

Multicast Inference of Temporal Loss Characteristics

Vijay Arya^a, N.G. Duffield^b, Darryl Veitch^c

^a*National ICT Australia (NICTA), Victorian Research Laboratory,
University of Melbourne, Australia*

^b*AT&T Labs–Research, Florham Park, NJ, USA*

^c*ARC Special Research Centre on Ultra-Broadband Information Networks
(CUBIN), Dept of EEE, University of Melbourne, Australia
(CUBIN is an affiliated program of NICTA)*

Abstract

Multicast-based inference has been proposed as a method of estimating average loss rates of internal network links, using end-to-end loss measurements of probes sent over a multicast tree. We show that, in addition to loss rates, temporal characteristics of losses can also be estimated. Knowledge of temporal loss characteristics has applications for services such as VoIP which are sensitive to loss bursts, as well as for bottleneck detection. Under the assumption of mutually independent, but otherwise general, link loss processes, we show that probabilities of arbitrary loss patterns, mean loss-run length, and even the loss-run distribution, can be recovered for each link. Alternative estimators are presented which trade-off efficiency of data use against implementation complexity. A second contribution is a novel method of reducing the computational complexity of estimation which can also be used by existing MINC estimators. We analyse estimator performance using a combination of theory and simulation.

Key words: network tomography, end-to-end measurement, loss inference, MINC

1 Introduction

After an initial period of relatively slow take-up, native multicast services are now becoming a necessary part of Internet Service Providers' offerings. For instance, Multicast Virtual Private Networks [1] and video distribution to the home [2] are two applications of increasing importance, whose recent availability and growth has rekindled interest in scalable multicast performance measurement. On the other hand, direct measurement of all network links of

interest remains a challenge of scale. Whereas ISPs do conduct active measurements between hosts located in router centers (~ 10 - 100 for large providers), pushing these measurements out to the customer edge of the network would involve instrumenting a far larger number of customer access points. In the video service example, hundreds of local video distribution networks, each with thousands of subscribers, may need to be covered.

These considerations motivate the use of tomographic methods to infer performance of internal network links based on end-to-end measurements over multicast trees. A body of work on Multicast Inference of Network Characteristics (MINC) has shown how to infer average packet loss rates [3,4] and delay distributions [5] of network links, and even the network topology itself [6]. More recently [7], tunneling has been used to create a multicast measurement overlay with physical network edge points acting as branch points and this new measurement technology has also rekindled interest in using tomographic methods.

This paper aims to expand the capabilities of loss tomography, which currently only provides a means of estimating average loss rates. It is well known that packet traffic is both bursty and exhibits temporal dependence. Therefore average loss rates can not provide a finer grained view on loss, in particular, temporal characteristics such as the probability of two successive losses, or the mean duration of loss-runs (a group of successive losses, flanked by successful transmissions). The distribution of loss-run durations is important for services such as VOIP or video transmission which are sensitive to loss bursts [8]. Forward Error Correction (FEC) is used to mitigate the effects of such bursts; knowing their typical duration informs the selection of FEC parameters that determine the loss burst duration that may be corrected for. If the correctable duration is too short, some errors will not be recoverable; if too long, increased delays and bandwidth consumption will result.

In this paper we show how the probability of arbitrary patterns of loss, or inversely, ‘passage’, can be estimated for each link in a multicast tree, under an assumption of link independence, but arbitrary temporal structure within links. We pay particular attention to calculating the joint probability of two successive transmissions, and show that, quite generally, together with the average loss rate this is sufficient to determine the mean length of loss-runs. Alternative estimators are presented which trade-off efficiency of data use against implementation complexity.

A second contribution is a novel method, *subtree partitioning*, of reducing the computational complexity of MINC based estimation, by transforming any multicast tree into a virtual binary tree for estimation purposes. The simplification applies very generally, for example it can be used to revisit the standard problem of estimating marginal link loss rates, as well as the temporal or joint

loss probabilities we study here. Technically it results in estimators which only require the solution of either linear or quadratic equations to recover the passage probabilities of shared paths between receivers, avoiding the need for root finding of higher order polynomials which arises in approaches (based on the MLE under Bernoulli losses) when dealing with trees with higher node degree than binary (see [3]).

We know of no prior work concerned with tomographic inference of temporal loss characteristics (as distinct from the inference of average loss rates). We describe briefly two pieces of prior work related to the subtree partitioning approach, in the special case of estimating marginal link passage probabilities. First, we observe that this approach can be associated with a procedure of converting an arbitrary tree to a binary by inserting zero-loss links into it. A related approach was proposed for multicast based topology estimation in [6]. Second, a variant of the MINC approach was recently proposed in [9], which constructed an estimator as an explicit function of the measured data as means of avoiding numerical root finding. In fact, we show using experiments that subtree-partition results in lower asymptotic variance (as the number of probes grows large) than the explicit estimator of [9].

Note that all temporal loss estimation methods of this paper may, in principle, be extended to groups of unicast packets emulating multicast packets, in the same manner as [10, 11] extend [3] for inferring average loss rates.

The paper is structured as follows. Section 2 details our assumptions on loss over the tree, and Section 3 then explores what they imply. Section 4 defines the estimators we use, and their performance is explored in Section 5 using simulation. Theoretical results on the estimators are provided in Section 6, before concluding in Section 7.

2 Tree Model and Probe Processes

In this section we justify and describe how we model the loss process over a tree, and derive the key consequences.

2.1 Model Definition

Tree Model Let $\mathcal{T} = (\mathbf{V}, \mathbf{L})$ denote the logical multicast tree consisting of a set of nodes \mathbf{V} and links \mathbf{L} . Let $0 \in \mathbf{V}$ denote the root node and let $\mathbf{R} \subset \mathbf{V}$ be the set of leaf nodes. A link is an ordered pair $(\mathbf{k}, \mathbf{j}) \in \{\mathbf{V} \times \mathbf{V}\}$ representing a logical link (one or more physical links) from node \mathbf{k} to node \mathbf{j} . The set of

children of a node k is denoted by $\mathbf{d}(k)$, thus $\mathbf{d}(k) = \{j \in V : (k, j) \in L\}$. All nodes have at least two children, except the root (just one) and the leaves (none), see fig 1(a). Let $\mathbf{U} = V \setminus \{0\}$ denote the set of all non-root nodes of \mathcal{T} . For each node $k \in \mathbf{U}$ there is a unique node $j = f(k)$, the father of k , such that $(k, j) \in L$. We define $f^n(k)$ recursively by $f^n(k) = f(f^{n-1}(k))$, with $f^0(k) = k$. Now $\forall k \in \mathbf{U}$, the set of ancestors of k is $\mathbf{a}(k) = \{j \in \mathbf{U} : f^\ell(k) = j, \ell \geq 0\}$. Note $k \in \mathbf{a}(k)$. We define $\mathbf{a}(0) = 0$. For convenience, we refer to link $(f(k), k)$ simply as link k . We say that j is a descendant of k if $k \in \mathbf{a}(j), k \neq j$.

Probes on the Tree An infinite stream of probe packets indexed by $i \in \mathbb{Z}$ is dispatched from the root node 0 . Each probe that arrives at a node k results in a copy being sent to each of its children. However, a given probe i may be lost in transmission on any link k . The passage of probes down the tree is modeled by two stochastic processes $X_k = \{X_k(i)\}$ and $Z_k = \{Z_k(i)\}$ for each node k . The process X_k acts as a bookkeeping or ‘probe survival’ process: it records which probes reach node k and which do not. The loss process Z_k determines whether probes will be lost or transmitted as they attempt to traverse link k .

