R(k)$; otherwise $y_k^* = 0$.) When $k \prec j$, let $d(j, k)$ denote that child of j that is an ancestor (or possibly equal to) k , i.e., $d(j, k) = \{i \in d(j) : i \succeq k\}$.

Theorem 5: When $y_k^* = 1$, $\hat{\gamma}_{k,\alpha}(x^*) = 1$. When $y_k^* < 1$,

$$\hat{\gamma}_{k,\alpha}(x^*) = \frac{c_k - b_k}{c_{d(h,k)}} \prod_{k \prec i \preceq d(k,h)} \left\{ \alpha_i \prod_{j \in d(i) \setminus d(i,k)} c_j \right\} \quad (23)$$

where $h = h(k, x^*)$, and for $k \preceq i \prec h$,

$$b_k = \mathbf{P}_\alpha[\vee_{j \in R(k)} X_j = 0 \mid X_{f(k)} = 1], \quad c_k = \begin{cases} 1, & \text{if } R(k, x^*) = \emptyset \\ \mathbf{P}_\alpha[\vee_{j \in R(k, x^*)} X_j = 0 \mid X_{f(k)} = 1] & \text{otherwise} \end{cases} \quad (24)$$

It was found in [3] that the problem of determining the α for a tree with complete data factors into the problem of solving a set of depth two tree inference problems, one for each node $k \in V \setminus R$. For each leaf k one constructs the logical tree with root ρ having single child k , and $d(k)$ leaf-children. Furthermore, for a general tree, the problem could be mapped onto that for a binary tree by the insertion of lossless links. However, this method cannot be applied when there is missing data. This is because the form (23) for $\hat{\gamma}_{k,\alpha}^{(\ell)}$ includes variables from receivers other than those descended from k .

D. Topology and data conditions

Theorems 3 and 4 required the iterators $\hat{\gamma}^{(\ell)}$ to lie in the domain \mathcal{G} . In this section we specify conditions on the data in order for these requirements to hold. When the conditions do not hold, in some cases we can adjust the problem by passing to one or more subtrees of the original tree.

D.1 Non-identifiable subtrees.

We eliminate from consideration subtrees on which no data is missing but whose leaves were reached by no probes. For $k \in V$, let D_k denote the event that for some probe i , $X_j^{(i)}(T^{(i)}) \neq 0$ for some $j \in R(k)$. We will assume

$$D_k \text{ occurs for all } k \in V \quad (25)$$

If (25) does not hold, the following procedure reduces the inference problem to one for which it does. If D_k fails, we remove from further consideration the subtree $\mathcal{T}(k)$ rooted at k . If this pruning leaves the parent $f(k)$ with only one offspring j , the remaining tree is no longer a logical multicast tree. To make it so we remove the link $(f(k), j)$ and identify the nodes j and $f(k)$. Consequently we can only be able to identify the characteristics of the composite link joining j to $f^2(j)$ of the original tree. Performing these operations for all k at which D_k fails, we obtain a tree for which (25) holds. When $e\mathcal{T}(k)$ is pruned, we flag all $\{\alpha_j : j \preceq k\}$ as unknown, except when k is a leaf but $D_{f(k)}$ holds. In this case we may estimate $\hat{\alpha}_k = 0$.

D.2 Links with perfect transmission.

Let D'_k denote the complement of the event $\{X_j^{(i)}(T^{(i)}) = 1, \forall j \in R(k), i = 1, \dots, n\}$. When D'_k fails, lossless transmission is reported for all probes to all receivers in $R(k)$. The effect is to position $\mathcal{K}(\hat{\gamma}_\alpha)$ on the boundary of \mathcal{A} , since it follows that $\mathcal{K}_j(\hat{\gamma}_\alpha) = 1$ for all $j \in R(k)$. Although this is not a problem for computation, it would us out of the domain of application of Theorems 2, 3 and 4. As a remedy, we can set $\hat{\alpha}_j = 1$ for all nodes $j \in V(k)$ where D'_k fails, and restrict consideration to the complementary α_i in \mathcal{A} .

D.3 Model Conformance.

We also need a condition to ensure that estimated quantities $\hat{\gamma}$ lie in \mathcal{G} . Let D''_k be the event that k has children $j, \ell \in d(k)$ such that $X_j^{(i)}(T^{(i)}) = 1$ and $X_\ell^{(i)}(T^{(i)}) \neq 0$. If D''_k fails each probe is observed on no more than one subtree of k , and hence the conditional probe losses on different subtrees descended from k are dependent, given a probe reaching k . Henceforth we assume

$$D''_k \text{ holds for all any } k \in V \setminus R. \quad (26)$$

If D''_k fails, we adjust the tree by removing the link $(f(k), k)$ from the tree and identifying its endpoints k and $f(k)$. In the original tree, we will only be able to identify the characteristics of the composite links joining $f(k)$ to the children $d(k)$. The procedure is iterated if necessary until (26) holds.

Conditions (25) and (26) enable us to fulfill some assumptions in Theorems 3 and 4. We will henceforth assume that they hold.

Lemma 1: When (25) and (26) hold, $\hat{\gamma}_\alpha \in \mathcal{G}$ for any $\alpha \in \mathcal{A}$.

E. Example: the two-receiver tree

In the simplest case we can establish unimodality of \mathcal{L}_c directly, and thus conclude convergence of the EM iterates to the incomplete data MLE. Consider the two receiver tree with root ρ having a single child 1 whose children are the leaf nodes 2 and 3. In the two receiver tree, we enumerate $\Omega = \{11, 01, 10, 00\}$ and $\Omega_0^* = \{11, 01, 10, 00, 1u, u1, 0u, u0\}$. It is not difficult to determine that the $\hat{\gamma}_{k,\alpha}(x^*)$ are as given in Table I. The EM iterates are then $(\mathcal{K}_1(\hat{\gamma}_\alpha), \mathcal{K}_2(\hat{\gamma}_\alpha), \mathcal{K}_3(\hat{\gamma}_\alpha))$ of $(\alpha_1, \alpha_2, \alpha_3)$ are then

$$\mathcal{K}_1(\hat{\gamma}_\alpha) = \frac{\hat{\gamma}_{2,\alpha}\hat{\gamma}_{3,\alpha}}{\hat{\gamma}_{2,\alpha} + \hat{\gamma}_{3,\alpha} - \hat{\gamma}_{1,\alpha}}, \quad \mathcal{K}_2(\hat{\gamma}_\alpha) = \frac{\hat{\gamma}_{2,\alpha} + \hat{\gamma}_{3,\alpha} - \hat{\gamma}_{1,\alpha}}{\hat{\gamma}_{3,\alpha}}, \quad \mathcal{K}_3(\hat{\gamma}_\alpha) = \frac{\hat{\gamma}_{2,\alpha} + \hat{\gamma}_{3,\alpha} - \hat{\gamma}_{1,\alpha}}{\hat{\gamma}_{2,\alpha}}$$

Theorem 6: In the two-receiver tree, the incomplete data likelihood function \mathcal{L} is unimodal, and hence $\hat{\alpha}^{(\ell)}$ converges to the incomplete data MLE provided that $\mathcal{K}(\hat{\gamma}^{(\ell)}) \in \mathcal{A}$ for all ℓ .

IV. NETWORK INFERENCE ALGORITHM

In order to carry out inference on measured data, we express the calculation of $\hat{\gamma}$ in Theorem 5 as a algorithm. We start by constructing b_k, c_k and y_k^* recursively. Clearly the b_k satisfy

$$b_k = \begin{cases} \bar{\alpha}_k, & k \in R, \\ \bar{\alpha}_k + \alpha_k \prod_{j \in d(k)} b_j, & k \in U \setminus R \end{cases} \quad (27)$$

The c_k satisfy a similar recursion:

$$c_k = \begin{cases} 1, & k \in R, x_k^* = \mathbf{u} \\ \bar{\alpha}_k, & k \in R, x_k^* \neq \mathbf{u} \\ \bar{\alpha}_k + \alpha_k \prod_{j \in d(k)} c_j, & k \in U \setminus R \end{cases} \quad (28)$$

x^*	$\hat{\gamma}_{1,\alpha}(x^*)$	$\hat{\gamma}_{2,\alpha}(x^*)$	$\hat{\gamma}_{3,\alpha}(x^*)$
11	1	1	1
10	1	1	0
01	1	0	1
00	0	0	0
1u	1	1	α_3
u1	1	α_2	1
0u	$\frac{\alpha_1 \bar{\alpha}_2 \alpha_3}{\bar{\alpha}_1 + \alpha_1 \bar{\alpha}_2}$	0	$\frac{\alpha_1 \bar{\alpha}_2 \alpha_3}{\bar{\alpha}_1 + \alpha_1 \bar{\alpha}_2}$
u0	$\frac{\alpha_1 \alpha_2 \bar{\alpha}_3}{\bar{\alpha}_1 + \alpha_1 \bar{\alpha}_3}$	$\frac{\alpha_1 \alpha_2 \bar{\alpha}_3}{\bar{\alpha}_1 + \alpha_1 \bar{\alpha}_3}$	0

TABLE I
COEFFICIENTS $\hat{\gamma}_{k,\alpha}(x^*)$ FOR THE TWO-RECEIVER TREE

The y_k^* satisfy the recursion

$$y_k^* = \begin{cases} x_k^*, & k \in R \\ \bigvee_{j \in d(k)}^* y_j^*, & k \in V \setminus R \end{cases} \quad (29)$$

We formally specify an algorithm for the calculation of the $\hat{\gamma}_{k,\alpha}$ in Figure 1. The main procedure comprises two phases. In the first phase, `set_abc`, calculates the y_k^* , b_k and c_k passing up the tree from the leaves. The second phase, `set_g`, then calculates the $\hat{\gamma}_{k,\alpha}$ traversing the tree from the root ρ downwards. h_k plays the role of $d(k, h)$ while e plays the role of $\alpha_i \prod_{j \in d(i) \setminus \{d(i,k)\}} c_j$. On a given path down the tree, `flag` = 1 until a node k with $y_k^* = 0$ is first encountered. `flag` = 0 on all calls to `set_g` below k . The identity of the node $h(i, x^*)$ is then maintained in calls at nodes i below the child j of k (lines 10–13).

