# Rice University, ECE Department, Technical Report TR-0002
# January 21, 2000
# Network Inference from Passive Unicast Measurements

Mark Coates and Robert Nowak *
*Department of Electrical and Computer Engineering, Rice University*
*6100 South Main Street, Houston, TX 77005–1892*
*Email: {mcoates, nowak}@ece.rice.edu, Web: www.dsp.rice.edu*

January 21, 2000

**Abstract**

The fundamental objective of this work is to determine the extent to which unicast, end-to-end network measurement is capable of determining internal network losses. We show that it is not possible to determine internal losses based solely on unicast, end-to-end measurement. However, by identifying and incorporating reasonable prior information or constraints, we demonstrate that it is possible to resolve these losses. The major contributions of this report are three-fold: we formulate a measurement procedure for network loss inference based on end-to-end packet pair measurements, we identify suitable prior probability models for network inference, and we develop a novel factor graph framework for inference calculation. Simulation experiments demonstrate the potential of our new framework.

*This work was supported by the National Science Foundation, grant no. MIP–9701692, the Army Research Office, grant no. DAAD19-99-1-0349, the Office of Naval Research, grant no. N00014-00-1-0390, and Texas Instruments.

# 1 Introduction

In large-scale networks, end-systems cannot rely on the network itself to cooperate in characterizing its own behavior. Network tomography (the inference of internal network behavior based on end-to-end network measurements) is a vital component in current efforts to transform large-scale internetworks into well-understood and predictable systems. Several groups have considered the network tomography problem [7, 16, 17, 18, 2, 13]; while promising, these methods require special support from the network in terms of either cooperation between hosts, internal network measurements, or multicast capability. Many networks do not currently support multicast due to its scalability limitations (routers need to maintain per group state), and lack of access control. In contrast, unicast inference is easily deployable on a scalable commercial infrastructure.

There is an urgent need for flexible measurement and inference tools that allow us to obtain useful information from end-to-end network measurements. To accomplish this we require new algorithms that fuse end-to-end traffic statistics with prior information about network behavior — a capability not provided by existing methods. We propose a new theory and tool for gauging internal network loss characteristics solely from unicast, end-to-end measurements, requiring no special-purpose network support and taking advantage of the wealth of information available in existing network traffic. There is a potential for new edge-based strategies of network prediction and control based on the novel theoretical framework for network inference developed in this report. A key strength of our methodology is that it can deliver not only point estimates and confidence intervals, but also probability distributions for network parameters of interest. This provides the complete characterization of the accuracy and reliability of inferred network behavior that is necessary for modelling, maintenance, and service provisioning.

The inherent structure of networks makes this problem ideally suited to the new field of factor graph analysis. We propose an inference framework based on probabilistic factor graphs and Bayesian analysis. We employ two types of statistics in our analysis — single packet losses and joint packet pair losses. The joint statistics are crucial because they capture key temporal correlations. These spatio-temporal statistics provide an enormous amount of information about internal network behavior; however harnessing this information for practical inference is a daunting task. Factor graphs enable us both to visualize the relationships between statistics and network parameters and to greatly simplify the tomography problem through probability factorization [8]. Probabilistic factor graph representations thus provide a theoretical and computational foundation for our spatio-temporal network analysis methodology.

## 1.1 Related Work

Several groups have considered the problem of estimating source-destination rates based on individual link-level traffic measurements [16, 17, 18]. These techniques are useful in network design, routing, and optimization, but require special support from the network to collect link level statistics. Conversely, others have made efforts to measure loss and delay statistics between participating hosts or routers based on the exchange of unicast probes [1, 4, 12]. There have also been recent proposals for network tomography based on end-to-end measurements. The Felix project employs linear decomposition techniques to determine network topology [3]. Two other approaches use active, multicast probing to gather packet correlation statistics and infer internal network characteristics [7, 2, 13]. Here we propose an alternative approach that does not require special support from the network (such as multicast capability) and enables the fusion of end-to-end traffic statistics with prior information about the network.

### 1.2 Bayesian Network Modelling and Analysis

One of the fundamental objectives of this work is to determine the extent to which unicast, end-to-end (edge-based) measurement is capable of determining internal network losses. By end-to-end we mean measurements of packet losses along entire paths, from source to receiver(s). Without special (and usually unrealistic) assumptions it can be shown that strictly speaking the internal loss rates on individual links are not uniquely determined by end-to-end statistics. That is, more than one internal loss configuration can give rise to the same end-to-end measurement. This is a classical example of an underdetermined or ill-posed inverse problem. Thus, the objective stated above boils down to the following question: *Can we identify reasonable prior information or constraints to uniquely resolve the losses based on end-to-end measurement?*

Ill-posed inverse problems similar to the network tomography problem arise routinely in many fields of science and engineering [10]. Determination of useful solutions to such problems depends on the clever incorporation of prior knowledge or constraints. For example, in the field of image processing a commonly used technique is to incorporate prior information reflecting the high probability of similarity between the gray levels of neighboring pixels [9]. In this report, we aim to develop a similar probabilistic modeling procedure that captures the expected characteristics of real-world networks. Perhaps the most simple possibility that comes to mind, following the image modeling approach, is to model the losses in spatially neighboring links as (probabilistically) correlated in some fashion (e.g., if one link is experiencing heavy losses, then its neighbors are also probably very lossy.) However, we argue that this type of modeling is unrealistic in many real situations and, moreover, may lead to grossly inaccurate loss estimates.