Node/Bookkeeping Processes The bookkeeping processes $\{X_k(i) : i \in \mathbb{Z}\}, k \in V$ describe the progress of the probes down the tree: $X_k(i) = 1$ if probe i reaches node k , and 0 otherwise. If a packet is lost on link k , it cannot reach node k nor other nodes or links descended from it. At the source, $X_0(i) = 1, \forall i$.

Modelling Link Loss Ideally, the state of links over the tree could be characterized by a set of underlying continuous time random processes $\{Z_k(t) : t \in \mathbb{R}\}, k \in \mathbf{U}$. If a packet attempts to traverse link k at time t and $Z_k(t) = 1$, then it is successfully transmitted, whereas if $Z_k(t) = 0$, it is lost. We consider an abstracted form of the problem, where the probe index i plays the role of discrete time. In this picture, the discrete time loss processes are well defined for **all** i , with the interpretation that $Z_k(i)$ determines whether loss *would* have occurred or not, had probe i been present.

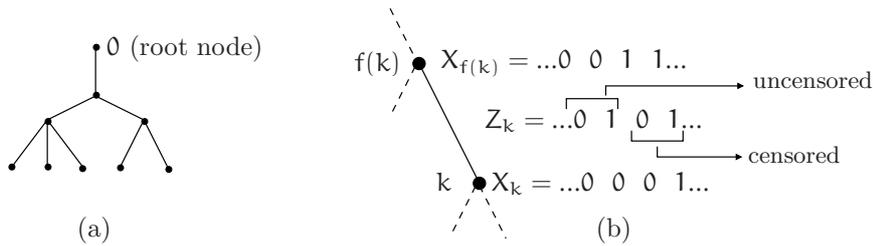


Fig. 1. (a): Example of a logical tree, (b) Loss model: the loss process $\{Z_k(i)\}$ on link k acts deterministically on the input process $X_{f(k)}$ at node $f(k)$ to produce the node process X_k .

Discrete Time Loss Model We define a binary loss process $\{Z_k(i)\}$ for each link $k \in \mathbf{U}$. It acts deterministically on the node process at each index i

as:

$$X_k(\mathbf{i}) = Z_k(\mathbf{i})X_{f(k)}(\mathbf{i}) \quad (1)$$

as illustrated in Figure 1(b). Hence, for \mathbf{i} fixed, $\{X_k(\mathbf{i}), k \in \mathbf{V}\}$ exhibits the spatial Markov structure that, conditional on the father variable $X_k(\mathbf{i})$ at node k , the child node variables $\{X_j(\mathbf{i}) : j \in \mathbf{d}(k)\}$ are independent of each other, with

$$\begin{aligned} \Pr[X_j(\mathbf{i}) = 0 | X_k(\mathbf{i}) = 0] &= 1 \\ \Pr[X_j(\mathbf{i}) = b | X_k(\mathbf{i}) = 1] &= \Pr[Z_j(\mathbf{i}) = b], \quad b \in \{0, 1\}. \end{aligned}$$

We assume the following loss dependence structure:

Spatial: The loss processes $\{Z_k\}$, $k \in \mathbf{U}$ are mutually independent.

Temporal: For each $k \in \mathbf{U}$, $\{Z_k(\mathbf{i})\}$ is stationary and ergodic.

Thus we assume that packet loss occurs independently across different links, but that within each link losses can be correlated, with parameters that in general depend on the link. Although potential mechanisms for spatial loss correlation have been identified (e.g. TCP synchronization [12]), we believe that diversity of loss rates and round-trip times will prevent such mechanisms from operating in links which carry large numbers of flows; see [3]. The assumption of ergodicity means that empirical time averages of functions of the process will converge to the true expected value almost surely, ensuring our estimators have desirable properties.

2.2 Consequences and Examples

Eq. (1) implies that $X_k(\mathbf{i})$ is determined by the product of loss processes over its ancestor nodes: $X_k(\mathbf{i}) = \prod_{j \in \mathbf{a}(k)} Z_j(\mathbf{i})$. It follows that X_k inherits the stationarity and ergodicity properties of the loss processes.

Bernoulli: For each link k the $Z_k(\mathbf{i})$ are i.i.d., a Bernoulli process characterised by the average passage probability $\alpha_k = \Pr[Z_k(\mathbf{i}) = 1]$. This is the model primarily used in prior MINC work [3]. In this simple case, the X_k are also Bernoulli, characterized by $\Pr[X_k(\mathbf{i}) = 1] = \prod_{j \in \mathbf{a}(k)} \alpha_j$.

On-Off: For each link k , define a loss-run variable L_k and a pass-run variable P_k , each taking values in $\{1, 2, \dots\}$. The link loss process Z_k is On-Off if sample paths are formed by laying down runs of 0's, interleaved with runs of 1's, according to independent copies of L_k and P_k . In general, a product of independent On-Off processes is not On-Off. However, the sub-class with geometrically distributed L_k taking 1 in the On state is closed under independent products.

Finally, note that the relation $\Pr[X_k(\mathbf{i}) = 1] = \prod_{j \in a(k)} \alpha_j$ allows the means, variances, as well as the covariance function $\nu_k(\cdot)$ to be calculated straightforwardly:

$$E[X_k(\mathbf{i})] = \prod_{j \in a(k)} \Pr[Z_k(\mathbf{i}) = 1], \quad \nu_k(\mathbf{h}) = \prod_{j \in a(k)} E[Z_j(\mathbf{i})Z_j(\mathbf{i}+\mathbf{h})] - \left(\prod_{j \in a(k)} E[Z_k(\mathbf{i})] \right)^2.$$

Assume $E[Z_j(\mathbf{i})Z_j(\mathbf{i} + \mathbf{h})] = 0$ for \mathbf{h} larger than some lag $\mathbf{h}(j)$, and let $\mathbf{h}^* = \max_j \mathbf{h}(j)$. Then $\nu_k(\mathbf{h}) = 0$ for $\mathbf{h} > \mathbf{h}^*$, so that the memory of the node processes cannot exceed that of the loss processes generating it.

3 Temporal Characteristics of Loss Processes

Before embarking on estimation, we need to establish which temporal characteristics of the processes Z_k one would wish to estimate. We choose an ambitious objective, the probability of observing any arbitrary loss pattern on a link in a general, non-parametric setting.

Sufficiency of Joint Passage Probabilities: One can express the joint probabilities of any stationary binary loss process Z_k in terms of the joint passage or transmission probabilities, that is those involving state 1 only. For example the distribution of $\{Z_k(\mathbf{i}), \dots, Z_k(\mathbf{i} + \mathbf{s})\}$ with $Z_k^{\mathbf{i}} = 0$, can be re-expressed as

$$\begin{aligned} & \Pr[Z_k(\mathbf{i}) = 0, Z_k(\mathbf{i} + 1) = \mathbf{b}_{\mathbf{i}+1}, \dots, Z_k(\mathbf{i} + \mathbf{s}) = \mathbf{b}_{\mathbf{i}+\mathbf{s}}] \\ &= \Pr[Z_k(\mathbf{i} + 1) = \mathbf{b}_{\mathbf{i}+1}, \dots, Z_k(\mathbf{i} + \mathbf{s}) = \mathbf{b}_{\mathbf{i}+\mathbf{s}}] \\ &- \Pr[Z_k(\mathbf{i}) = 1, Z_k(\mathbf{i} + 1) = \mathbf{b}_{\mathbf{i}+1}, \dots, Z_k(\mathbf{i} + \mathbf{s}) = \mathbf{b}_{\mathbf{i}+\mathbf{s}}]. \end{aligned} \tag{2}$$

The above technique can be applied recursively to eliminate all occurrences of random variables taking value 0. Therefore, the probability of observing a loss pattern on a link is accessible provided all necessary joint passage probabilities on that link can be estimated.