We note there is some redundancy in the algorithms, which can be avoided in implementations. b_k and c_k need not be calculated at nodes k for which $y_k^* = 1$, since these values are not used. The c_k depend only on the missing data indicator $t(x^*)$, and so need be calculated once for each set incomplete outcomes $\{x^* \in \Omega^* : t(x^*) = t\}$ having the same missing data indicator t . The b_k do not depend on x^* , and so may be calculated once in advance; in particular $b_k = c_k$ when x^* has no missing data, i.e., when $x_k^* \neq u \forall k \in R$. Lastly, the y_k^* need only be calculated once for each probe with distinct x^* , and once at the start of the sequence of iterations.

V. IDENTIFIABILITY AND MISSING DATA

We address the question of identifiability, i.e., whether there exists a unique set of model parameters giving rise to a given distribution of observable data. The multicast inference method exploits correlations between end-to-end measurements on intersection paths. Conversely, we expect that if the sets of receivers on which data from a given probe is observable are insufficiently rich, it will not be possible to infer the loss rates on all links. We give below a simple example that demonstrates this. In this section we shall derive conditions—between the topology and the subsets at which data is observable—that must be satisfied in order that the model parameters can be identified.

Consider a parameterized family of distribution $\{P_\phi : \phi \in \Phi\}$ with vector parameter ϕ , and let F be some function on Φ . We say that P_ϕ identifies $F(\phi)$ when $P_\phi = P_{\phi'}$ implies $F(\phi) = F(\phi')$. Here, F will be the

```

1. procedure main( $\mathcal{T}, \alpha, x^*$ )
2. set_ycb( $\mathcal{T}, \alpha, x^*, \rho$ );
3. set_g( $\mathcal{T}, \alpha, 1, \rho, \rho, 1$ );
4. return( $\{g_k : k \in V\}$ );

1. procedure set_ycb( $\mathcal{T}, \alpha, x^*, k$ )
2. if( $d(k) == \emptyset$ ) then {
3.    $y_k^* := x_k^*$ ;
4.    $b_k := \bar{\alpha}_k$ ;
5.   if( $x_k^* == u$ ) then  $\{c_k := 1;\}$ 
6.   else $\{c_k := \bar{\alpha}_k;\}$ 
7. }
8. else{
9.   foreach( $j \in d(k)$ ) {
10.    ( $y_j^*, b_j, c_j$ ) := set_ycb( $\mathcal{T}, \alpha, x^*, j$ );
11.   }
12.    $y_k^* := \bigvee_j y_j^*$ ;
13.    $b_k := \bar{\alpha}_k + \alpha_k \prod_{j \in d(k)} b_j$ ;
14.    $c_k := \bar{\alpha}_k + \alpha_k \prod_{j \in d(k)} c_j$ ;
15. }
16. return( $y_k^*, b_k, c_k$ );

1. procedure set_g( $\mathcal{T}, \alpha, e, k, h, \text{flag}$ )
2.
3. if( $y_k^* == 1$ ) then {
4.    $g_k := 1$ ;
5.   foreach( $j \in d(k)$ ) {
6.     set_g( $\mathcal{T}, \alpha, 1, j, \rho, 1$ );
7.   }
8. }
9.
10. else{
11.   if(flag == 1) then  $\{h_k := k;\}$ 
12.   else $\{h_k := h;\}$ 
13.   }
14.    $g_k := e(c_k - b_k)/c_{h_k}$ ;
15.   foreach( $j \in d(k)$ ) {
16.     set_g( $\mathcal{T}, \alpha, e\alpha_k \prod_{i \in d(k) \setminus \{j\}} c_i, j, h_k, 0$ );
17.   }
18. }

```

Fig. 1. Algorithms for determining $\hat{\gamma}_{k,\alpha}(x^*)$, as returned from the procedure main.

identity, or some other projection of components of ϕ . In an MAR model, $P_{\alpha,\theta}$ identifies (α, θ) iff

$$q_{\alpha,\theta}(x^*) = q_{\alpha',\theta'}(x^*), \forall x^* \in \Omega^* \Rightarrow (\alpha, \theta) = (\alpha', \theta'). \quad (30)$$

A simple example of MCAR data that is not identifiable is a two leaf tree in which, for each probe independently, data is missing from exactly one leaf. Then the only non-trivial equations (14) become

$$q_{\alpha,\theta}(1u) = \alpha_1\alpha_2\theta(1u), \quad q_{\alpha,\theta}(0u) = (1 - \alpha_1\alpha_2)\theta(0u), \quad q_{\alpha,\theta}(u1) = \alpha_1\alpha_3\theta(u1), \quad q_{\alpha,\theta}(u0) = (1 - \alpha_1\alpha_3)\theta(u0)$$

The RHS of these equations are invariant w.r.t. the transformations $\alpha_1 \mapsto k\alpha_1, \alpha_2 \mapsto \alpha_2/k, \alpha_3 \mapsto \alpha_3/k$.

With each $S \subset R$ we associate the minimal logical multicast tree $\mathcal{T}_S = (V_S, L_S)$ that spans the root ρ and S . This is obtained by first finding the minimum spanning tree of ρ and S in \mathcal{T} . The branch points in the

spanning tree, together with ρ and S , form the node set V_S . To define L_S , the parent $f_S(k)$ in \mathcal{T}_S , of each node in $U_S := V_S \setminus \{\rho\}$, is the \succ -minimal j in V_S such that $j \succ k$ in \mathcal{T} . A path in \mathcal{T} that connects two nodes in V_S is called an S -segment. $K_S(i) = \{j \in V : i \preceq j \prec f_S(i)\}$ is the S -segment terminating at $i \in U_S$. Given $i \in V$, $\kappa_S(i)$ denote the node in V_S that terminates the S -segment containing i , i.e., that for which $i \in K_S(\kappa_S(i))$. Likewise, $\alpha_S(i) = \prod_{j \in K_S(i)} \alpha_j$ denotes the composite transmission probability along the segment $K_S(i)$. $\alpha_S = \{\alpha_S(i) : i \in U_S\}$ will denote the collection of such probabilities.

Let D_S be the $\#U_S \times \#U$ incidence matrix of the nodes of U in the segments of \mathcal{T}_S , i.e., $D_{S,jk} = 1$ if $k \in K_S(j)$ and 0 otherwise. Setting $\beta_S(k) = \log \alpha_S(k)$ and $x_k = \log \alpha_k$ we have that

$$\beta_S = D_S x \quad (31)$$

for any $S \subset R$. Before stating and proving results on identifiability, we note that there exists at least one solution, $\log \alpha$, to (31). Let $\mathbb{P}_{\alpha, \theta, S}$ denote the distribution of the reports from nodes in S . We give two conditions for identifiability of α .

Theorem 7: Let \mathcal{T} be a canonical loss tree and $\{S_i\}_{i=1}^m$ a collection of subsets of R .

(i) $U = \cup_{i=1}^m U_{S_i}$ if and only if the equations $\{\beta_{S_i} = D_{S_i} x\}_{i=1}^m$ have a unique solution x .

(ii) Assume $\mathbb{P}_{\alpha, \theta, S_i}$ identifies α_{S_i} for each i . Then $\{\mathbb{P}_{\alpha, \theta, S_i}\}_{i=1}^m$ identifies α iff either (and hence both) of the conditions of part (i) are satisfied.

Uniqueness of the solution to (31) is determined by the structure of the D_S , which depend only on the topology and the choice of the S , not on β_S . Consequently, when uniqueness holds, it holds for any additive metric. One can devise a test for identifiability based on path length in terms numbers of links. And if α is not identifiable, the procedure can be modified to determine which links can be solved.

