

What Signals Do Packet-Pair Dispersions Carry?

Xiliang Liu[†], Kaliappa Ravindran[†], and Dmitri Loguinov[‡]

[†]The City University of New York, New York, NY 10016
xliu@gc.cuny.edu, ravi@cs.ccny.cuny.edu

[‡]Texas A&M University, College Station, TX 77843
dmitri@cs.tamu.edu

Abstract—Although packet-pair probing has been used as one of the primary mechanisms to measure bottleneck capacity, cross-traffic intensity, and available bandwidth of end-to-end Internet paths, there is still no conclusive answer as to what information about the path is contained in the output packet-pair dispersions and how it is encoded. In this paper, we address this issue by deriving closed-form expression of packet-pair dispersion in the context of a single-hop path and general bursty cross-traffic arrival. Under the assumptions of cross-traffic stationarity and ASTA sampling, we examine the statistical properties of the information encoded in inter-packet spacings and derive the asymptotic average of the output packet-pair dispersions as a closed-form function of the input dispersion. We show that this result is different from what was obtained in prior work using fluid cross-traffic models and that this discrepancy has a significant impact on the accuracy of packet-pair bandwidth estimation.

Index Terms—Stochastic Process, Queuing Theory, Active Measurement, Bandwidth Estimation.

I. INTRODUCTION

Sending probing packets to measure network path characteristics has been a common practice in the Internet since the late 1980's. There are two categories of probing-based measurements: delay based and dispersion based. In the first category, path characteristics such as per-hop capacity, queuing delay, and link utilization are inferred based on the RTT or one-way delay of *individual* packets [1], [6], [11], [14], [20]. In the second category, the dispersion of packet-pairs is traditionally used to infer bottleneck capacity [3], [4], [5], [7], [10], [13], [22], [23]; however, recent approaches also use packet-pairs/trains to measure cross-traffic and available bandwidth of an end-to-end path [2], [8], [9], [12], [18], [25], [26]. It is straightforward to identify the information encoded in the delay of individual probing packets. Hence, delay-based estimation techniques are theoretically validated and measurement difficulties are mostly due to practical issues [20], [24]. On the other hand, it is far more difficult to characterize the information contained in the output dispersion of probing packet-pairs. Consequently, apart from the practical issues, dispersion-based measurement techniques are yet to be fully justified for general cross-traffic conditions.

There has been a fair amount of research effort to characterize the information encoded in packet-pair spacing. However, previous analysis either relied on constant-rate fluid cross-traffic models [5], [17], or provided answers only partially suitable for generic bursty cross-traffic [3], [9], [21], [25]. In

this paper, we provide a more accurate, yet concise characterization of packet-pair probing in the context of a single-hop path and non-fluid cross-traffic. We identify three stochastic processes related to cross-traffic arrival and show that packet-pair probing essentially inspects the sample-paths of these three processes and constructs the output dispersion signal based on their random sampling. We derive several closed-form expressions to describe this construction procedure and call our characterization of packet-pair probing the “sampling and constructing” model.

Under the assumption of cross-traffic stationarity, we examine the statistical properties of the probing signals encoded in inter-packet spacing and derive the asymptotic average of the output dispersion as a closed-form function of the input dispersion. We show that the result deviates from what was previously obtained using constant-rate fluid cross-traffic and that this deviation has a significant adverse impact on packet-pair bandwidth estimation techniques.

We list the terminology used in the paper in Table I. In section II, we summarize related work of packet-pair analysis and report our earlier modeling attempt. In section III, we introduce our “sampling-and-constructing” model to characterize packet-pair probing. Based on our model, we examine the statistical characteristics of encoded probing signals in section IV. We derive the asymptotic average of output dispersions in section V and show that its deviation from the fluid result has an adverse impact on packet-pair bandwidth estimation in section VI. Finally, we present our concluding remarks in section VII.

II. BACKGROUND

A. Related Work

The earliest packet-pair analysis dated back to 1988, when Jacobson [10] examined the packet-pair spacing in the absence of cross-traffic and obtained the following result

$$\delta' = \begin{cases} \frac{s}{C} & \delta \leq \frac{s}{C} \\ \delta & \delta > \frac{s}{C} \end{cases}. \quad (1)$$

where δ and δ' are the input and output spacings of the packet-pair, respectively, s is the probing packet size, and C is the bottleneck capacity of the path. Note that when the input spacing δ is small, the output spacing δ' contains information about C .