Mean Run Lengths: Let P_k be a random variable taking the marginal distribution of the length of a ‘pass-run’ (a run of Z_k in state 1, flanked by 0’s):

$$\begin{aligned} E[P_k] &= \sum_{j \geq 1} j \Pr[P_k = j] = \sum_{j \geq 1} \Pr[P_k \geq j] \\ &= \sum_{j \geq 1} \Pr[Z_k(j) = 1, \dots, Z_k(1) = 1 \mid Z_k(0) = 0, Z_k(1) = 1] \\ &= \sum_{j \geq 1} \frac{\Pr[Z_k(j) = 1, \dots, Z_k(1) = 1, Z_k(0) = 0]}{\Pr[Z_k(0) = 0, Z_k(1) = 1]} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j \geq 1} \frac{\Pr[Z_k(j)=1, \dots, Z_k(1)=1] - \Pr[Z_k(j)=1, \dots, Z_k(0)=1]}{\Pr[Z_k(0)=0, Z_k(1)=1]} \\
&= \frac{\Pr[Z_k(1)=1]}{\Pr[Z_k(1)=1] - \Pr[Z_k(0)=1, Z_k(1)=1]}.
\end{aligned}$$

The last sum is absolutely convergent because $E[P_k]$ is finite. This formula makes intuitive sense as the ratio of the expected proportion of time spent in pass-runs (per time index) divided by the expected frequency of transitions into state 1 (the expected number of pass-runs per time index). A similar calculation gives the mean loss-run length:

$$\mu_k = E[L_k] = \frac{1 - \Pr[Z_k(1)=1]}{\Pr[Z_k(1)=1] - \Pr[Z_k(0)=1, Z_k(1)=1]}. \quad (3)$$

Thus, the natural constraints of binary processes makes the mean loss run length accessible even without parametric help, provided we can estimate the simplest joint passage probabilities, those for a single: $\Pr[Z_k(i) = 1]$ and a pair: $\Pr[Z_k(i) = 1, Z_k(i+1) = 1]$, of probes. It turns out that these are also sufficient to estimate all parameters of a first order Markov chain. Finer details of the loss-run distribution can also be obtained, for example $\Pr[L_k \geq 2]$ can be obtained using the joint passage probabilities of one, two, and three consecutive probes.

4 Estimation of Temporal Loss Characteristics

In the previous section we discussed what the temporal link parameters could be, and highlighted the importance of joint passage probabilities. In this section we show how to estimate these from receiver data.

4.1 From Path Passage to Link Passage Probabilities

We define the probe index set $I = \{i_1, i_2, \dots, i_s\}$ (not necessarily contiguous), and the random vectors:

$$Z_k(I) = [Z_k(i_1), \dots, Z_k(i_s)], \quad X_k(I) = [X_k(i_1), \dots, X_k(i_s)].$$

We will use $\{Z_k(I) = \mathbb{B}\}$ to denote the event $\{Z_k(i_1) = \mathbf{b}_1, \dots, Z_k(i_s) = \mathbf{b}_s\}$, where $\mathbb{B} = [\mathbf{b}_1, \dots, \mathbf{b}_s] \in \{0, 1\}^{|I|}$.

We define the *link pattern passage probability* $\alpha_k(I)$ as

$$\alpha_k(I) = \Pr[Z_k(I) = \mathbb{1}] = \Pr[X_k(I) = \mathbb{1} | X_{f(k)}(I) = \mathbb{1}] \quad (4)$$

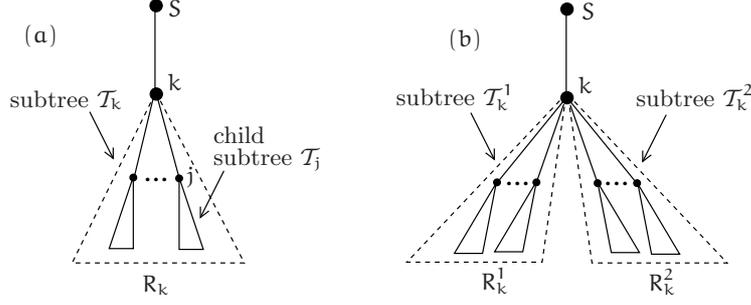


Fig. 2. The subtree \mathcal{T}_k with root at node k . (a): node k has a number of child subtrees rooted at the child nodes $\{d(k)\}$. (b): subtree partitioning of \mathcal{T}_k . The child subtrees of k are partitioned into two virtual subtrees $\mathcal{T}_k^1, \mathcal{T}_k^2$, anchored at k , with receivers R_k^1, R_k^2 .

where $\mathbb{1} = [1, \dots, 1]$. From Eq. (1), this holds for arbitrary I . The event $\{\mathbb{X}_k(I) = \mathbb{1}\}$ can be thought of as a *pattern* of probes, corresponding to the index set I , which survived to node k . By stationarity, $\alpha_k(I) = \alpha_k(i+I)$, where $i+I \equiv \{i+i_1, i+i_2, \dots, i+i_s\}$. E.g., $\alpha_k(\{i\}) = \alpha_k(\{1\})$, $\alpha_k(\{i, i+1\}) = \alpha_k(\{1, 2\})$.

Because of temporal dependence, $\alpha_k(I)$ is not in general $(\alpha_k\{i\})^s$, however thanks to spatial independence, $\forall k \in \mathcal{V}$ the *path pattern passage probability* $A_k(I)$ is simply the product of the corresponding link probabilities:

$$A_k(I) = \Pr[\mathbb{X}_k(I) = \mathbb{1}] = \alpha_k(I)A_{f(k)}(I) = \prod_{\ell \in a(k)} \alpha_\ell(I). \quad (5)$$

Since probes are sent from the root node, $A_0(I) = 1$.

Our aim is to estimate the link passage probability $\alpha_k(I)$, for arbitrary probe patterns I , for each link $k \in \mathcal{U}$. Eq. (5) shows they can be extracted if we know the $A_k(I)$ for all $k \in \mathcal{U}$. We therefore proceed to derive expressions for these.