We say that *complete data is available* from a subset S if $\theta(x^*) > 0$ for all x^* such that $R(\rho, x^*) = S$, i.e., for which reports are received from all receivers in S and no others. Let \mathcal{S}_θ denote the set of subsets S of R for which complete data is available, and Θ_c the set of missingness parameters θ with $U = \cup_{S \in \mathcal{S}_\theta} U_S$.

Theorem 8: Restrict the parameter space to $\mathcal{A} \times \Theta_c$ and assume data is MCAR. Then $\mathbb{P}_{\alpha, \theta}$ identifies (α, θ) .

Although we do not have a corresponding result for general MAR models, Theorem 8 is sufficient to enable further analysis of simple models in the following sections.

VI. ASYMPTOTICS FOR LARGE NUMBERS OF PROBES

Let $\check{\alpha} = \arg \max_{\alpha} \mathcal{L}(\alpha)$ denote the incomplete data MLE arising from (17). In this section we examine the asymptotic properties of $\check{\alpha}$ as the number of probes n grows, without specific reference to the EM algorithm.

Theorem 9: Assume MCAR. The incomplete data MLE $\check{\alpha}$ is consistent, i.e., $\lim_{n \rightarrow \infty} \check{\alpha} = \alpha$ almost surely.

We now describe the asymptotic variance of $\check{\alpha}$ for large numbers of probes n in the regime of small loss probabilities $\bar{\alpha}$. We calculate the expected Fisher information matrix for the incomplete data problem, i.e., the matrix $\mathcal{I}(\alpha, \theta) = [\mathcal{I}^{ij}(\alpha, \theta)]_{i,j \in U}$, where $\mathcal{I}^{ij}(\alpha) = -\mathbb{E} \frac{\partial^2 \mathcal{L}(\alpha)}{\partial \alpha_i \partial \alpha_j}$. Under conditions that we establish below, the inverse of $\mathcal{I}(\alpha)$, suitably rescaled, is the asymptotic variance of $\check{\alpha}$.

Our approach is to decompose the Fisher information matrix as a sum over subtrees for which complete data is present at the leaves. In the original incomplete data problem for the logical multicast topology \mathcal{T} , the counts $\mathbf{n}_S = \{m(x^*) \mid R(\rho, x^*) = S\}$, for each $S \subset R$, can be considered as a set of counts of complete outcomes on \mathcal{T}_S stemming from those probes for which reports were received only from nodes in S . Thus the

incomplete data log-likelihood function can then be decomposed as follows:

$$\mathcal{L}(\alpha) = \sum_{S \subset R: S \neq \emptyset} \mathcal{L}_c(\mathcal{T}_S, \mathbf{n}_S, \alpha_S), \quad \text{where } \mathcal{L}_c(\mathcal{T}_S, \mathbf{n}_S, \alpha_S) = \sum_{x^*: R(\rho, x^*)=S} m(x^*) \log p_{\alpha_S}(x_S^*) \quad (32)$$

and $x_S^* = \{x_k^* : k \in S\}$ is the data in x^* that is observable at S . The corresponding decomposition of the expected Fisher information matrix is

$$\mathcal{I}^{ij}(\alpha) = \sum_{S \subset R: S \neq \emptyset} \sum_{k, \ell \in U_S} \mathcal{I}_S^{k\ell}(\alpha_S) \frac{\partial \alpha_S(k)}{\partial \alpha_i} \frac{\partial \alpha_S(\ell)}{\partial \alpha_j} \quad (33)$$

where $\mathcal{I}_S^{jk}(\alpha_S) = -\mathbb{E} \frac{\partial^2 \mathcal{L}_c(\mathcal{T}_S, \mathbf{n}_S, \alpha_S)}{\partial \alpha_S(j) \partial \alpha_S(k)}$. Let $N_S = n \mathbb{P}[R(T^{(i)}(X^{(i)})) = S]$ be the mean number of probes with data observable exactly at S , and $W_S(i) = \sum_{j \in K_S(\kappa_S(i))} \bar{\alpha}_j$ the sum of link loss rates on the S -segment containing i .

Theorem 10: (i) When $\theta \in \Theta_c$ and hence when $\check{\alpha}$ is consistent, $\sqrt{n}(\check{\alpha} - \alpha)$ converges in distribution to a mean zero multivariate Gaussian random variable with covariance matrix $n^{-1} \mathcal{I}(\alpha)$.

(ii) When data is MCAR, $\mathcal{I}^{ij} = \sum_{S \subset R: S \neq \emptyset} \frac{N_S}{W_S(i)} \delta_{\kappa_S(i)\kappa_S(j)} + O(1)$ as $\bar{\alpha} \rightarrow 0$.

Example: uniform report transmission. Let reports be transmitted independently with uniform probability $p \in (0, 1]$. Then $N_S = np^{\#S} \bar{p}^{\#R - \#S}$. For each $S \subset R$, and node $\ell \in U_S$, let $C_S(\ell)$ denote the matrix on U with entries $C_S^{ij}(\ell) = 1/W_S(i)$ if $i, j \in K_S(\ell)$ and 0 otherwise. For $s \in \{1, \dots, \#R\}$ let $C_s = \sum_{S: \#S=s} \sum_{\ell \in U_S} C_S(\ell)$. Then

$$\mathcal{I} = nC \cdot (1 + O(\bar{\alpha})), \quad \text{where } C = \sum_{s=1}^{\#R} p^s (1-p)^{\#R-s} C_s \quad (34)$$

Let P_K denote the orthogonal projection onto the nullspace of a symmetric matrix K , and recursively define matrices $K_1, \dots, K_{\#R}$ by $K_1 = C_1$, and

$$K_s = P_{K_1 + \dots + K_{s-1}} C_s P_{K_1 + \dots + K_s}, \quad K_1 = C_1. \quad (35)$$

Let r_0 denote the minimal s for which $P_{K_1 + \dots + K_s} = 1$. Since $C_{\#R} = 1$, such a $r_0 \leq \#R$ exists.

Proposition 1: $p^{r_0} C^{-1}$ converges to the pseudo-inverse of K_{r_0} as $p \rightarrow 0$.

Let I_2 denote the Fisher information arising from measurements on binary subsets, i.e., I_2 is the sum obtained by restricting (33) to binary subsets S .

Proposition 2: (i) $\mathcal{I} \geq \tilde{\mathcal{I}}_2 > 0$, and hence $0 < \mathcal{I}^{-1} \leq \tilde{\mathcal{I}}_2^{-1}$, in the order of positive linear operators.

(ii) Proposition 1 holds with $r_0 = 2$.

Thus we have established:

Proposition 3: Assume independent report loss with uniform probability p . Then $\sqrt{n}(\check{\alpha} - \alpha)$ converges to multivariate Gaussian random variable with mean zero and covariance $G(\alpha, p)$, where $\lim_{p \rightarrow 0} p^2 G(\alpha, p) = KI_2 + O(\|\bar{\alpha}\|^2)$ as $\bar{\alpha} \rightarrow 0$, where KI_2 is the pseudo-inverse of K_2 .

Proposition 3 suggests that we approximate the variance of $\check{\alpha}_k$ by $(KI_2)_{kk}/(np^2)$ when p and $\bar{\alpha}$ are small, and n is large.

Example: uniform report transmission from binary trees Consider the family of binary trees with 2^r leaves, $r = 1, 2, \dots$, with small uniform link loss probabilities $\bar{\alpha}$ and uniform small report transmission rate p . Let

$$\begin{aligned}
v(1) &= \left\{ \frac{1}{3}, \frac{1}{3} \right\} \\
v(2) &= \left\{ \frac{1}{8}, \frac{21}{40}, \frac{2}{5} \right\} \\
v(3) &= \left\{ \frac{3}{80}, \frac{49}{240}, \frac{25}{42}, \frac{3}{7} \right\} \\
v(4) &= \left\{ \frac{1}{96}, \frac{43}{672}, \frac{27}{112}, \frac{91}{144}, \frac{4}{9} \right\} \\
v(5) &= \left\{ \frac{5}{1792}, \frac{33}{1792}, \frac{5}{64}, \frac{21}{80}, \frac{36}{55}, \frac{5}{11} \right\} \\
v(6) &= \left\{ \frac{3}{4096}, \frac{187}{36864}, \frac{133}{5760}, \frac{153}{1760}, \frac{73}{264}, \frac{209}{312}, \frac{6}{13} \right\}
\end{aligned}$$

TABLE II
FIRST SIX SETS OF COEFFICIENTS $v(r)$

$v(r) = (v_1(r), \dots, v_{r+1}(r))$ denote the set of unique diagonal elements of $KI_2/\bar{\alpha}$, the j^{th} element determining the asymptotic variance on links j nodes away from the root. The first six $v(r)$ were computed using Mathematica, and are reported in Table II. In all cases the estimator variance rises in a given tree on moving away from the root, except falling slightly at a leaf link as compared with its parent. At a given distance from the root, the link variance decreases as the tree depth increases. Both these trends can be understood by the intuition that variance should decrease when data is available from larger subtrees below a given link of interest. Considering the root and leaf links only, the values in Table II are consistent with the forms

$$v_1(r) = \frac{r}{r+2} \frac{1}{2^{2(r-1)}}, \quad v_{r+1}(r) = \frac{r}{2r+1} \quad (36)$$

VII. CONVERGENCE RATES FOR THE EM ITERATES

We now consider convergence of the EM iterates themselves. Let \mathcal{M} denote the map on \mathbb{R}_+^U that implements the iteration, i.e., such that $\hat{\alpha}^{(\ell+1)} = \mathcal{M}(\hat{\alpha}^{(\ell)})$. A Taylor expansion of the iterative map gives

$$\hat{\alpha}^{(\ell+1)} - \alpha \approx \nabla \mathcal{M} \cdot (\hat{\alpha}^{(\ell)} - \alpha) \quad (37)$$

where $\nabla M_{ij} = \frac{\partial M_i}{\partial \alpha_j}$ is the gradient of \mathcal{M} . A standard result [17, §3.93] expresses $\nabla \mathcal{M} = (1 - \mathcal{I}_c^{-1} \mathcal{I})$, with \mathcal{I}_c the complete data information matrix and \mathcal{I} the incomplete data information matrix from (33). The convergence ratio of the iteration is taken as the maximum eigenvalue λ for $\nabla \mathcal{M}$.