We propose instead a network modeling framework based on the correlation between unconditional (single packet) losses and conditional (back-to-back) packet losses. Based on theoretical queue models, we show that the unconditional and conditional loss probabilities are coupled. Throughout the remainder of the report we work with "success" probabilities (probability of non-loss) instead of loss probabilities. This provides a more convenient mathematical parameterization of the problem, and the probability of loss is simply one minus the probability of success. We regard the two types of success probabilities as random variables themselves and model the coupling between these random variables with a joint probability density function. To state it very simply, the conditional success probability of a packet, given that the preceding packet was successfully received, is lower than the unconditional success probability (probability of a packet being lost irrespective of whether or not the preceding packet was received). This relationship has also been verified experimentally in real networks [11]. Through the use of carefully designed joint prior probability models for the conditional and unconditional success probabilities that reflect this relationship, we develop a framework for the statistical estimation of internal success probabilities based solely on end-to-end measurement.

The report is organized as follows. In Section 2, we discuss the basic inference problem in more generality. In Section 3, we formally define our measurements and prior probability models for network success probabilities. In Section 4, we describe the factor graph modeling and inference process. In Section 5, we give a simulation experiment. In Section 6, we draw conclusions and discuss avenues for future work.

## 2 End-to-end measurements and tomography

When we restrict ourselves to edge-based measurement of a tree topology, the simplest data we can collect are counts of the number $n_i$ of packets sent from the source $S_0$ to a receiver $R_i$ and the number $m_i$ of these that were lost. Using these measurements and the independence assumptions we can form maximum likelihood estimates $\{\widehat{\rho_i}\}$ of the true source-to-receiver path success rates $\{\rho_i\}$, where $\widehat{\rho_i} \equiv \frac{m_i}{n_i}$. These estimates converge to the true success rates as the amount of collected data grows large. However, there is no unique mapping of the path-level success probabilities to the link-level success rates $\boldsymbol{\alpha}$; we are faced with an unidentifiable system. This is clear even for the simple triad network of Figure 1. We can collect measurements for only two paths in this network, which is clearly insufficient to resolve the three link success parameters uniquely. The system would be identifiable if we knew one of the success probabilities beforehand, or if we could guarantee that one of the three links was perfect. In general, we observe that each internal node in the logical tree introduces an extra degree of rank deficiency in the system of (linear) equations that relate (log) path success rates to (log) link success rates. This means that there are an infinite number of link success probability vectors $\boldsymbol{\alpha}$ that give rise to the same path success probability vector $\boldsymbol{\rho}$. Hence, no matter how many data are collected, we cannot fully resolve the individual link probabilities.

As the simplest packet count measurements do not provide sufficient information for identifiability, we turn our attention to more sophisticated measurements. Measurements made using back-to-back packet pairs provide an opportunity to generate spatio-temporal statistics of the system. By back-to-back packet pairs we mean two packets that are sent one after the other by the source, possibly destined for different receivers, but sharing a common set of links in their paths. The reason for introducing back-to-back measurements is that we anticipate correlation between the link-level success probabilities of closely-spaced packets travelling along the same links and hope to exploit this correlation. Evidence for such correlation has been provided by observations of the Internet; packet losses are observed to occur in bursts [5, 11]. On the basis of this evidence, we model the success probabilities of packets belonging to different packet pairs as independent,

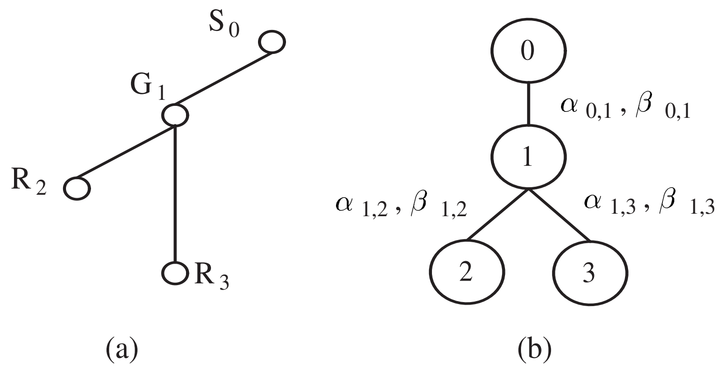


Figure 1: (a) A simple network involving a source $S_0$, a router $G_1$, and two receivers $R_2$ and $R_3$. (b) Its logical tree representation with unconditional ($\alpha$) and conditional ($\beta$) success probabilities parameterising each link.

but introduce a new parameter to capture the correlation of packet success rates within pairs.