In real networks, cross-traffic is often non-negligible. To take into account the effect of cross-traffic, Dovrolis *et al.* [5]

[‡]Supported by NSF grants CCR-0306246, ANI-0312461, CNS-0434940.

TABLE I
TERMINOLOGY

Term	Definition
Network path	Sequence of interconnected FIFO store and forward hops
Hop capacity	Transmission speed of the hop in bits per second
Bottleneck capacity	The minimum hop capacity along a network path
Narrow hop	The hop with the smallest capacity along a network path
Cross-traffic intensity	The average arrival rate of cross-traffic in some time interval
Hop available bandwidth	The hop's residual capacity after transmitting cross-traffic in some time interval
Path available bandwidth	The minimum hop available bandwidth along the path
Tight hop	The hop with the minimum available bandwidth along a path
Encoded probing signal	The information contained in packet-pair output dispersions

and Melander *et al.* [17] studied the relationship between the packet-pair input and output rate using a constant-rate fluid cross-traffic model. In a single-hop path, their results can be summarized as follows

$$r' = \begin{cases} C \frac{r}{r + \lambda} & r \geq C - \lambda \\ r & r < C - \lambda \end{cases}, \quad (2)$$

where $r = s/\delta$ and $r' = s/\delta'$ are the input and output rates of packet-pairs, respectively, λ is the cross-traffic intensity, and C is again the hop capacity. Translating (2) into its spacing version, we get

$$\delta' = \begin{cases} \frac{s}{C} + \frac{\lambda\delta}{C} & \delta \leq \frac{s}{C - \lambda} \\ \delta & \delta > \frac{s}{C - \lambda} \end{cases}. \quad (3)$$

This shows that when the input rate is higher than hop available bandwidth, there will be a deterministic multiplicative signal λ/C and a deterministic additive signal s/C encoded in the output spacing δ' . Again note that this result embraces (1) as a special case when $\lambda = 0$.

Applying mathematical induction to all hops along the path, we get the following relation between the input and output rate for an arbitrary multi-hop path

$$r' = \begin{cases} C \frac{r}{r + \lambda} & b \geq r \geq C - \lambda \\ r & r < C - \lambda \end{cases}, \quad (4)$$

where b is the second minimum hop-available-bandwidth along the path, C is the capacity of the tight hop, and λ is the cross-traffic intensity at the tight hop. This relation leads to the recent measurement proposal TOPP [18], [16], which is a technique to infer available bandwidth and tight-link capacity.

Realistic cross-traffic is always bursty and its intensity is never a time-invariant constant. Therefore, a natural question becomes how to generalize results (1)-(4) to accommodate bursty cross-traffic. We summarize the main results of previous studies next.

To interpret his Internet measurement observations of the probing packet RTT phase plot, Bolot (1993) [3] adopted a single-hop path with bursty cross-traffic in his analysis. Bolot showed that the packet-pair dispersion reflects the amount of traffic workload arrived at the router between the pair when the router does not idle between their arrivals.

Hu *et al.* (2003) [9] did a similar analysis and proposed a spacing formula under the condition when the packet-pair

share the same queuing period (the same condition used in Bolot's analysis)

$$\delta' = \frac{s}{C} + \frac{y\delta}{C}, \quad (5)$$

where y is a random variable reflecting the cross-traffic intensity in the duration between the arrivals of the probing pairs.

Both Bolot and Hu pointed out that when the input spacing δ is large enough, the output spacing δ' , although random, becomes equal to δ on average. In other words, the input spacing is only contaminated by some additive zero-mean signal.

Pasztor *et al.* (2002) [21] identified several types of signals encoded in the packet dispersions. They differentiated packet-pairs falling into the same router busy period from those that do not. In the former case, a multiplicative signal related to cross-traffic called *distribution signature* is encoded into the output spacings; while in the latter case, an additive zero-mean white noise signal called *accumulation signature* is encoded.

B. Our Earlier Attempt

We now report our earlier attempt of packet-pair probing modeling inspired by conventional wisdom. Although it turned out to be of little success, a comparison of this approach with the one presented later in this paper helps understand the problem better.

Based on previous insights, we tried to interpret packet-pair probing in single-hop path with general cross-traffic using the following signal model

$$\delta' = \begin{cases} \frac{s}{C} + \frac{y\delta}{C} & \text{JQ} \\ \delta + w & \text{DQ} \end{cases}, \quad (6)$$

where w is a zero-mean random variable, y is a λ -mean random variable, and λ is the long-term average of cross-traffic intensity. JQ (Joint Queuing) and DQ (Disjoint Queuing) are the conditions under which a packet-pair share the same router busy period or fall into different queuing period respectively.