Our analysis of the convergence ratio is confined to the regime treated in Section VI, namely that of independent report transmission with small probability p , and small link loss probabilities $\bar{\alpha}$. In this regime, we have seen that $(\mathcal{I}_c^{-1})_{ij} = n^{-1}(\bar{\alpha}_i \delta_{ij} + O(\|\bar{\alpha}\|^2))$ and so from (34) $\nabla \mathcal{M}(\alpha)_{ij} = \delta_{ij} - \bar{\alpha}_i C_{ij} + O(\|\bar{\alpha}\|^2)$. Let $\mathcal{E}(X)$ denote the set of eigenvalues of a matrix X .

Proposition 4: Assume independent report loss with probability $p \in (0, 1)$ and small uniform probe loss rate $\bar{\alpha}$. The convergence rate λ for the EM algorithm obeys $\lambda = 1 - p^2 \kappa + O(p^2(p + \bar{\alpha}))$, where κ is the minimum non-zero eigenvalue of $\bar{\alpha} K_2$.

VIII. EXPERIMENTS AND SIMULATIONS

In order to evaluate the performance of the missing data inference algorithm, we conducted two types of simulation. First, we used simulation we used model-based simulation in which the model for missing data

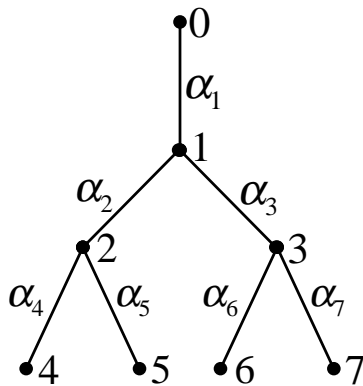


Fig. 2. 4-RECEIVER BINARY TREE: used in model based simulation of Section VIII-A.

indicators conformed with the assumed MAR property. Second, we used a network-based implementation of the RTCP-based reporting mechanism outlined in Section II-F. In this case the missing data indicators are not known to conform to the MAR model. This enabled us to test robustness of the algorithm with respect to violations of the MAR hypothesis that might occur in a real network application.

A. Model-based simulation

We conducted model based simulations on a balanced binary tree with 4 receivers, illustrated in Figure 2. Probe losses were independent with a uniform loss rate per link. Receiver reports were generated at each receiver for each probe and were subject to independent loss with a uniform probability. We conducted 100 separate cumulation runs, each of 100,000 probes. Initialization of $\hat{\alpha}^{(0)}$ used (18). The termination criterion for the EM algorithm was that successive iterates $\hat{\alpha}_k^{(\ell)}$ should differ by less than 10^{-4} on each link k .

Figure 3(left) shows the mean and error bars for 95% confidence of link loss estimates obtain using the missing data algorithm. We also display the corresponding quantities for the complete data estimator applied to only those probes for which complete reports were available. In both cases the mean estimate is close to the model loss rate, i.e. $\bar{\alpha}_k = 0.01$. But note the rapid widening of error bars for the full data algorithm, compared with the missing data algorithm, as the report transmission probability decreases. From Prop. 3 we expect the standard error to diverge as p^{-1} for the missing data algorithm, regardless of the topology. However, in a 4-leaf tree the number of probes with complete data is proportional to p^4 . Hence we expect the standard error to diverge as p^{-2} , with faster divergence for trees with more receivers. In this example, for p less than 0.4, the error bars encompass loss rate 0: the inferred loss from complete becomes statistically indistinguishable from 0.

Figure 3(right) breaks down the standard error according to the location of the link in the topology links 1, 2 and 4 being representative of links respectively 0, 1 and 2 links removed from the root. The experiment standard error show close agreement with the theoretical values obtained by inverting the information matrix \mathcal{I} in (34). We also show the small p approximation obtained using Proposition 3 and the values $v(2)$ from Table II. The approximation remains reasonable even for quite large p .

B. RTCP-based experiments

The RTCP-based simulations used data gathered from a network-based implementation of loss reporting. Loss reports are embedded in RTCP feedback packets; any collector listening to these can then perform

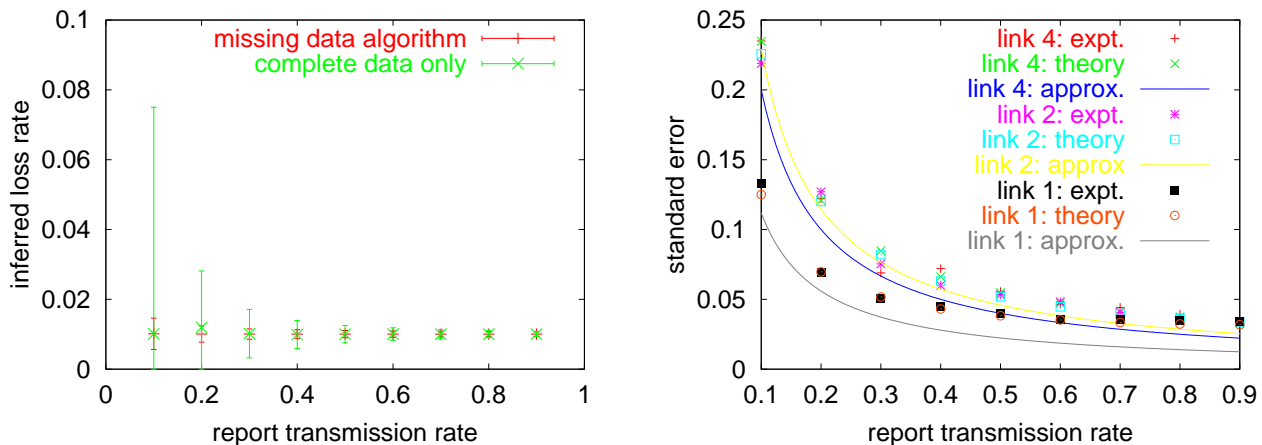


Fig. 3. VARIANCE OF MISSING DATA ESTIMATOR IN 4-RECEIVER UNIFORM PROBE AND PACKET LOSS MODEL: Over 100 simulation runs each of 100,000 probes, uniform link loss $\bar{\alpha} = 0.01$, probe transmission rate p from 0.1 to 0.9. LEFT: mean estimate with error bars for 95% confidence. Comparison with estimator using only probes with complete data. RIGHT: standard error depending on link location: experiment, theory and approximation.

inference. The basic RTCP reporting mechanism includes only the average loss rate based on sequence numbers of received packets. An extension of the report format allows the inclusion of a binary vector indicating receipt or otherwise of a set of packets.

According to the RTP standard [20], the total report volume over all receivers should not exceed 5% of the source rate. RTCP clients estimate their share of this based upon the reports they hear from the other receivers, and limit report frequency and size accordingly. Consequently, for a sufficiently large number of receivers, it will not be possible to report on all probes. Missingness arises then by two mechanisms: the omission of certain probes from reporting, and the loss of report packets during transmission to the collector.

The implementation of extended RTCP-based reporting used in this study has a simulation mode that enables it to report on packet losses generated on a model topology according to a Bernoulli loss model, rather than due to packet loss in a real network. The probe source was chosen to have the characteristics of a GSM audio stream that could act as a probe source in real networks, sending packets at a rate of 50 per second. Since probe losses follow the assumed statistical model, only the missing data indicators can potentially exhibit departures from our model assumptions. Report thinning and transmission then takes place in the manner described above.

We collected traces from a 32 receiver balanced binary tree for which the link loss rates were chosen independently with a uniform distribution between 1% and 10%. The trace comprised reports on 11,956 probe packets, encompassing about 4 minutes at 50 packets per second. The mean number of reports received for a given probe was 18.8, so that the proportion of missing reports was $1 - 18.8/32 = 41.3\%$. The maximum number of reports per probe was 29, i.e. no probe had complete data. Figure 4(left) shows a scatter plot of the 63 pairs of (actual, inferred) loss rates. The agreement is quiet close, with tight clustering around the line of slope 1 through the origin. The median relative error over all links was only 4.5%.