We now characterise the loss behaviour of the link between nodes $j$ and $k$ by two parameters: the unconditional probability $\alpha_{j,k}$ of receiving a packet from $j$ at $k$, and the conditional probability $\beta_{j,k}$ of receiving a packet given that it was the second of a packet pair traversing the link between $j$ and $k$ and that the first packet was received (not dropped). Internet loss observations (see [15] lead us to conjecture that for most networks each conditional probability $\beta_{j,k}$ is greater than or equal to the corresponding unconditional success probability $\alpha_{j,k}$. This is intuitively reasonable since the condition that the first packet in a pair is successfully received suggests that it is probable that the link is not presently experiencing heavy cross-traffic and losses. In fact, based on experimental data [11], it is not unreasonable to expect that in many cases ${}_{j,k}$ is known for each link, then there is a unique mapping from the known (or estimated) $\boldsymbol{\rho}$ and $\boldsymbol{\beta}$ vectors to an $\boldsymbol{\alpha}$ vector, implying identifiability of the system and enabling the formation of robust maximum likelihood estimates of the $\alpha_{j,k}$ values.[1]

In the case where the $\{\beta_{j,k}\}$ are not known, the problem of unidentifiability remains. The back-to-back measurements provide us with more data (if there are $N$ receivers then there are $N^2$ back-to-back receiver "pairings"), resulting in $N$ count statistics per receiver. However, many of these measurements are redundant, involving the same links and related losses. In general, although there may be fewer unknown parameters than back-to-back statistics, the resulting system of equations is underdetermined.

The parameterized triad logical tree in Figure 1(b) illustrates the packet pair approach. For example, statistics collected from the first packets in back-to-back pairs with the first packet sent to node 2 and the second to node 3 provide information about the relationship between $\alpha_{0,1}$, $\alpha_{1,2}$ and the unconditional (single packet) path success denoted by $\rho^u_{0,2}$; statistics from the second

[1]We can consider the multicast success inference procedure proposed in [7] as being equivalent to the case of total correlation ($\{\beta_{j,k}\}$ are all identically one). An example clarifies this interpretation: multicast probes sent through the triad network of Figure 1 provide exactly the same information as back-to-back packet pairs if $\beta_{0,1}$ is equal to one. Given $\beta_{0,1}$, we can uniquely determine the remaining $\alpha_{j,k}$ and $\beta_{j,k}$ from the end-to-end packet-pair measurements.

probe (back-to-back packet) provide information about the relationship between $\beta_{0,1}$, $\alpha_{1,3}$ and the conditional path success denoted by $\rho^c_{0,3}$. However, as noted above, even perfect knowledge of all the $\rho^u$ and $\rho^c$ parameters does not produce a unique mapping to the $\alpha$ and $\beta$ values.

In order to uniquely determine the network parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ we must incorporate additional information into the estimation process. Specifically, we advocate the use of prior knowledge of parameters to help guide or "regularize" the estimation process. The Bayesian approach to parameter inference provides us with a formal means for combining measurements and prior information. It is important to emphasise that the modelling we will now describe is *a priori* modelling, i.e., before any measurements are made. First of all, it is our contention that in the most general situations it is unreasonable to assume specific prior information about the unconditional probabilities $\boldsymbol{\alpha}$ is available. That is, in general we cannot assume that any one link is "good" or "bad" and, moreover, we cannot assume any specific relationship (deterministic or probabilistic) between link losses. However, empirical observations of Internet traffic and theoretical results for network models suggests that the assignment of partially-informative prior probabilities to the conditional probabilities $\boldsymbol{\beta}$ is reasonable. This is a key insight and major contribution of this report: *joint prior probability modeling of the network parameters* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$*, as described next, enables identification of the most probable network configuration, based soley on (unicast, end-to-end) packet-pair measurements.*

Specifically, based on physical considerations of the network we assume that a conditional probability $\beta_{j,k}$ will be larger than its corresponding $\alpha_{j,k}$. We can formally characterise this belief by placing a partially informative (conditional) prior probability distribution $p\left(\beta_{j,k}|\alpha_{j,k}\right)$ on $\beta_{j,k}$ *given* $\alpha_{j,k}$. The distribution $p\left(\beta_{j,k}|\alpha_{j,k}\right)$ models the (uncertain) relationship between $\alpha_{j,k}$ and $\beta_{j,k}$. The least presumptive (non-informative) prior probability model that satisfies the basic physical constraints of the problem (*i.e.,* probabilities lie between 0 and 1 and $\beta_{j,k} \geq \alpha_{j,k}$) is: the $\alpha_{j,k}$ parameters are independent and uniformally distributed over $[0, 1]$ and the $\beta_{j,k}$ parameters are mutually independent and conditionally uniformally distributed over $[\alpha_{j,k}, 1]$. We will further discuss the motivation for this choice and others in the forthcoming sections.