This characterization can accommodate the fluid model as a special case where both random variables w and y degenerate to deterministic constants. Also note that, in the fluid model, a packet-pair falls into the same queuing period if and only if $\delta < s/(C - \lambda)$. In bursty cross-traffic, it is easy to verify that when $\delta < s/C$, the JQ condition is always satisfied. However, when $\delta > s/C$, there is usually no deterministic relationship between δ and the JQ condition. For any given $\delta > s/C$, JQ

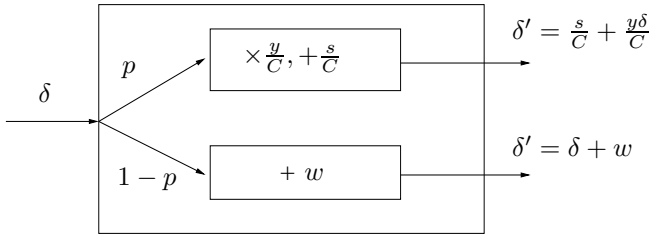


Fig. 1. Our unifying signal model inspired by previous work.

becomes a random event that occurs with certain probability. This can be schematically illustrated using a unifying signal model in Fig. 1.

The unifying model shows that the forwarding hop can be viewed as a stochastic mixture of two independent random systems. With probability p , the input signal δ passes through the JQ system, where multiplicative random signal y/C and additive deterministic signal s/C are stamped on δ . With probability $1 - p$, the input signal δ passes through the DQ system, where additive white noise signal w is stamped on δ . Although this characterization makes intuitive sense, we found that it is not accurate. We discuss the explanation in Section IV.

III. THE “SAMPLING AND CONSTRUCTING” NATURE OF PACKET-PAIR PROBING

A. Formulation of Cross-Traffic Arrival

In this paper, we focus on single-hop packet-pair probing. We assume infinite buffer capacity, FIFO queuing, and a work-conserving discipline for the forwarding hop. For the composite of cross-traffic and probing traffic, we assume simple traffic arrival, i.e., at most one packet arrives at any time instance. For cross-traffic alone, we identify three sample-paths which play crucial roles in determining the nature of packet-pair probing. They are the sample-paths of cross-traffic intensity process, hop workload-difference process, and available bandwidth process. We next present a rigorous formulation for these three elements and show the basic relationship among them.

Definition 1: Cross-traffic is driven by the packet counting process $\{N(t), 0 \leq t < \infty\}$ and the packet size process $\{S_n, 1 \leq n < \infty\}$. The cumulative traffic arrival $\{V(t), 0 \leq t < \infty\}$ is a random process counting the total volume of data¹ received by the hop up to time instance t

$$V(t) = \sum_{n=1}^{N(t)} S_n.$$

Note that $V(t)$ and $N(t)$ are right continuous, meaning that the packet arriving at t is counted in $V(t)$.

Definition 2: We define $\{Y_\delta(t), 0 \leq t < \infty\}$ as the process indicating the average cross-traffic arrival rate in the interval $(t, t + \delta]$

$$Y_\delta(t) = \frac{V(t + \delta) - V(t)}{\delta}.$$

and call it “ δ -interval cross-traffic intensity process”.

¹In this paper, packet size and data are measured in bits.

The second critical element is hop workload-difference process.

Definition 3: Hop workload process $\{W(t), 0 \leq t < \infty\}$ indicates the sum at time instance t of service times of all packets in the queue and the remaining service time of the packet in service.

Definition 4: We define $\{D_\delta(t), 0 \leq t < \infty\}$ as the process indicating the difference between the hop workload at time t and $t + \delta$

$$D_\delta(t) = W(t + \delta) - W(t),$$

and call it “ δ -interval workload-difference process”.

The third important process is the available bandwidth process.

Definition 5: Hop utilization process $\{U(t), 0 \leq t < \infty\}$ is an on-off process associated with $\{W(t)\}$

$$U(t) = \begin{cases} 1 & W(t) > 0 \\ 0 & W(t) = 0 \end{cases} \quad (7)$$

and δ -interval hop idle process

$$I(t, t + \delta) = I_\delta(t) = \delta - \int_t^{t+\delta} U(x) dx \quad (8)$$

is a process indicating the total amount of idle time of the forwarding hop in $[t, t + \delta]$. We further call time interval $[t, t + \delta]$ a “hop busy period” if $I_\delta(t) = 0$ and a “hop idle period” if $I(t + \delta) - I(t) = \delta$.