Figure 4(right) displays the median, 5th and 95th percentile of the relative error over all links as a function of the size of a subsets of probes used for inference. Note that even with 2000 probes the relative error typically less than 50%. Hence we can expect to identify the lossiest links with measurements over a duration less than 1 minute. The median error is only about 13% for this number of probes.

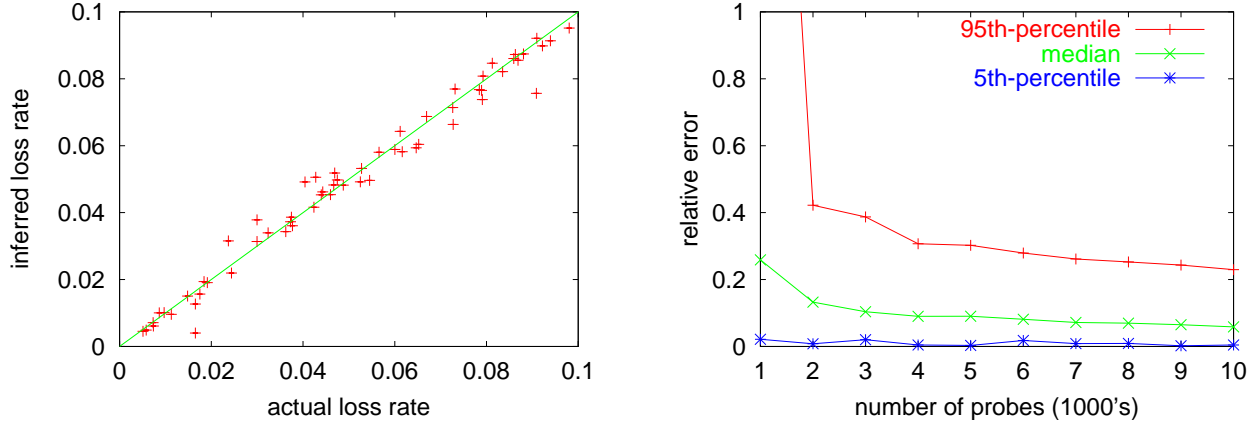


Fig. 4. INFERENCE FROM RTCP TRACE DATA ON 32 RECEIVER BINARY TREE. LEFT: Scatter plot of inferred vs. actual loss rates for full trace of 11,956 probes. RIGHT: Median, 5th and 95th percentiles of relative error over all links as function of number of probes.

IX. CONCLUSIONS

In this paper we have extended the multicast based method for inferring network internal loss from end-to-end measurements that was first proposed in [3]. The extension was motivated by the proposed use of the RTCP transport protocol to transmit measurements. This will inevitably lead to missing data, either through the need to thin data, or due to loss of reports in transmission. We applied the EM algorithm to generate an iterative approximation to the corresponding MLE, and analyzed its convergence rate for a class of models. We evaluated the approach on data derived from model simulations, and also on traces gathered from an implementation of the RTCP-based report transfer method. These results showed (i) variance reduction as compared with an ad hoc inference approach; (ii) accuracy of inferred loss rates compared with model or directly measured rates in the simulation; and (iii) robustness of the approach under potential departures from the model assumptions on the missingness statistics in the RTCP-based application. In the RTCP-based experiments the median estimation error on a 32 receiver tree was only about 13% for 2,000 probes, and was typically less than about 50%. Thus inference sufficiently accurate to identify the lossiest links could be performed on measurements collected over about a minute.

Future work is planned in two directions. First, we want to apply the same general methodology to the estimation of other internal characteristics, such as delay, utilization and topology itself, by adapting the framework of the present paper to work on estimation of these quantities with complete data, as performed in [4], [10], [16]. A second direction is to develop more specific models of the missing data mechanism that could be used in a parametric approach to estimation with missing data. Lastly, we intend to publish elsewhere details of the RTCP reporting mechanism that motivated this study; see [2].

Acknowledgement We thank Francesco Lo Presti for some useful suggestions.

X. PROOFS OF THEOREMS

Proof of Theorem 2: We first render $\mathcal{L}_c(\alpha)$ into the canonical form of a standard exponential family. Denote by $\mathbf{0}$ and $\mathbf{1}$ the elements x of Ω with x_k all 0 or all 1 respectively. For $x \in \Omega$, denote by $W'(x)$ those nodes $k \in U$ for which $x_j = 0$ for all $j \in R(k)$. Let $W(x)$ be the \succeq -maximal elements of $W'(x)$. Note that $W(\mathbf{1}) = \emptyset$. For each $k \in U$ and $i \in \{0, 1\}$ define $q_k(i) = \mathbb{P}_\alpha[X_j = i, \forall j \in R(k) | X_{f(k)} = 1]$. Define new

parameters $\{\delta_k : k \in U\}$ by $\delta_k = \log(q_k(0)/q_k(1))$. Observe that

$$\frac{p_\alpha(x)}{p_\alpha(\mathbf{1})} = \prod_{k \in W(x)} \frac{q_k(0)}{q_k(1)} = \prod_{k \in W(x)} e^{\delta_k}. \quad (38)$$

We interpret the product over \emptyset for $x = \mathbf{1}$ as 1. The map taking \mathcal{A} to its image Δ under the change of parameters $\alpha \mapsto \delta$ is invertible. To see this note that given $\sum_{x \in \Omega} p_\alpha(x) = 1$, (38) fixes the $p_\alpha(x)$ in terms of the δ . These in turn determine the γ_k , and hence the α_k , by Theorem 2.

Writing $n(\mathbf{1}) = n - \sum_{x \neq \mathbf{1}} n(x)$, and recalling that $W(\mathbf{1}) = \emptyset$, we find

$$\mathcal{L}_c(\alpha) = \sum_{x \in \Omega} n(x) \sum_{k \in W(x)} \delta_k + n \log p_\alpha(\mathbf{1}) = \sum_{k \in U} N_k \delta_k - nc(\delta),$$

where $N_k = \sum_{x \in \Omega: k \in W(x)} n(x)$, and $c(\delta)$ is the reparameterization of $-\log p_\alpha(\mathbf{1})$ in terms of δ :

$$c(\delta) = \log \sum_{x \in \Omega} \prod_{k \in W(x)} e^{\delta_k}. \quad (39)$$

The expression (39) has the form of a standard exponential family, with the log-likelihood expressed in terms of the natural parameters $\delta = \{\delta_k : k \in U\}$, sufficient statistics $N = \{N_k : k \in U\}$, and cumulant $c(\delta)$. Since $c(\delta)$ is finite for all $\delta \in \mathbb{R}^U$ the family can be considered full. However, the parameter space of interest is the open subset $\Delta \subset \mathbb{R}^U$ that is the image of \mathcal{A} under the reparameterization $\alpha \mapsto \delta$.

Since the mapping $\mathcal{A} \rightarrow \Delta : \alpha \mapsto \delta$ is invertible, the parameters δ are identifiable by Theorem 2(v), and hence the exponential family is affinely independent. (A simple argument shows that natural parameters in an open set are identifiable iff the exponential family is affinely independent). A well-known result (see e.g. [14, Ex. 6.6.3.]) for standard exponential families then says that the MLE is the solution δ' of

$$N_k = \mathbb{E}_{\delta'}[N_k], \quad k \in U \quad (40)$$

provided this δ lies in the interior of Δ . But clearly $\sum_{j \geq k} N_j = n(1 - \hat{\gamma}_k)$, and hence finding the solution to (40) is equivalent to finding the solution α' to

$$\gamma_k = \mathbb{E}_{\alpha'}[\hat{\gamma}_k], \quad k \in U. \quad (41)$$

Provided $\hat{\gamma} \in \mathcal{G}$, then by Theorem 1, $\hat{\alpha}$ is the unique such solution, and hence if it lies in \mathcal{A} it is the MLE. ■

Proof of Theorem 3: Observe that $\mathbb{E}_\alpha[\mathcal{L}_c(\alpha') | \mathbf{m}] = \sum_{x \in \Omega} \mathbb{E}_\alpha[n(x) | \mathbf{m}] \log p_{\alpha'}(x)$. Hence maximizing $Q_{\hat{\alpha}^{(\ell)}, \alpha'}$ over α' is equivalent to finding the complete-data MLE, but with $n(x)$ replaced by $\mathbb{E}_{\hat{\alpha}^{(\ell)}}[n(x) | \mathbf{m}]$ throughout. In particular, $\hat{\gamma}_k$, being a linear combination of the $n(x)$, gets replaced by $\hat{\gamma}_k^{(\ell)}$. Now if $\hat{\alpha}^{(\ell)} \in \mathcal{A}$ then it is not hard to see that $\hat{\gamma}_{\hat{\alpha}^{(\ell)}}(x^*) \in \mathcal{G}$ for each x^* , and hence $\hat{\gamma}^{(\ell)} \in \mathcal{G}$ since \mathcal{G} is convex. The result then follows from Theorems 1 and 2. ■

Proof of Theorem 4: The condition says that the sequence of EM iterates exists in \mathcal{A} . Parts (i) and (ii) follow from Theorem 6 of [22]. This is because: (a) by Theorem 1(iii), $\hat{\alpha}^{(\ell+1)}$ is stationary for $\alpha \mapsto Q(\alpha, \hat{\alpha}^{(\ell)})$; (b) $\nabla_\alpha Q(\alpha', \alpha)$ is clearly continuous for $(\alpha', \alpha) \in \mathcal{A} \times \mathcal{A}$; and (c) By Theorem 2, \mathcal{L}_c comes from a regular exponential family and hence $\|\hat{\alpha}^{(\ell+1)} - \hat{\alpha}^{(\ell)}\| \rightarrow 0$ as $\ell \rightarrow \infty$; see remark 3(vi) in [22]. Part (iii) then follows from Corollary 1 in [22] ■

Proof of Theorem 5: For the proof we observe the following Markov property of the probe process X :

(M) Conditioned on $X_i = 1$, the distributions of the sets of variables $\{X_k : k \in R(i)\}$ are independent for different $i \in d(k)$.