The prior model described above, and others like it, introduce probabilistic coupling between the $\alpha$ and $\beta$ parameters, which, in conjunction with back-to-back measurement, provides a means for determining the most probable network configuration; i.e., most probable network parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. The key to our approach is that we identify the most probable joint $\alpha_{j,k}$ and $\beta_{j,k}$ pairs rather than attempting to identify them independently. The combination of prior probability models and measured data is formally carried out using the Bayesian probability calculus, and we determine *a posteriori* probability distributions for the network parameters. The posterior distributions can be used to identify the most probable parameter settings, as well as confidence intervals for these estimates.

# 3 Formalisation of the Adopted Model and Measurements

In this section, we formally define both our model for the network and the measurements that we make. We consider the analysis of networks comprised of a single source and multiple receivers, and represent them (using the same notation as in [7]) with a logical tree $\mathcal{T} := (\mathcal{V}, \mathcal{L})$, consisting of the set of nodes $\mathcal{V}$ that represent the source, routers and receivers in the network, and the set of links $\mathcal{L}$ connecting the nodes. Each node in the logical tree has a single parent and a number of children. End-to-end measurements made on an isolated subpath (a subpath consisting of two or more links in which internal nodes have only one child) do not provide sufficient information to resolve the individual losses in the isolated subpath. Thus, if isolated subpaths exist in the network under study, we remove the internal subpath nodes during the formation of the logical tree and use a single composite link to represent the isolated subpath. All nodes in the logical tree thus have at least two children, apart from the source (one child) and the receivers (no children).

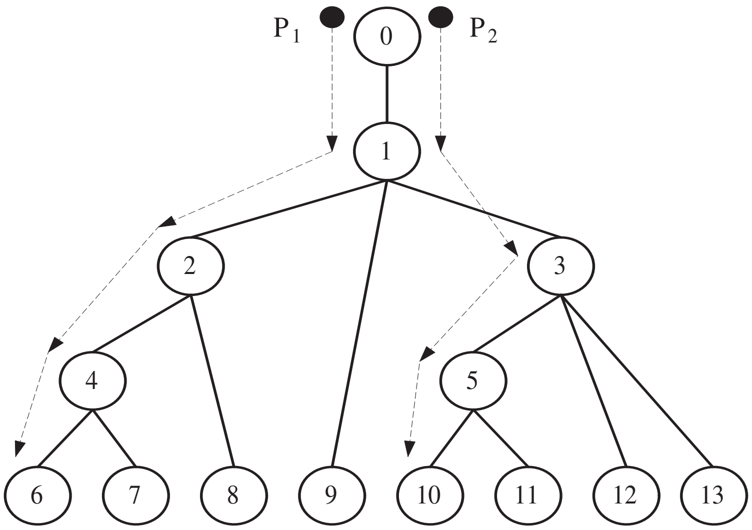


Figure 2: An example of a depth-4 tree. The arrows depict the paths traversed by the first packet ($P_1$) and second packet ($P_2$) of a $(6, 10)$ packet pair.

Figure 1(b) depicts the logical tree of the simplest non-trivial network (a triad). Figure 2 depicts a depth-4 general tree that we will use to illustrate the form of the measurements and the likelihood function. We use $\mathcal{R}$ to represent the set of leaf nodes (receivers) and $\mathcal{T}(k) := (\mathcal{V}(k), \mathcal{L}(k))$ to denote the subtree rooted at node $k$. $\mathcal{R}(k) := \mathcal{R} \cap \mathcal{V}(k)$ denotes the set of receivers that are descendants of node $k$. $\mathcal{P}(j, k) \subset \mathcal{L}$ denotes the set of links included in the shortest path from node $j$ to node $k$. The level $l(j)$ of node $j$ is defined as the cardinality of the set $\mathcal{P}(0, j)$.

Our goal is to estimate the success rates on individual links within the tree structure. If we wish to estimate success probabilities for the entire tree, then for each pair of receivers $(r_j, r_k)$, we gather the following statistics using (*active*) back-to-back probing pairs or (*passive*) back-to-back pairs arising naturally in source traffic generation. The term "$(j, k)$-pairs" is used to denote those pairs in which the first packet was destined for node $j$ and the second for receiver $k$. The nature of a $(6, 10)$ pair is shown by the two arrows in Figure 2: the first packet $(P_1)$ is sent to node 6, and the second $(P_2)$ to node 10.