4.2 Path Passage Probabilities: General Temporal Estimators

Consider a node $k \in \mathcal{U}$ as in the tree in Figure 2(a). It is the root of the subtree \mathcal{T}_k of \mathcal{T} which has receivers R_k . For a probe sent from the source, we define the following random variables and corresponding random vector:

$$Y_k(i) = \bigvee_{r \in R_k} X_r(i), \quad \mathbb{Y}_k(I) = [Y_k(i_1), \dots, Y_k(i_s)] \quad (6)$$

where \bigvee denotes bitwise OR. Hence $Y_k(i) = 1$ if there exists a receiver in R_k which receives the i -th probe (else 0). We now define, $\forall k \in \mathcal{U}$,

$$\gamma_k(I) = \Pr[\mathbb{Y}_k(I) = \mathbb{1}] \quad (7)$$

$$\beta_k(I, \mathbb{B}) = \Pr[\mathbb{Y}_k(I) = \mathbb{B} | \mathbb{X}_{f(k)}(I) = \mathbb{1}]. \quad (8)$$

Thus $\gamma_k(I)$ is the probability that, for each probe index in I , there is at least one receiver in R_k which receives it. On the other hand $\beta_k(I, \mathbb{B})$ is the probability that the receivers in R_k collectively see probe arrivals described by \mathbb{B} at indices I , conditional on all probes (at these indices) being present at node $f(k)$ (the definition is based on $f(k)$ rather than k to simplify expressions). In the special case of $\mathbb{B} = \mathbb{1}$, the following important relationship links \mathbf{A} , β , and γ :

$$\begin{aligned} \beta_k(I, \mathbb{1}) &= \Pr[Y_k(I) = \mathbb{1} | X_{f(k)}(I) = \mathbb{1}] = \frac{\Pr[Y_k(I) = \mathbb{1}, X_{f(k)}(I) = \mathbb{1}]}{\Pr[X_{f(k)}(I) = \mathbb{1}]} \\ &= \frac{\Pr[Y_k(I) = \mathbb{1}]}{\Pr[X_{f(k)}(I) = \mathbb{1}]} = \frac{\gamma_k(I)}{\mathbf{A}_{f(k)}(I)}. \end{aligned} \quad (9)$$

So far, the generalisation from single probes to probe patterns has been natural and remarkably straightforward. This is not true for the next two properties, which give expressions for $\beta_k(I, \mathbb{B})$ in the cases $\mathbb{B} = \mathbb{1}$ and $\mathbb{B} \neq \mathbb{1}$.

Property 1 (Parent-Child) The following relationship holds between β_k and the $\{\beta_j, j \in d(k)\}$ of the children of k . For convenience we relabel the children as $j = 1, 2 \dots d_k$ where $d_k = |d(k)|$. We have

$$\begin{aligned} \beta_k(I, \mathbb{1}) &= \alpha_k(I) \sum_{\substack{\{\mathbb{B}_1, \dots, \mathbb{B}_{d_k}\} \\ \text{s.t. } \bigvee_j \mathbb{B}_j = \mathbb{1}}} \prod_{j=1}^{d_k} \beta_j(I, \mathbb{B}_j) \\ &= \alpha_k(I) \left(1 - \prod_{j=1}^{d_k} (1 - \beta_j(I, \mathbb{1})) + \sum_{\substack{\{\mathbb{B}_1 \neq \mathbb{1}, \dots, \mathbb{B}_{d_k} \neq \mathbb{1}\} \\ \text{s.t. } \bigvee_j \mathbb{B}_j = \mathbb{1}}} \prod_{j=1}^{d_k} \beta_j(I, \mathbb{B}_j) \right) \end{aligned} \quad (10)$$

where we have first used the fact that an OR over all receivers can be decomposed as an OR over child subtrees, and then the property that such subtrees are conditionally independent. The ‘first’ $(1 - \prod_j)$ term in Eq. (10) corresponds to all child subtree receiver events $\{\mathbb{B}_j\}$ such that $\mathbb{B}_j = \mathbb{1}$ for at least one j . In the traditional case where $I = \{i\}$, this is the only term, since it is the same event as $\bigvee_j \mathbb{B}_j = \mathbb{1}$. It enables recursion over the tree, since $\beta(I, \mathbb{1})$ terms appear on both sides. In the temporal case this is not true, since which receivers see the probe can be different for each i in I . As a result, there are extra terms with $\beta_j(I, \mathbb{B}_j)$ factors with $\mathbb{B}_j \neq \mathbb{1}$.

Property 2 (Recursion over index sets with $\mathbb{B} = \mathbb{1}$) One can express $\beta_k(I, \mathbb{B} \neq \mathbb{1})$ in terms of $\beta_k(I', \mathbb{1})$, where $I' \subseteq I$. For instance, if $\mathbb{B} = [\mathbf{b}_1 =$

$0, \mathbf{b}_2, \dots, \mathbf{b}_s]$, and $I' = \{\mathbf{i}_2, \dots, \mathbf{i}_s\}$, $\mathbb{B}' = [\mathbf{b}_2, \dots, \mathbf{b}_s]$, then we have

$$\begin{aligned} \beta_{\mathbf{k}}(I, \mathbb{B}) &= \Pr[\Upsilon_{\mathbf{k}}(I) = \mathbb{B} | \mathbf{X}_{f(\mathbf{k})}(I) = \mathbb{1}] \\ &= \Pr[\Upsilon_{\mathbf{k}}(I') = \mathbb{B}' | \mathbf{X}_{f(\mathbf{k})}(I') = \mathbb{1}] - \Pr[\Upsilon_{\mathbf{k}}(I) = [\mathbf{1}, \mathbf{b}_2, \dots, \mathbf{b}_s] | \mathbf{X}_{f(\mathbf{k})}(I) = \mathbb{1}] \\ &= \beta_{\mathbf{k}}(I', \mathbb{B}') - \beta_{\mathbf{k}}(I, [\mathbf{1}, \mathbf{b}_2, \dots, \mathbf{b}_s]) \end{aligned} \quad (11)$$

which has eliminated the zero at \mathbf{i}_1 . We used the fact that $\Pr[\Upsilon_{\mathbf{k}}(I') = \mathbb{B}' | \mathbf{X}_{f(\mathbf{k})}(I) = \mathbb{1}] = \Pr[\Upsilon_{\mathbf{k}}(I') = \mathbb{B}' | \mathbf{X}_{f(\mathbf{k})}(I') = \mathbb{1}]$, which holds regardless of the index removed from I , thanks to Eq. (1). The above can be applied recursively to eliminate all zeroes, resulting in terms of the form $\beta_{\mathbf{k}}(I', \mathbb{1})$, $I' \subseteq I$, $|I| - z(\mathbb{B}) \leq |I'| \leq |I|$, where $z(\mathbb{B})$ denotes the number of zeroes in \mathbb{B} . In general

$$\beta_{\mathbf{k}}(I, \mathbb{B} \neq \mathbb{1}) = (-1)^{z(\mathbb{B})} \beta_{\mathbf{k}}(I, \mathbb{1}) + \delta_{\mathbf{k}}(I, \mathbb{B}), \quad (12)$$

where $\delta_{\mathbf{k}}(I, \mathbb{B})$ is the appropriate summation of $\beta_{\mathbf{k}}(I', \mathbb{1})$ for index sets $I' \subset I$. For example, if $I = \{1, 2, 3\}$, $\mathbb{B} = \{0, 0, 1\}$, then $\beta_{\mathbf{k}}(I, \mathbb{B}) = \beta_{\mathbf{k}}(I, \mathbb{1}) + (\beta_{\mathbf{k}}(\{3\}, \mathbb{1}) - \beta_{\mathbf{k}}(\{1, 3\}, \mathbb{1}) - \beta_{\mathbf{k}}(\{2, 3\}, \mathbb{1}))$. Eq. (12) can be used in (10) to remove all $\mathbb{B}_j \neq \mathbb{1}$ terms, leaving only terms of $\mathbb{B}_j = \mathbb{1}$ type. By applying (9) to each of these, we obtain