If $y_k^* = 1$, then $\bigvee_{j \in R(k)} x_j = 1$, and hence $\widehat{\gamma}_{k,\alpha}(x^*) = 1$. Suppose instead that $y_k^* < 1$. Let $Q(k)$ denote the event that $X_j = 0$ for all $j \in R(k, x^*)$ if $R(k, x^*) \neq \emptyset$, otherwise take $Q(k)$ as the universal set in the underlying probability space. Since $y_h^* = 1$ for $h = h(k, x^*)$, then $\{X^* = x^*\} \subset \{X_h = 1\}$ and hence $\{X^* = x^*\} = \{X_h = 1\} \cap Q(d(h, k)) \cap \{X_j^* = x_j^* : j \in R \setminus R(d(h, k))\}$. This yields that

$$\mathbb{P}_\alpha[\bigvee_{j \in R(k)} X_j = 1 \mid X^* = x^*] = \frac{\mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(d(k, h)) \mid X_h = 1]}{\mathbb{P}_\alpha[Q(d(k, h)) \mid X_h = 1]}, \quad (42)$$

where we have used property (M), and the fact that $\mathbb{P}[A \mid B \cap C \cap D] = \mathbb{P}[A \mid C \mid B]$ when A and C are conditionally independent of D given B .

The denominator in (42) is just $c_{d(h,k)}$. To treat the numerator, observe that

$$\begin{aligned} & \mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(i) \mid X_{f(i)} = 1] \\ &= \mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(d(k, i)) \cap \{X_i = 1\} \cap_{j \in d(i) \setminus d(k, i)} Q(j) \mid X_{f(i)} = 1] \\ &= \mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(d(k, i)) \mid X_i = 1 \mid X_{f(i)} = 1] \mathbb{P}_\alpha[X_i = 1 \mid X_{f(i)} = 1] \prod_{j \in d(i) \setminus d(k, i)} \mathbb{P}_\alpha[Q(j) \mid X_i \mid X_{f(i)}] \\ &= \mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(d(k, i)) \mid X_i = 1] \alpha_i \prod_{j \in d(i) \setminus d(k, i)} c_j. \end{aligned} \quad (43)$$

Here we have used $Q(j) = \bigcap_{i \in d(j)} Q(i)$, the Markov property (M), and the fact that $\{X_k = 1\} \subset \{X_{f(k)} = 1\}$. Applying (43) repeatedly to the numerator of (42), we obtain the form $\mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(k) \mid X_{f(k)} = 1] \prod_{k \prec i \preceq d(k, h)} \alpha_i \prod_{j \in d(i) \setminus \{d(i, k)\}} c_j$. The first term in this product is $\mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \cap Q(k) \mid X_{f(k)} = 1] = c_k \mathbb{P}_\alpha[\{\bigvee_{j \in R(k)} X_j = 1\} \mid Q(k) \mid X_{f(k)} = 1] = c_k(1 - \mathbb{P}_\alpha[\bigvee_{j \in R(k)} X_j = 0 \mid X_{f(k)} = 1]) / c_k = c_k - b_k$, and the stated result (23) follows. ■

Proof of Lemma 1: We first show that for each $k \in V$ there is some $x^* \in \Omega_0^*$ with $m(x^*) > 0$ for which $\widehat{\gamma}_{k,\alpha}(x^*) > 0$. By condition (25) there is x^* with $m(x^*) > 0$, for which either $x_j^* = 1$ or $x_j^* = u$ for some $j \in R(k)$. In the former case $\widehat{\gamma}_{k,\alpha}(x^*) = 1 -$ or $x_j^* = u$; in the latter case it is not hard to see in Theorem 5 that $c_k > b_k$ and hence $\widehat{\gamma}_{k,\alpha}(x^*) > 0$. Since $\widehat{\gamma}_{k,\alpha}(x^*) \geq 0$ for all $x^* \in \Omega_0^*$, then $\widehat{\gamma}_{k,\alpha} > 0$.

By (26) then for each $k \in V \setminus R$ there exists x^* with $m(x^*) > 0$ such that $\widehat{\gamma}_{k,\alpha}(x^*) = 1$, and children j, ℓ of k for which $\widehat{\gamma}_{j,\alpha}(x^*) = 1$ while $\widehat{\gamma}_{\ell,\alpha}(x^*) > 0$. Hence $\widehat{\gamma}_{k,\alpha}(x^*) < \sum_{j \in d(k)} \widehat{\gamma}_{j,\alpha}(x^*)$. Since by definition $\widehat{\gamma}_{k,\alpha}(x^*) \leq \sum_{j \in d(k)} \widehat{\gamma}_{j,\alpha}(x^*)$ for all x^* , we have $\widehat{\gamma}_{k,\alpha} < \sum_{j \in d(k)} \widehat{\gamma}_{j,\alpha}$. Taking these relations over all relevant k , $\widehat{\gamma}_\alpha \in \mathcal{G}$. ■

Proof of Theorem 6: Reparameterizing the incomplete data likelihood function \mathcal{L} in terms of γ , we obtain

$$\begin{aligned} \mathcal{L} &= n(11) \log(\gamma_2 + \gamma_3 - \gamma_1) + n(10) \log(\gamma_1 - \gamma_3) + n(01) \log(\gamma_1 - \gamma_2) + n(00) \log(1 - \gamma_1) \\ &\quad + n(1u) \log(\gamma_2) + n(u1) \log(\gamma_3) + n(0u) \log(1 - \gamma_2) + n(u0) \log(1 - \gamma_3) \end{aligned}$$

Writing this as $\mathcal{L} = \sum_{x^* \in \Omega_0^*} n(x^*) \mathcal{L}_{x^*}$, it suffices to show that $-\mathcal{L}_{x^*}$ is jointly convex in $\gamma_1, \gamma_2, \gamma_3$ for each x^* . This follows from the fact, established by direct computation, that the principal minors of the

Hessian matrices $-\partial^2 \mathcal{L}_{x^*} / \partial \gamma_i \partial \gamma_j$ for $i, j = 1, 2, 3$ are non-negative. Convergence then follows from Theorem 4 and the standing assumptions (25) and (26). ■

The method used in the proof appears not to extend to more general trees, not even binary ones.

Proof of Theorem 7: (i) We establish that if $U = \cup_{i=1}^m U_{S_i}$, then there is a unique solution x . We do this by contradiction. Suppose that there exists a node $k \in V$ such that there does not exist a unique value for x_k . Pick any such \prec -maximal k . By assumption, k terminates some segment $K_{S_i}(k)$ and hence $x_k = \beta_{S_i}(k) - \sum_{k \prec j \prec f_{S_i}(k)} x_j$. By the maximality assumption, all terms on the RHS are unique, and hence so is x_k .

We establish now that if there is a unique solution x , then $U = \cup_{i=1}^m U_{S_i}$. This is done by contradiction. First, note if the solution x is unique, it must be $x_k = \log \alpha_k$. Second, note that we only need to consider a branch point $k \in U \setminus R$. Assume there exists $k \notin \cup_{i=1}^m U_{S_i} \setminus R$. Then every segment containing $v \in d(k)$ also contains k . Using $x_k = \log \alpha_k$ we can reduce the family of equations $\{\beta_{S_i} = D_{S_i} x\}_{i=1}^m$ to the following set of equations describing the behavior of x_k and x_v , $v \in d(k)$,

$$x_k + x_v = \beta'_v, \quad \forall v \in d(k) \quad (44)$$

where β'_v is the log of the probability of a reception of a packet at $v \in d(k)$ given that it was received at $f(k)$. The β'_v are determined by the link probabilities other than α_k and α_v , $v \in d(k)$, and the original β_{S_i} . However, these equations do not have a unique solution, resulting in a contradiction.