We measure the number $n_{j,k}$ of $(j, k)$-pairs in which the first packet was successfully received and also count the number $m_{j,k}$ of these $n_{j,k}$ pairs in which the second packet was also successfully received. Similarly, we record $n_{k,j}$ and $m_{k,j}$ from the $(k, j)$-pairs. By considering all receiver pairs, we generate two sets of measurements $\mathcal{M} := \{m_{j,k}; j, k \in \mathcal{R}\}$ and $\mathcal{N} := \{n_{j,k}; j, k \in \mathcal{R}\}$.

We model the loss processes on separate links as mutually independent. Although spatial dependence (correlated success probabilities on neighbouring links) may be observed in networks due to common traffic, such dependence highly circumstantial and cannot be readily incorporated in a model that is intended to be generally applicable to a variety of networks. Bolot *et al.* proposed Markovian models of packet loss in [6] based on observations of Internet traffic. Although such models do not fully account for the extended loss bursts observed in [11], we adopt a similar approach for modelling the packet loss processes on each link (the model is reminiscent of that used to explore temporal dependence in [7]).

Let us assume a total of $N$ packet pairs are sent to each $(j, k)$ receiver pair. We assume that separate packet pairs are sufficiently spaced so that the link loss processes for the first packets of all pairs can be modelled as a set of mutually independent Bernoulli processes $X(n) = \{X_{j,k}(n);\ (j, k) \in \mathcal{L}\}$. $X_{j,k}(n)$ is the value of the first packet loss process on link $(j, k)$ for the $n$-th packet pair, and takes a value 0 to indicate a loss occurred and 1 to indicate successful transfer. $X_{j,k}(n)$ is Bernoulli with probability $\alpha_{j,k}$ of being in state 1. The link loss processes for the second packets in the pairs, denoted $Y_{j,k}(n)$, are conditionally independent given $X(n)$. If $X_{j,k}(n) = 1$, then $Y_{j,k}(n)$ is Bernoulli with probability $\beta_{j,k}$ of being in state 1.

Using this model, it is possible to write an expression for the probability of observing an $(m_{j,k}, n_{j,k})$ pair[2]:

$$\begin{aligned} Q(j,k) &= \prod_{(s,t)\in\mathcal{P}(0,r)} \beta_{s,t} \prod_{(s,t)\in\mathcal{P}(r,k)} \alpha_{s,t} \\ p\left(m_{j,k}|n_{j,k},\alpha_{j,k},\beta_{j,k}\right) &\propto \left(Q(j,k)\right)^{m_{j,k}} \left(1 - Q\left(j,k\right)\right)^{n_{j,k}-m_{j,k}} \end{aligned}$$

where $r = \max_p \{l(p) : j, k \in \mathcal{R}(p)\}$ is the node at which the two paths diverge. The likelihood of observing $\mathcal{M}$ and $\mathcal{N}$ is then:

$$p\left(\mathcal{M}|\mathcal{N}, \boldsymbol{\alpha}, \boldsymbol{\beta}\right) = \prod_{j,k\in\mathcal{R}} p\left(m_{j,k}|n_{j,k},\alpha_{j,k},\beta_{j,k}\right)$$

The system is not identifiable, and there is not a unique pair of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ vectors that maximise

[2]The proportionality in the expression indicates that there is a constant factor that is independent of the parameters, but necessary as a normalisation factor for constructing a valid density. As the factor does not affect any comparative calculations, it can be safely neglected.

this likelihood. In a Bayesian framework we can combine the likelihood with our prior knowledge or beliefs to form a posterior density:

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathcal{M}, \mathcal{N}) \;\propto\; p(\mathcal{M} | \mathcal{N}, \boldsymbol{\alpha}, \boldsymbol{\beta})\, p(\boldsymbol{\beta}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha})$$

The generation of this density requires the specification of a prior density on the $\alpha_{j,k}$ parameters and a conditional prior density on the $\beta_{j,k}$ parameters given $\boldsymbol{\alpha}$.

As discussed in the previous section, we have no prior knowledge about the $\alpha$ parameters in a general setting, so a non-informative prior density must be selected. A specific success probability $\alpha_{j,k}$ is modelled as uniformally distributed over the range 0 to 1, i.e., $p(\alpha_{j,k}) = U[0,1]$ and $p(\boldsymbol{\alpha}) = \prod_{(j,k)\in\mathcal{L}} p(\alpha_{j,k})$. It is however reasonable to assume that $\beta_{j,k}$ is greater than $\alpha_{j,k}$. A obvious choice is model the $\{beta_{j,k}\}$ as independent with non-informative uniform densities, $p(\beta_{j,k}|\alpha_{j,k}) = U[\alpha_{j,k}, 1]$. However, there are other reasonable choices for the prior $p(\beta_{j,k}|\alpha_{j,k})$, as discussed later in Section 5.