$$\begin{aligned} \frac{\gamma_{\mathbf{k}}(I)}{\mathbf{A}_{\mathbf{k}}(I)} &= 1 - \prod_{j=1}^{d_{\mathbf{k}}} \left(1 - \frac{\gamma_j(I)}{\mathbf{A}_{\mathbf{k}}(I)} \right) \\ &+ \sum_{\substack{\{\mathbb{B}_1 \neq \mathbb{1}, \dots, \mathbb{B}_d \neq \mathbb{1}\} \\ \text{s.t. } \forall_j \mathbb{B}_j = \mathbb{1}}} \prod_{j=1}^{d_{\mathbf{k}}} \left((-1)^{z(\mathbb{B}_j)} \frac{\gamma_j(I)}{\mathbf{A}_{\mathbf{k}}(I)} + \delta_j(I, \mathbb{B}_j) \right), \end{aligned} \quad (13)$$

an expression relating the observables $\gamma_{\mathbf{k}}(I)$ and $\{\gamma_j(I), j \in \mathbf{d}(\mathbf{k})\}$ to the desired path pattern probability $\mathbf{A}_{\mathbf{k}}(I)$. Note, $\forall \mathbf{k} \in \mathbf{R}$, $\mathbf{A}_{\mathbf{k}}(I) = \gamma_{\mathbf{k}}(I)$.

Recall that δ_j is a sum over terms involving index sets $I' \subset I$, and therefore **does not** contain $\mathbf{A}_{\mathbf{k}}(I)$. It is therefore easy to see that (13) is equivalent to a polynomial of degree $d_{\mathbf{k}}$ in $\mathbf{A}_{\mathbf{k}}(I)$ which can in principle be solved for $\mathbf{A}_{\mathbf{k}}(I)$. Eq. (13) is one of our main results. It is a full extension of the expression for path passage probability $\mathbf{A}_{\mathbf{k}}$ to a general temporal case $\mathbf{A}_{\mathbf{k}}(I)$.

The price of accessing the probability of a probe pattern I is the need to first calculate those for smaller patterns $I' \subset I$. This can be done recursively, beginning with $I = \{\mathbf{i}\}$ for which the sum in (13) vanishes, reducing it to the known polynomial expression in $\mathbf{A}_{\mathbf{k}}(\mathbf{i})$ of degree $d_{\mathbf{k}} - 1$ (see [3]).

Example: let $I = \{1, 2\}$ for a binary tree ($d_{\mathbf{k}} = 2, \forall \mathbf{k} \in \mathbf{U} \setminus \mathbf{R}$). Consider a branch node \mathbf{k} with children $j = 1, 2$. Let $I' = \{1\}$ and assume $\mathbf{a} = \mathbf{A}_{\mathbf{k}}(I')$ is already known. Let $p^{11} = \gamma_{\mathbf{k}}(I)$, $p^1 = \gamma_{\mathbf{k}}(I')$, and $p_j^{11} = \gamma_j(I)$, $p_j^1 = \gamma_j(I')$ for the children $j = 1, 2$. The events captured by the sum in (13) are $\{\mathbb{B}_1, \mathbb{B}_2\} =$

$\{\{1, 0\}, \{0, 1\}\}$, and $\{\{0, 1\}, \{1, 0\}\}$ and the quadratic polynomial in $\mathbf{A} = \mathbf{A}_k(\mathbf{I})$ is:

$$(2p_1^1 p_2^1) \mathbf{A}^2 + (\alpha^2(p_1^{11} + p_2^{11} - p^{11}) - 2\alpha(p_1^{11} p_2^1 + p_2^{11} p_1^1)) \mathbf{A} + \alpha^2 p_1^{11} p_2^{11} = 0. \quad (14)$$

4.3 Subtree Partitioning

In practice, solving polynomials arising from (13) implies numerical root finding for each node, and each $\mathbf{I}' \in \mathbf{I}$, which can be computationally slow. Our second main contribution of this paper is a computationally effective method for estimating the joint passage probabilities on trees of arbitrary order.

The idea is to combine child subtrees of node k into two sets, indexed by $j \in \{1, 2\}$, corresponding to two virtual subtrees \mathcal{T}_k^1 and \mathcal{T}_k^2 of \mathcal{T}_k with receivers \mathbf{R}_k^j (see Fig. 2(b)) (details of the allocation do not affect what follows). We therefore define additional quantities

$$\tilde{\mathbf{Y}}_k^j(\mathbf{i}) = \bigvee_{r \in \mathbf{R}_k^j} \mathbf{X}_r(\mathbf{i}), \quad \tilde{\mathbf{Y}}_k^j(\mathbf{I}) = [\tilde{\mathbf{Y}}_k^j(\mathbf{i}_1), \dots, \tilde{\mathbf{Y}}_k^j(\mathbf{i}_s)], \quad \tilde{\gamma}_k^j(\mathbf{I}) = \Pr[\tilde{\mathbf{Y}}_k^j(\mathbf{I}) = \mathbf{1}] \quad (15)$$

and the corresponding conditional probabilities:

$$\tilde{\beta}_k^j(\mathbf{I}, \mathbb{B}) = \Pr[\tilde{\mathbf{Y}}_k^j(\mathbf{I}) = \mathbb{B} | \mathbf{X}_k(\mathbf{I}) = \mathbf{1}] = \frac{\tilde{\gamma}_k^j(\mathbf{I})}{\mathbf{A}_k(\mathbf{I})}. \quad (16)$$

Since sets of disjoint child subtrees are still conditionally independent, β_k can be related to $\tilde{\beta}_k^1$ and $\tilde{\beta}_k^2$ (previously this was to $\{\beta_j, j \in \mathbf{d}(k)\}$). Then following the same sequence of steps as before, we get

$$\frac{\gamma_k(\mathbf{I})}{\mathbf{A}_k(\mathbf{I})} = 1 - \prod_{j=1}^2 \left(1 - \frac{\tilde{\gamma}_k^j(\mathbf{I})}{\mathbf{A}_k(\mathbf{I})} \right) + \sum_{\substack{(\mathbb{B}_1 \neq \mathbf{1}, \mathbb{B}_2 \neq \mathbf{1}) \\ \text{s.t. } \mathbb{B}_1 \vee \mathbb{B}_2 = \mathbf{1}}} \prod_{j=1}^2 \left((-1)^{z(\mathbb{B}_j)} \frac{\tilde{\gamma}_k^j(\mathbf{I})}{\mathbf{A}_k(\mathbf{I})} + \delta_k^j(\mathbf{I}, \mathbb{B}_j) \right) \quad (17)$$

This leads to a quadratic expression (or linear if $\mathbf{I} = \{\mathbf{i}\}$) in $\mathbf{A}_k(\mathbf{I})$ for all nodes irrespective of their degree, which can be solved explicitly. By making general trees appear binary in this way, numerical root finding is eliminated for numerous MINC estimators, including that in [3] for loss rates, and the more general temporal ones introduced here.

4.4 Simpler Temporal Estimators

The two previous estimators based on Eqs. (13) and (17) require recursive computations of $A_k(I')$ for smaller index sets, beginning with $I' = \{i\}$. We now introduce three alternative estimators which avoid this by carefully utilizing only a subset of receiver events that imply $X_k(I) = \mathbb{1}$. Consequently, they may yield higher variance but are easier to implement and computationally light.