(ii) The solution x to $\{\beta_{S_i} = D_{S_i} x\}_{i=1}^m$ is not unique if and only if there is more than one set of link probabilities α giving rise to the same $\{\alpha_{S_i}\}_{i=1}^m$ and hence the same $\{P_{\alpha, \theta, S_i}\}_{i=1}^m$. ■

Proof of Theorem 8: With MCAR the missingness probabilities do not depend on any data, and so we can replace $\theta(x^*)$ with $\theta(t(x^*))$. Since $\sum_{x^*: t(x^*)=t'} p_{\alpha}^*(x^*) = 1$ for all $t' \in \{0, 1\}^R$, $q_{\alpha, \theta}(x^*) = q_{\alpha', \theta'}(x^*)$, $\forall x^* \in \Omega^*$ implies $\theta = \theta'$.

Consider now any $S \subset R$ for which complete data is available. Then $\theta(t(x^*))$ are equal and hence strictly positive for any x^* with $R(\rho, x^*) = S$. Hence $q_{\alpha, \theta}(x^*) = q_{\alpha', \theta'}(x^*)$ implies $p_{\alpha}(x^*) = p_{\alpha'}(x^*)$. By Theorem 1(iv), α_S is identifiable. Then α is identifiable by Theorem 7(ii) since $\theta \in \Theta_c$. ■

Proof of Theorem 9: The proof of mirrors that of Theorem 3(ii) in [3] which in turn uses Lemma 7.54 in [19]. Although not mentioned there, the latter result requires identifiability. This follows from the MCAR assumption and Theorem 8. ■

Proof of Theorem 10: We first show that \mathcal{I}_S is positive definite if complete data is available from S . By standard arguments (see e.g. Prop 2.84 in [19]) one writes $\mathcal{I}_S^{jk}(\alpha_S) = \text{Cov}_{\alpha, \theta} \left(\frac{\partial \mathcal{L}_c(\mathcal{I}_S, \mathbf{n}_S, \alpha_S)}{\partial \alpha_S(j)}, \frac{\partial \mathcal{L}_c(\mathcal{I}_S, \mathbf{n}_S, \alpha_S)}{\partial \alpha_S(k)} \right)$. If $\mathcal{I}_S^{jk}(\alpha_S)$ is not positive definite, there exists some nonzero $c \in \mathbb{R}^{U_S}$ for which $c \cdot \mathcal{I}_S(\alpha_S) \cdot c = \text{Var}_{\alpha, \theta} \left(c \cdot \frac{\partial \mathcal{L}_c(\mathcal{I}_S, \mathbf{n}_S, \alpha_S)}{\partial \alpha_S} \right) = 0$. This happens if $c \cdot \frac{\partial \mathcal{L}_c(\mathcal{I}_S, \mathbf{n}_S, \alpha_S)}{\partial \alpha_S} = 0$, almost surely, or, equivalently, if $c \cdot \frac{\partial \log p_{\alpha_S}(x_S^*)}{\partial \alpha_S} = 0$ for all x_S^* , since $\theta(x^*) > 0$ by assumption. Repeating the argument of Theorem 4(ii) in [3], this requires $c = 0$.

Observe that $\frac{\partial \alpha_S}{\partial \alpha} = B(\alpha_S) D_S B^{-1}(\alpha)$ where $B(\alpha)$ is the diagonal matrix with entries from α . Thus $\mathcal{I} \geq \sum_{S \in \mathcal{S}_\theta} B^{-1}(\alpha) D_S^T B(\alpha_S) \mathcal{I}_S(\alpha_S) B(\alpha_S) D_S B^{-1}(\alpha)$ (as positive linear operators). Since the kernel of a sum of non-negative definite operators is the intersection of their kernels, the sum is positive definite iff $\cap_{S \in \mathcal{S}_\theta} \ker(D_S) = \{0\}$, i.e., iff the equations $\{\beta_S = D_S x\}_{S \in \mathcal{S}_\theta}$ have a unique solution x , which is guaranteed by Theorem 7(i) and the assumption $\theta \in \Theta_c$.

(ii) Let $S \in \mathcal{S}_\theta$. Note $\alpha_S(k)/\alpha_i = D_{S,ki}(1 + O(\|\bar{\alpha}\|))$. With data MCAR, the $\theta(x^*)$ are equal for any x^* with $R(\rho, x^*) = S$, and hence $\text{Em}(x^*) = N_S p_\alpha^*(x^*) = N_S p_{\alpha_S}(x_S^*)$. From Theorem 5 of [3], $N_S^{-1} \mathcal{I}_S(\alpha_S)$ has inverse ν_S for which $\nu_S^{k\ell} = \bar{\alpha}_S(k) \delta_{k\ell} + O(\|\bar{\alpha}_S\|^2)$. Hence $\mathcal{I}_S^{k\ell}(\alpha_S) = \frac{N_S}{\bar{\alpha}_S(k)} (\delta_{k\ell} + O(\|\bar{\alpha}_S\|))$. Finally, observe that $\bar{\alpha}_S(\kappa_S(i)) = W_S(i) + O(\|\bar{\alpha}\|^2)$. Putting these together with (33) the result follows. ■

Proof of Proposition 1: Rewrite $C = \sum_{s=1}^{\#R} p^s C'_s$ where $C'_s = \sum_{1 \leq t \leq s} C_t \cdot (-1)^{s-t} \binom{n-t}{s-t}$. Observe that the definition of K_s is invariant under replacement of C_s by C'_s in (35) since $P_{K_1+\dots+K_{s-1}} C_{s'} P_{K_1+\dots+K_{s-1}} = 0$ for $s' < s$. Let W_s denote the unitary matrix that diagonalizes K_s and set $W = \prod_{i=1}^{r_0} W_s$. (Since the ranges of the K_s are disjoint, the various W_s commute). Then up to unitary transformation under W , we can write C is block diagonal form

$$C = W \begin{pmatrix} pA_{11}(p) & p^2A_{12}(p) & \cdots & p^{r_0}A_{1r_0}(p) \\ p^2A_{21}(p) & p^2A_{22}(p) & \cdots & p^{r_0}A_{2r_0}(p) \\ \vdots & \vdots & \ddots & \vdots \\ p^{r_0}A_{r_01}(p) & p^{r_0}A_{r_02}(p) & \cdots & p^{r_0}A_{r_0r_0}(p) \end{pmatrix} W^T \quad (45)$$

where $A_{ij}(p) = A_{ji}^T(p)$, and each submatrix A_{ij} converges to some A_{ij} as $p \rightarrow 0$ such that the block matrix with A_{ss} as the s^{th} diagonal element and zero elsewhere is unitarily equivalent to K_s under W . By construction, the A_{ss} are invertible, and hence so are the $A_{ss}(p)$ for sufficiently small p . The block diagonal representation of C can be inverted inductively as follows. Assume the block submatrix $B_{[s-1][s-1]}$ comprising the first $s-1$ blocks can be inverted and that $B_{[s-1][s-1]}^{-1}$ is $O(p^{1-s})$ as $p \rightarrow 0$. This condition is trivially satisfied for $s=2$. Now write $B_{[s][s]}$ as a two-by-two superblock matrix

$$B_{[s][s]} = \begin{pmatrix} B_{[s-1][s-1]} & B_{[s-1]s} \\ B_s{}_{[s-1]} & p^s A_{ss}(p) \end{pmatrix} \quad (46)$$

Then $B_{[s][s]}^{-1} = D_{s,1}(p) \cdot D_{s,2}^{-1}(p)$ where, denoting $W_s = B_{[s-1][s-1]}^{-1} B_{[s-1]s} A_{ss}^{-1}(p)$ and W_s^T its transpose,

$$D_{s,1}(p) = \begin{pmatrix} B_{[s-1][s-1]}^{-1} & -p^{-s} W_s \\ -p^{-s} W_s^T & p^{-s} A_{ss}^{-1}(p) \end{pmatrix}, \quad D_{s,2}(p) = \begin{pmatrix} 1 - p^{-s} B_{[s-1]s} W_s^T & 0 \\ 0 & 1 - p^{-s} B_s{}_{[s-1]} W_s \end{pmatrix}$$

Now $B_{[s-1]s} = O(p^s)$ as $p \rightarrow 0$, from which it follows that

$$B_{[s][s]}(p) = \begin{pmatrix} 0 & 0 \\ 0 & p^{-s} A_{ss}^{-1} \end{pmatrix} + O(p^{1-s}). \quad (47)$$

Consequently, $B_{[ss]}^{-1} = O(p^{-s})$ as $p \rightarrow 0$, completing the induction step. The statement of the Proposition then follows from (47) by taking $s = r_0$. ■

Proof of Proposition 2: (i) In the given model, $N_S > 0$ for all subsets S of R , and hence $\mathcal{I} \geq \mathcal{I}_2$. Since each node in $U \subset R$ has at least two descendent leaves $U = \cup_{\#S=2} U_S$ and hence $\mathcal{I}_2 > 0$ by an argument similar to that in Theorem 7(i). Now if $A \geq B > 0$ then $0 < 1 - a \leq 1 - b < 1$ where $a = A/(2\|A\|)$ and $b = B/(2\|A\|)$. Since $\|1 - a\|, \|1 - b\| < 1$, $a^{-1} = (1 - (1 - a))^{-1} = \sum_{i=0}^{\infty} (1 - a)^i < \sum_{i=0}^{\infty} (1 - b)^i = b^{-1}$, whence the result follows.