## 4 Implementation: A Factor Graph Approach

Factor graphs (or Bayesian networks) enable us both to visualize the relationships between statistics and network parameters and to greatly simplify the tomography problem through probability factorization [8]. The graphical structure of the probabilistic factor graph representations thus provides a theoretical and computational foundation for our network loss analysis methodology. This factorization facilitates a computationally efficient probability propagation strategy to determine the posterior success probabilities, optimally fusing the prior probabilities with the measurement statistics. Indeed, without such factorization, the problem becomes computationally overwhelming for networks of reasonable size. The application of factor graphs is naturally suited not only to the estimation problem presented in this report, but to a variety of network problems, because of their corresponding graphical structures.

We do not intend to provide a detailed account of how factor graphs are derived or how to apply graphical methods for inferring posterior probabilities (see [8] for details). We merely want to illustrate the principles underpinning the factor graph framework, and therefore focus on the problem of inferring success probabilities across the internal links in the triad network of Figure 1. Figure 3(a) depicts the form of the packet-pair used to collect the statistics $m_{2,3}$ and $n_{2,3}$, as described in the previous section. The packet-pair consists of the first packet $P_1$ destined for node 2 and the second packet $P_2$ destined for node 3. We also consider (unconditional) single packet statistics. Let $n_2$ denote a number of single packets sent to node 2 and let $m_2$ denote the number of these actually received. For clarity, we consider the factor graph when only these statistics, and the statistics $n_2$ and $m_2$ have been collected (using single probe packets sent to node 2). If the other measurable statistics ($n_3$, $m_3$, $n_{3,2}$, and $m_{3,2}$) are available, then the resulting factor graph is a relatively straightforward extension of the one we derive.

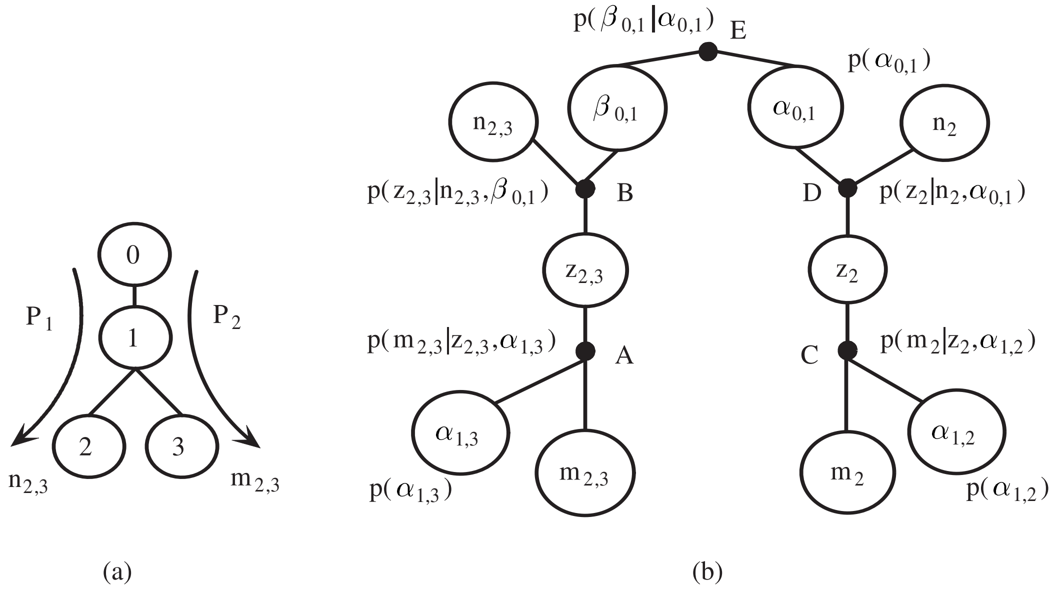


Figure 3: (a) The structure of (2-3) packet-pairs, and the recorded statistics. (b) The factor graph representation available from experiments that measure $m_{2,3}$, $n_{2,3}$ and $m_2$, $n_2$.

The factor graph in Figure 3(b) is derived by considering the dependencies between all network parameters and measured statistics. The hollow circles represent the parameters/data in the experiment; the filled circles (labelled $A$–$E$) indicate dependencies between these parameters. The right branch of the graph arises from the consideration of the $m_2$ and $n_2$ statistics, and the left branch is determined from the $n_{2,3}$ and $m_{2,3}$ statistics. The branches are connected because of our assumption of a dependency between the values of $\beta_{0,1}$ and $\alpha_{0,1}$. The factor graph for this experiment introduces two auxiliary variables ($z_{2,3}$ and $z_2$) that represent unobserved or "hidden" data and enable a simplified factorization of the probabilistic network model. In this case, $z_2$ is the number of the $n_2$ probes that reached node 1, and $z_{2,3}$ is the number of the $n_{2,3}$ probes that reached node 1. The conditional probabilities at $A$–$D$ describe the functional relationship between the parameters connected to the nodes. Due to the assumed model, these conditional probabilities are binomial, and depend only on a (potentially conditional or unconditional) success probability parameter of the link of interest. Our prior beliefs are embedded at $E$ which assigns the prior distribution $p(\beta_{0,1}|\alpha_{0,1})$ as the functional relationship between the conditional and unconditional success parameters on the $(0, 1)$ link.