We define:

$$Y_k^\vee(I) = \bigvee_{j \in d(k)} \mathbb{1}_{\{Y_j(I) = \mathbb{1}\}}, \quad \gamma_k^\vee(I) = \Pr[Y_k^\vee(I) = 1] \quad (18)$$

Here $Y_k^\vee(I) = 1$ iff there exists at least one child subtree of T_k where each probe at index in I is received by at least one of its receivers. The innovation is the use of the receiver event $\{Y_k^\vee(I) = 1\}$, rather than $\{Y_k(I) = \mathbb{1}\}$, to extract $A_k(I)$. The probability of this event is precisely that expressed by the first term of (13), and so there is no need to consider cases where $B \neq \mathbb{1}$. It follows that one simply has

$$\frac{\gamma_k^\vee(I)}{A_k(I)} = 1 - \prod_{j \in d(k)} \left(1 - \frac{\gamma_j(I)}{A_k(I)}\right). \quad (19)$$

This is formally identical to the first term in (13), which as we know leads to a polynomial of order $d_k - 1$ in $A_k(I)$.

One can combine this estimator with subtree partitioning, resulting in an equation which is linear for all I , with explicit solution

$$A_k(I) = \frac{\tilde{\gamma}_k^1(I)\tilde{\gamma}_k^2(I)}{\tilde{\gamma}_k^1(I) + \tilde{\gamma}_k^2(I) - \gamma_k^\vee(I)} \quad k \in \mathcal{U} \setminus \mathcal{R}, \quad A_k(I) = \gamma_k(I) \quad k \in \mathcal{R}. \quad (20)$$

Finally, one can construct an even simpler estimator, also with an explicit solution, by temporally extending the estimator of [9]. Let

$$Y_k^\wedge(I) = \bigwedge_{j \in d(k)} \mathbb{1}_{\{Y_j(I) = \mathbb{1}\}}, \quad \gamma_j^\wedge(I) = \Pr[Y_j^\wedge(I) = 1] \quad (21)$$

where \wedge is bitwise AND. Here $Y_j^\wedge(I) = 1$ iff for **each** child subtree of T_k , each probe at index in I is received by at least one of its receivers. Then we have,

$$A_k(I) = \left(\frac{\prod_{j \in d(k)} \gamma_j(I)}{\gamma_k^\wedge(I)} \right)^{1/(|d(k)|-1)}. \quad (22)$$

Even this estimator can be combined with subtree partitioning, however the result turns out to be the estimator of Eq. (20).

4.5 Estimator Names, Relations and Definitions

We name the five estimators of $A_k(I)$ based on above Eqs. as follows: GM :(13), GS :(17), SM :(19), SS :(20), SE :(22). For $I = \{i\}$, for arbitrary trees, estimators GM and SM are equivalent, as are GS and SS, resulting in three estimators which we simply call M, S, and E. For binary trees, for arbitrary I, GM and GS are equivalent, as are SM and SS. All five are equivalent for binary trees when $I = \{i\}$.

We now complete the estimator definitions. From data consisting of n trials, $\gamma_k(I)$, etc. for any I of interest, can be estimated using the empirical frequencies $\hat{\gamma}_k(I) = \sum_{i=0}^{n-s} Y_k(i+I)/(n-s+1)$. Next, the $\hat{\gamma}_k(I)$ are used to define an estimator $\hat{A}_k(I)$ for $A_k(I)$. In case of estimators with explicit solution, this is done by substituting into the relevant expression for $A_k(I)$, otherwise by finding the unique root (see [3]) in $[0, 1]$ of the polynomial numerically. Then, following Eq. (5), the joint link passage probabilities are estimated by setting $\hat{\alpha}_k(I) = \hat{A}_k(I)/\hat{A}_{f(k)}(I)$. Lastly, following Eq. (3), the mean loss run length μ_k is estimated as $\hat{\mu}_k = (1 - \hat{\alpha}_k)/(\hat{\alpha}_k - \hat{\alpha}_k[11])$, where $\alpha_k = \alpha_k(\{1\})$ and $\alpha_k[11] = \alpha_k(\{1, 2\})$.

5 Experiments

We illustrate the performance of the estimators using Monte Carlo simulations which obey our loss process assumptions from Section 2.1. We use two different families of loss processes: a 1-st order Markov chain to provide an example of a gentle departure from Bernoulli assumptions, and an On-Off process with geometric pass-runs and Zipf distributed loss-runs, to provide much stronger temporal dependence. We use a Zipf with a finite support over $[1, 1000]$, a distribution which has a power-law shaped tail over a wide range of values, but which has all moments finite. Each family is parameterised by two parameters: the marginal passage probability α_k , and the mean loss-run length μ_k .

We first use the general estimator GM and estimate: α_k (passage probability of a probe), $\alpha_k[11]$ (joint passage probability of a pair of consecutive probes), and μ_k (mean loss-run length) for each link k in the tree. Figure 3 provides benchmark results for a two receiver binary tree using $n = 10,000$ probes. The loss processes of the three links are identically distributed and as results for each are similar since the tree is small, we show results for shared link k only, and drop the subscript k .

Plot (a) in the figure shows curves of the relative error $(\hat{\alpha} - \alpha)/\alpha$ (this combines bias and variance) of α as a function of μ , using Markovian losses. The

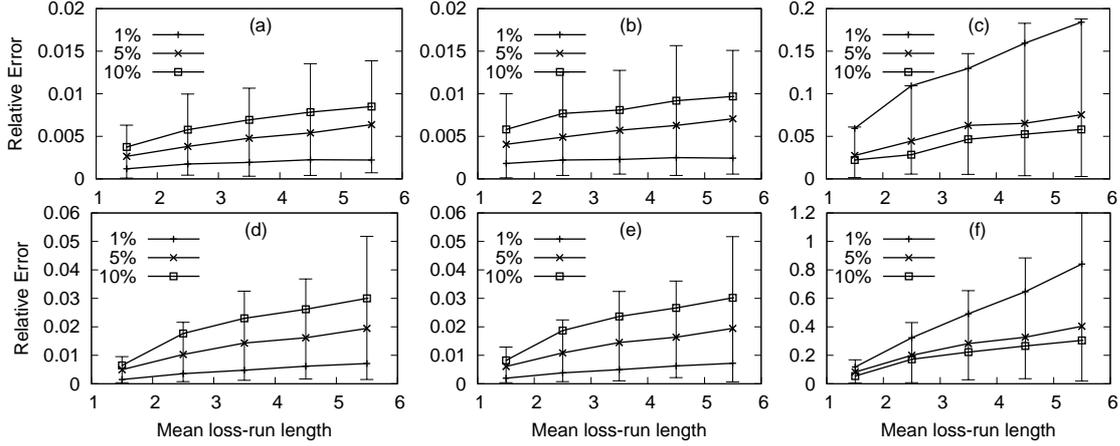


Fig. 3. Two-receiver tree. Relative error in the estimation of shared link loss parameters for 10k probes. Plots (a), (d): estimates of α , (b), (e): of $\alpha[11]$, (c), (f): of mean loss-run length μ .