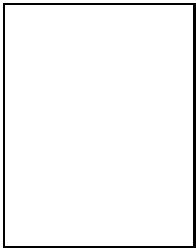
(ii) Consider inference performed by using only measurements from binary sets S . $\mathcal{I}_2 = pQ_1 + p^2Q_2$ for some Q_1 and Q_2 independent of p , and hence \mathcal{I}_2^{-1} is $O(p^{-2})$. By (i), $\mathcal{I}^{-1} \leq \mathcal{I}_2^{-1}$, which precludes $r_0 > 2$ in Proposition 1. We conclude $r_0 = 2$ by showing that K_1 has 0 as an eigenvalue, for then $r_0 > 1$. Assume that

the root ρ as a unique child 1. If not, partition the \mathcal{T} into disjoint subtrees with nodes descended from each child of ρ , then apply the following argument to each subtree. Let v denote the element of R^U with $v(1) = 1$, $v(j) = -1$ for $j \in d(1)$, and $v(k) = 0$ otherwise. Observe that for each $k \in R$, the $C_{\{k\}}(k)_{ij}$ are equal for $i, j \in \{1, d(k, 1)\}$. Since $C_1 = \sum_{k \in R} C_{\{k\}}$, then $C_1 \cdot v = 0$. ■

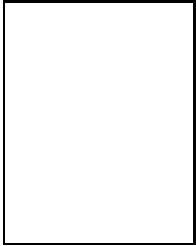
Proof of Proposition 4: $\max \mathcal{E}(\nabla \mathcal{M}) = 1 - \bar{\alpha} \min \mathcal{E}(C) + O(\bar{\alpha})$. From Propositions 1 and 2 we know that C takes the block diagonal form (45) with $r_0 = 2$. From this it follows that each eigenvalue of C takes the form $p^i v_i(p) + O(p^{i+1})$ for some $i \in \{1, 2\}$, where $\lim_{p \rightarrow 0} v_i(p) \in \mathcal{E}(A_{ii})$. Since $0 \notin \mathcal{E}(A_{ii})$, $p^{-2} \min \mathcal{E}(C) \rightarrow \min \mathcal{E}(A_{22})$ as $p \rightarrow 0$, and hence the result follows. ■

REFERENCES

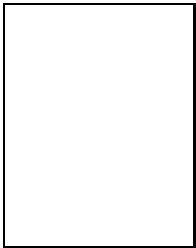
- [1] A. Adams, T. Bu, R. Cáceres, N.G. Duffield, T. Friedman, J. Horowitz, F. Lo Presti, S.B. Moon, V. Paxson, D. Towsley, The Use of End-to-End Multicast Measurements for Characterizing Internal Network Behavior, IEEE Communications Magazine, May 2000.
- [2] R. Cáceres, N.G. Duffield, T. Friedman, “Impromptu measurement infrastructures using RTP”, Proc. IEEE Infocom 2002, New York, June 23-27, 2002.
- [3] R. Cáceres, N.G. Duffield, J. Horowitz and D. Towsley, “Multicast-Based Inference of Network Internal Loss Characteristics” IEEE Trans. on Information Theory, vol. 45, pp. 2462-2480, 1999.
- [4] R. Cáceres, N.G. Duffield, J. Horowitz F. Lo Presti, D. Towsley, “Statistical Inference of Multicast Network Topology”, in Proc. IEEE Conf. on Decision and Control, Phoenix, AZ, Dec. 1999.
- [5] R. Carter, M. Crovella, “Measuring bottleneck link-speed in packet-switched networks,” *Performance Evaluation*, 27&28, 1996.
- [6] M. Coates, R. Nowak. “Network loss inference using unicast end-to-end measurement”, *ITC Conf. IP Traffic, Modeling and Management*, 2000.
- [7] M.J. Coates and R. Nowak, “Network Delay Distribution Inference from End-to-end Unicast Measurement,” *Proc. IEEE ICASSP*, May 2001.
- [8] A.P. Dempster, N.M. Laird, D.B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm (with discussion)”, *J. Roy. Statist. Soc. Ser.*, vol. 39, pp. 1–38, 1977.
- [9] A.B. Downey. “Using pathchar to estimate Internet link characteristics,” *Proc. SIGCOMM’99* Sept. 1999.
- [10] N.G. Duffield and F. Lo Presti, “Multicast Inference of Packet Delay Variance at Interior Network Links”, in Proc. IEEE Infocom 2000, Tel Aviv, March 2000.
- [11] N.G. Duffield, F. Lo Presti, V. Paxson, D. Towsley, “Inferring link loss using striped unicast probes,” Proc. IEEE Infocom 2001.
- [12] T. Friedman, R. Cáceres, K. Almeroth, K. Sarac, “RTCP Reporting Extensions”, work in progress, January 2001.
- [13] V. Jacobson, Pathchar - A Tool to Infer Characteristics of Internet paths. For more information see <ftp://ftp.ee.lbl.gov/pathchar>
- [14] E.L. Lehmann and G. Casella, “Theory of Point Estimation”, Springer Texts in Statistics, Springer, New York, 1998.
- [15] R.J.A. Little and D.B. Rubin, “Statistical Analysis with Missing Data”, Wiley, New York, 1987
- [16] F. Lo Presti, N.G. Duffield, J. Horowitz and D. Towsley, “Multicast-Based Inference of Network-Internal Delay Distributions”, submitted for publication, September 1999.
- [17] G.J. McLachlan and T. Krishnan, “The EM algorithm and extensions”, Wiley, New York, 1997.
- [18] ns – Network Simulator. See <http://www-mash.cs.berkeley.edu/ns/ns.html>
- [19] M.J. Schervish, “Theory of Statistics”, Springer, New York, 1995.
- [20] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications,” RFC 1889, January 1996.
- [21] Y. Shavitt, X. Sun, A. Wool, B. Yener. “Computing the Unmeasured: An Algebraic Approach to Internet Mapping” IEEE Infocom 2001, Anchorage, Alaska, April 22-26, 2001.
- [22] C.F. Jeff Wu, “On the convergence properties of the EM algorithm”, *Annals of Statistics*, vol. 11, pp. 95–103, 1982



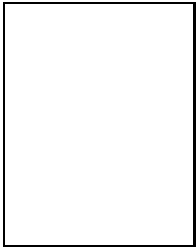
Nick Duffield (duffield@research.att.com) received his B.A. in natural sciences from Cambridge University, UK, in 1982, and his Ph.D. in mathematical physics from the University of London, UK, in 1987. He subsequently held post-doctoral and faculty positions in Heidelberg, Germany and Dublin, Ireland. He is a Technology Consultant in the Internet and Networking Research group at AT&T Labs–Research in Florham Park, NJ, USA. His current research focuses on Internet performance measurement, inference, and analysis.



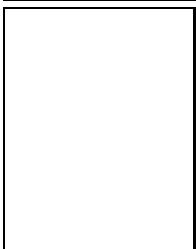
Joseph Horowitz (joe@math.umass.edu) received his B.S. degree from MIT in 1962, and his M.S. and Ph.D. degrees from the University of Michigan in 1963 and 1967, all in mathematics. He joined the Department of Mathematics and Statistics at the University of Massachusetts at Amherst in 1969, where he has been a professor since 1980. He has been a visiting faculty member at Stanford University, the University of Strasbourg, ETH-Zuerich, and the Technion (Haifa), and a Fulbright Research Fellow (1988, 1992) at the Indian Statistical Institute, New Delhi.



Don Towsley (towsley@cs.umass.edu) is a distinguished professor in the Department of Computer Science at the University of Massachusetts at Amherst. His research focuses on Internet modeling and analysis and network support for streaming media applications. He is a Fellow of the ACM and IEEE and has won several best paper awards, including the 1998 IEEE Communications Society William Bennett award.



Wei Wei (weiwei@cs.umass.edu) received the B.S. in Applied Mathematics from Peking University in 1992, the M. S. in Statistics from Texas A&M University in 2000. Currently he is a graduate student in Department of Computer Science at University of Massachusetts Amherst.



Timur Friedman (M 1997) (friedman@cs.umass.edu) received the A.B. in philosophy from Harvard College in 1987, the M.S. in technology management from Stevens Tech in 1991, and the M.S. in computer science from UMass Amherst in 1994. He is a candidate for the Ph.D in computer science from UMass Amherst in 2002. His dissertation is entitled "Scalable Estimation of Multicast Session Characteristics". He will be a postdoctoral fellow at the Laboratoire d'informatique de Paris 6 in 2002."