Once the appropriate factor graph has been derived, probability propagation algorithms (see [8]) can be used to generate posterior distributions for the network parameters. Prior distributions from the Beta-family are conjugate to the binomial likelihoods, so if we choose from this family, analytic expressions can be derived for the posterior distributions.

## 5 Simulation Experiments

We experimented using simulations based on the network in Figure 2link in the network to assume one of two state values, 0 representing congestion, and 1 representing a light traffic burden. At time instants $t \in T$, the state of each link was updated according to a Markov process. The transition probability matrix of the process governing the state of link $(u, v)$ was determined by drawing $\alpha_{u,v}$ from a uniform distribution $U[0, 1]$, and then drawing $\beta_{u,v}$ from $U[\alpha_{u,v}, 1]$; the matrix was designed so that if traffic were sent across the link it would experience a steady-state success probability of $\alpha_{u,v}$ and a conditional success probability of $\beta_{u,v}$. Packet-pair probes were sent in to the various receivers in an ordered fashion designed to extract an informative subset of the possible $m_{j,k}$ and $n_{j,k}$. The times at which the first packets of these pairs were sent were determined from a Poisson process, such that interarrival times were well-separated. The second packet in a pair was sent one time instant later. 1600 packet pairs were sent through the network, with the destinations designed so that there was a uniform distribution across the network of divergence nodes (the node at which the paths of the individual packets in the packet-pairs separated). Such a distribution guarantees an equal (prior) exploration of all network parameters.

Figure 4 depicts the result of one of the experiments. The posterior distribution of success probability was calculated for each link, and plotted in the boxes; the arrows mark the true values. The confidence that can be placed on an estimate is clearly dependent on the amount of data that can be collected; estimation of the success probabilities of links $(2, 4)$, $(4, 6)$, $(4, 7)$ are generated from packet-pairs involving a packet travelling from the source to either node 6 or 7, both of which are extremely lossy paths. One hundred random trials were conducted for each of the three priors (uniform, queue-based scale-invariant, and queue-based exponential) depicted in Figure **??**. When the success rate was estimated using the peaks of the generated distributions, the mean absolute errors were 0.084, 0.069, and 0.074, respectively.

## 6 Conclusions

We have demonstrated that unicast, end-to-end measurement is capable of determining internal network losses provided that prior information can be incorporated into the estimation process. We have identified reasonable prior information models to resolve the losses based on solely end-to-end measurement. The factor graph framework enables efficient inference methods based on probability propagation, and we believe that our preliminary experimental results support further investigation of this new methodology.

## References

[1] IPMA: Internet performance measurement and analysis. See www.merit.edu/ipma.