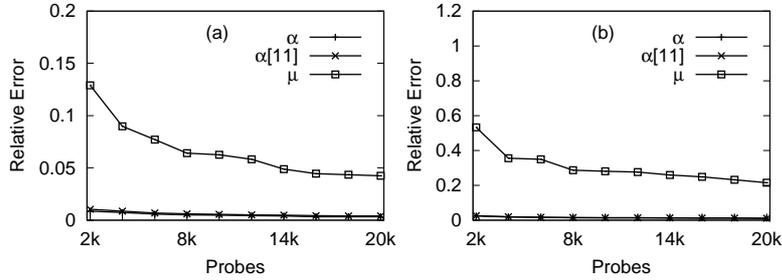


Fig. 4. Two-receiver tree. Relative error in the estimation of shared link loss parameters, variation with n for fixed α , μ .

errors are calculated by averaging over 100 independent replications (the grey confidence intervals gives an idea of the variation over these). Three curves are given corresponding to $\alpha = [0.01, 0.05, 0.1]$. We see that performance degrades both with increasing α and μ . Plot (d) tells a similar story for the On-Off case, but as expected with an increased error due to the high burstiness of the process.

The second column of Figure 3 reports on the estimation of $\alpha[11]$. The performance is quite satisfying, with errors being similar to those for α , despite the more challenging, temporal target. The third column gives errors in the estimation of μ . Because $\hat{\mu} = (1 - \hat{\alpha}) / (\hat{\alpha} - \hat{\alpha}[11])$, which involves a subtraction and a division of estimated quantities, we expect its variance to be greater than that of its components $\hat{\alpha}$ and $\hat{\alpha}[11]$, resulting in an increased error. This is indeed what we see. Finally, figure 4 fixes $\alpha = 0.05$ and $\mu = 3.5$, and shows how the relative error decreases with increasing n , confirming that we can recover even the more difficult μ given enough probes.

Figure 5 shows results for all 5 links in a binary tree with two levels and 3 receivers. We fix $\mu = 3$ for all links, and show how $\alpha[11]$ converges to the

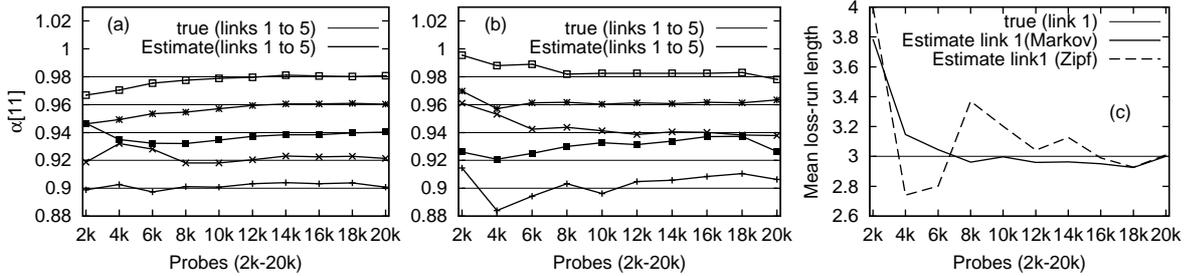


Fig. 5. Three-receiver binary tree with 5 heterogeneous links: Estimation of $\alpha[11]$ for all links: (a) Markov, (b) Zipf, (c) estimation of μ for shared link, (d) CDF comparison.

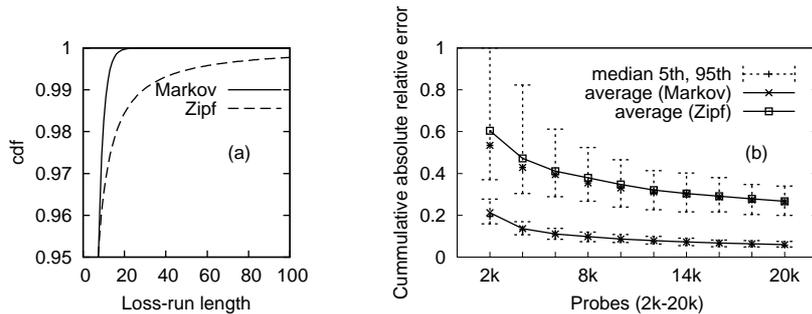


Fig. 6. (a): Comparison of CDFs for markovian and zipfian losses, (b): Estimation in general trees (non-binary real topologies from Rocketfuel).

correct value for each link as n increases, both for the Markov case (plot (a)) and the On-Off case (plot b). Plot (c) shows the slower convergence of $\hat{\mu}$ (shared link only).

Figure 6(a) gives the CDF of the loss-run distributions for Zipf and Markovian cases to indicate how ‘nasty’ the Zipf On-Off is compared to the Markov. In Figure 6(b), we show averaged results for 100 different 32 receiver logical trees, each built over the AT&T network topology collected from Rocketfuel [13]. The estimation error for μ (averaged over all links and all trees with a spread of link loss parameters: α uniformly over $[0.01, 0.05]$ and μ over $[1.5, 5.5]$) reduces as the number of probes n increases, though again slowly in the On-Off case.

In Figure 7, the variance of all five estimators is compared for a 3 receiver tertiary tree for Bernoulli, Markovian, and Zipf loss processes respectively. Across all three plots, we see that the standard errors for M and S are indistinguishable when estimating α . The same is true for GM and GS when estimating $\alpha[11]$. This shows that, for realistic loss rates, the use of subtree-partitioning allows computational efficiency without compromising on variance. On the other hand, E for α and the temporal SE for $\alpha[11]$ perform less well, although the difference becomes negligible for Zipf losses. More generally, as we move from independent Bernoulli losses to positively correlated Markov and Zipf losses, the difference between the general temporal estimators (GM, GS) which

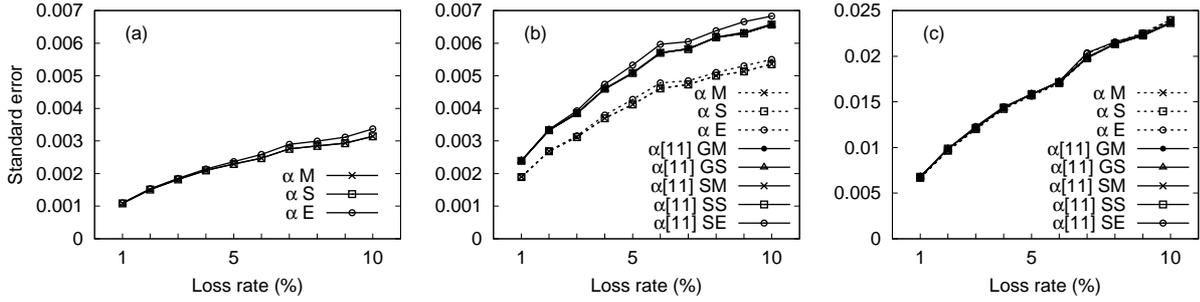


Fig. 7. Variance comparison of all estimators for Bernoulli, Markov and Zipf losses.

use all receiver events and the simpler ones (SM, SS, SE) which use only a subset, decreases. This can be understood by noting that the simpler estimators only miss events which are less likely to occur in positively correlated loss processes. This indicates that, even though estimation of general temporal characteristics is ambitious, in practice it can be performed almost as easily as estimation of average loss rates using the low-cost simpler temporal estimators, without significant variance penalty. For negatively correlated processes however, arguably of less practical relevance for the Internet, we expect differences between the general and simple estimators to be more marked. Finally, simulations confirm the expectation (plots omitted for space reasons), that the performance of subtree partitioning is enhanced as node degree increases, which is precisely when its computational advantages are most needed.

6 Analysis of Estimator Properties

In this section we determine some statistical properties of the above estimators.