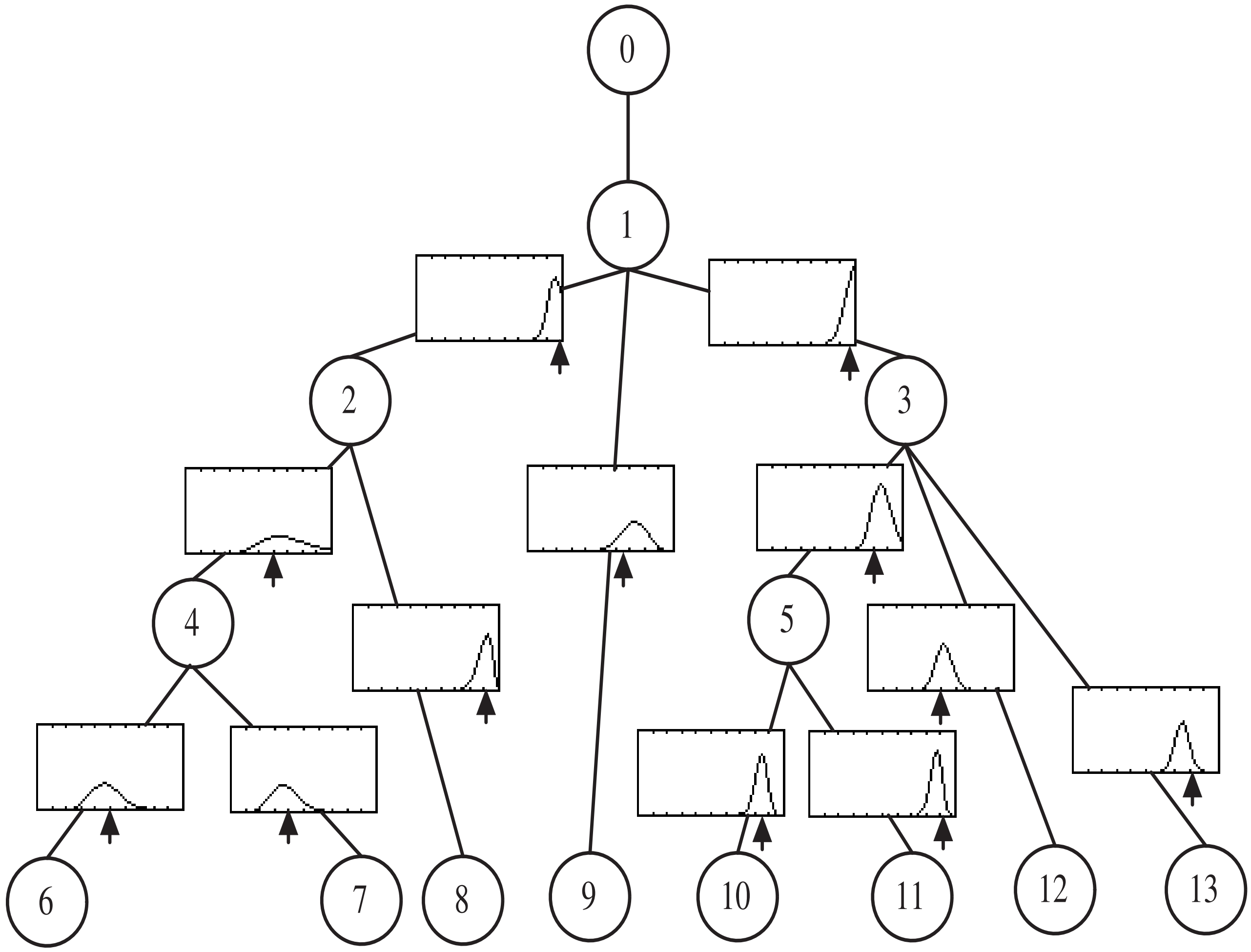


Figure 4: An example of the results of the experiment described in Section 5. 1600 packet pairs were sent to various receivers in order to generate posterior probability distributions of the success rates of the links. These are plotted in the boxes on the links; the arrows mark the true values.


[2] Multicast-based inference of network-internal characteristics (MINC). See gaia.cs.umass.edu/minc.

[3] Project Felix: Independent monitoring for network survivability. See govt.argreenhouse.com/felix.

[4] Surveyor: An infrastructure for Internet performance measurements. INET'99, June 1999. See io.advanced.org/surveyor.

[5] J-C. Bolot. End-to-end packet delay and loss behaviour in the Internet. In *Proc. SIGCOMM '93*, pages 289–298, Sept. 1993.

[6] J-C. Bolot and A. Vega Garcia. The case for FEC-based error control for packet audio in the internet. to appear in ACM Multimedia Systems.

[7] R. Cáceres, N. Duffield, J. Horowitz, and D. Towsley. Multicast-based inference of network-internal loss characteristics. *IEEE Trans. Info. Theory*, 45(7):2462–2480, November 1999.

[8] B. Frey. *Graphical Models for Machine Learning and Digital Communication.* MIT Press, Cambridge, 1998.

[9] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. 6(6):712–741, 1984.

[10] Finbarr O'Sullivan. A statistical perspective on ill-posed inverse problems. 1(4):502–527, 1986.

[11] V. Paxson. End-to-end Internet packet dynamics. *IEEE/ACM Trans. Networking*, 7(3):277–292, June 1999.

[12] V. Paxson, J. Mahdavi, A. Adams, and M. Mathis. An architecture for large-scale Internet measurement. *IEEE Communications*, 3:226–243, 1998.

[13] S. Ratnasamy and S. McCanne. Inference of multicast routing trees and bottleneck bandwidths using end-to-end measurements. In *Proceedings of INFOCOM '99*, New York, NY, March.

[14] D. Rubenstein, J. Kurose, and D. Towsley. Detecting shared congestion of flows via end-to-end measurement. To appear in Proc. ACM SIGMETRICS'00 (Santa Clara, CA, June 2000).

[15] C. Tebaldi and M. West. Bayesian inference on network traffic using link count data (with discussion). *J. Amer. Stat. Assoc.*, pages 557–576, June 1998.

[16] S. Vander Wiel, J. Cao, D. Davis, and B. Yu. Time-varying network tomography: router link data. In *Proc. Symposium on the Interface: Computing Science and Statistics*, Schaumburg, IL, June 1999.

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