Estimator Consistency All the estimators $\hat{A}_k(I)$ are constructed as follows. The true joint transmission probabilities $A_k(I)$ obey a relation $A_k(I) = g(\boldsymbol{\gamma}_k)$ where $\boldsymbol{\gamma}_k$ is a set of probabilities of leaf events, and g is a rational function continuous at $\boldsymbol{\gamma}_k$. The estimator $\hat{A}_k(I)$ is then $\hat{A}_k(I) = g(\hat{\boldsymbol{\gamma}}_k)$, where $\hat{\boldsymbol{\gamma}}_k$ are the empirical frequencies of the events that define $\boldsymbol{\gamma}_k$. Now by virtue of the stationary and ergodic assumption on the Z_k , $\hat{\boldsymbol{\gamma}}_k$ converges almost surely to $\boldsymbol{\gamma}_k$ as the number of probes n grows. Since g is continuous at $\boldsymbol{\gamma}_k$, $\hat{A}_k(I)$ converges almost surely to $A_k(I)$. Thus we have proved:

THEOREM If the Z_k are stationary and ergodic, the estimators $\hat{A}_k(I)$ are consistent, i.e., they converge almost surely to the $A_k(I)$ as $n \rightarrow \infty$.

Estimator Variance As well as consistency, it is important to understand estimator variability. We now outline how the asymptotic variance can be calculated. The complete solution will be left to future work.

We expect that as the number of probes $\mathbf{n} \rightarrow \infty$, estimator variance scales inversely with it, so that the rescaled error variable $\sqrt{\mathbf{n}}(\hat{\mathbf{A}}_k(\mathbf{I}) - \mathbf{A}_k(\mathbf{I}))$ converges in distribution to a Gaussian random variable with asymptotic variance σ^2 , allowing a comparison of estimator accuracy by comparing their asymptotic variances.

Our principal analytic tool is the δ -method (see [14], Ch. 7). Recall $\hat{\mathbf{A}}_k(\mathbf{I}) = \mathbf{g}(\hat{\boldsymbol{\gamma}}_k)$. Let $\hat{\boldsymbol{\gamma}}_k = (\hat{\gamma}_k^{(0)}, \dots, \hat{\gamma}_k^{(m)})$. By the central limit theorem, $\sqrt{\mathbf{n}}(\hat{\boldsymbol{\gamma}}_k - \boldsymbol{\gamma}_k)$ converges in distribution as $\mathbf{n} \rightarrow \infty$ to a Gaussian random variable with some covariance matrix \mathbf{C} where $\mathbf{C}_{ij} = \lim_{\mathbf{n} \rightarrow \infty} \mathbf{n}^{-1} \text{Cov}(\hat{\gamma}_k^{(i)}, \hat{\gamma}_k^{(j)})$. The δ -method then states that $\sqrt{\mathbf{n}}(\hat{\mathbf{A}}_k(\mathbf{I}) - \mathbf{A}_k(\mathbf{I}))$ is asymptotically Gaussian with variance $\sigma^2 = \sum_{ij} \mathbf{v}_i \mathbf{C}_{ij} \mathbf{v}_j$ where $\mathbf{v}_i = (\partial \mathbf{g}(\boldsymbol{\gamma}_k)) / \partial \gamma^{(i)}$. The \mathbf{v}_i are easily calculated, for example from differentiating the quadratic relation derived from (17) in the case of subtree partitioning. The \mathbf{C}_{ij} are more challenging since their computation involves both the covariances between the $\tilde{\mathbf{Y}}_k(\mathbf{I})$, $\tilde{\mathbf{Y}}_k^1(\mathbf{I})$ and $\tilde{\mathbf{Y}}_k^2(\mathbf{I})$ and the autocovariance functions amongst these over different times. The latter can be calculated given suitable assumptions, for example first order Markov, on the Z_k processes.

7 Conclusion

Packet traffic is bursty, and the performance of network applications depends on temporal metrics such as the duration of congestion events. In this paper we have shown how loss tomography techniques can be extended to include the estimation of temporal parameters. We show how arbitrary patterns of loss, and characterisations such as the mean loss-run length, can be recovered from each link, in a general, non-parametric context. We show that the estimators are consistent, and provide simulation illustrations of their relative errors over a range of parameters and realistic multicast tree topologies, using both Markovian and Zipfian On-Off link loss models.

We also provide a new technique to tackle general multicast trees in a computationally effective way: ‘subtree partitioning’, which effectively reduces a general tree to a binary one, for computational purposes. Simulation evidence suggests that the increase in variance due to the procedure is very small. The method is applicable both to traditional marginal loss rate estimation, and to the estimation of temporal characterisations of loss, and can also be used for delay tomography. We show how computation can also be significantly reduced by considering a simpler class of estimators which ignore certain events. By comparing against the general estimator, we show the resulting efficiency drop is negligible for reasonable loss rates.

In future work, the variance of the estimators presented here will be investi-

gated in more detail, including the impact of higher loss, higher node degree, larger and less homogeneous trees, as well as real traffic conditions.

References

- [1] AT&T Business Service Guide: AT&T VPN Service, See: <http://new.serviceguide.att.com/avpn.pdf> (February 2007).
- [2] Wikipedia, U-verse, <http://en.wikipedia.org/wiki/U-Verse> (February 2007).
- [3] R. Caceres, N. G. Duffield, J. Horowitz, D. Towsley, Multicast-Based Inference of Network-Internal Loss Characteristics, *IEEE Transactions on Information Theory*, 45 (1999) 2462–2480.
- [4] N. G. Duffield, V. Arya, R. Bellino, T. Friedman, J. Horowitz, D. Towsley, T. Turetti, Network tomography from aggregate loss reports, *Perform. Eval.* 62 (1-4).
- [5] F. Lo Presti, N. G. Duffield, J. Horowitz, D. Towsley, Multicast-based inference of network-internal delay distributions, *IEEE/ACM Transactions on Networking*, 10 (6) (2002) 761–775.
- [6] N. G. Duffield, J. Horowitz, F. Lo Presti, D. Towsley, Multicast topology inference from measured end-to-end loss, *IEEE Transactions on Information Theory*, 48 (1) (2002) 26–45.
- [7] Y. Gu, L. Breslau, N. G. Duffield, S. Sen, GRE Encapsulated Multicast Probing: A Scalable Technique for Measuring One Way Loss, in: *ACM Sigmetrics*, 2007.
- [8] J. Bolot, Characterising end-to-end packet delay and loss behavior in the internet, *Journal of High-Speed Networks*, 2 (3) (1993) 305–323.
- [9] N. G. Duffield, J. Horowitz, F. Lo Presti, D. Towsley, Explicit Loss Inference in Multicast Tomography, *IEEE Transactions on Information Theory*, (To appear).
- [10] N. G. Duffield, F. Lo Presti, V. Paxson, D. Towsley, Inferring link loss using striped unicast probes, in: *IEEE INFOCOM*, Anchorage, Alaska, 2001.
- [11] M. Coates, R. Nowak, Network loss inference using unicast end-to-end measurement, in: *ITC Seminar on IP Traffic, Measurement and Modeling*, 2000.
- [12] V. Jacobson, Congestion avoidance and control, in: *ACM SIGCOMM*, 1988.
- [13] N. Spring, R. Mahajan, D. Wetherall, T. Anderson, Measuring ISP topologies with Rocketfuel, *IEEE/ACM Transactions on Networking*, 12 (1) (2004) 2–16.
- [14] M. J. Schervish, *Theory of Statistics*, Springer, 1995.