Definition 6: We define $\{B_\delta(t), 0 \leq t < \infty\}$ as the process indicating the residual bandwidth in the time interval $[t, t + \delta]$

$$B_\delta(t) = C \left(1 - \frac{1}{\delta} \int_t^{t+\delta} U(x) dx \right) = \frac{I_\delta(t)C}{\delta}, \quad (9)$$

and call it “ δ -interval available bandwidth process”.

The following theorem describes the relationship among the three important processes.

Theorem 1: For all positive t and δ , the following holds

$$B_\delta(t) = C - Y_\delta(t) + \frac{D_\delta(t)C}{\delta}. \quad (10)$$

Proof: Note that the total hop idle time within the time interval $[t, t + \delta]$ is

$$I_\delta(t) = \frac{B_\delta(t)\delta}{C}. \quad (11)$$

The amount of data transmitted by the hop within the time interval $[t, t + \delta]$ is

$$\begin{aligned} V(t + \delta) - V(t) + (W(t) - W(t + \delta))C \\ = Y_\delta(t)\delta - D_\delta(t)C. \end{aligned}$$

Thus, the hop working time is

$$\frac{Y_\delta(t)\delta}{C} - D_\delta(t). \quad (12)$$

Since δ is the sum of hop working time and hop idle time. Adding up (11) and (12), we get

$$\delta = \frac{B_\delta(t)\delta}{C} - D_\delta(t) + \frac{Y_\delta(t)\delta}{C}. \quad (13)$$

Rearranging (13), we get the desired result. ■

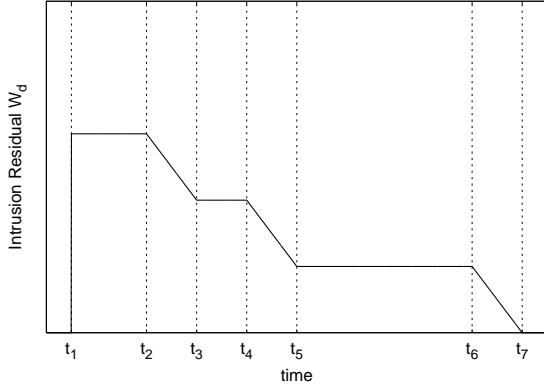


Fig. 2. Illustration of the intrusion residual function.

Packet-pair probing is essentially interacting with the *sample-paths* of the processes we just formulated. To prepare for the presentation of our main results, we next examine certain details about this interaction.

B. Probing Intrusion of Packet-Pairs

We use the triple $\langle a_1, \delta, s \rangle$ to denote a pair of probing packets p_1 and p_2 of the same size. The first element a_1 in the triple is the arrival time of the packet p_1 to the hop; δ is the inter-packet spacing; and s is the probing packet size. The arrival time of p_2 is $a_2 = a_1 + \delta$. The departure time of the probing packets from the hop are denoted by d_1 and d_2 . The output spacing is $\delta' = d_2 - d_1$. In terms of rate, the input and output probing rates are $r = s/\delta$ and $r' = s/\delta'$.

We use $\tilde{W}(t)$ and $\tilde{I}_\delta(t)$ to denote the workload sample-path and the hop idle sample-path associated with the superposition of cross-traffic and probing traffic. Note that traffic composition only increases hop workload. That is, for all t , $\tilde{W}(t) \geq W(t)$. Therefore, we define the following function to help understand this intrusion behavior of packet probing.

Definition 7: The *intrusive range* of the probing traffic into $W(t)$ is the set $\{t : \tilde{W}(t) > W(t)\}$. The *intrusion residual function* is $W_d(t) = \tilde{W}(t) - W(t)$.

Let us next examine the properties of function $W_d(t)$ to understand the intrusion behavior of a single probing packet. Before the arrival of the probing packet, $W_d(t) = 0$. It gets an immediate increment of s/C upon the packet arrival, where s is the packet size. In $W(t)$'s busy periods, $W_d(t)$ remains unchanged. In $W(t)$'s idle periods, $W_d(t)$ decreases linearly with slope -1 until it becomes 0, which marks the end of the intrusive range. Within the intrusive range, $W_d(t)$ is monotonically non-increasing. Fig. 2 illustrates this behavior, from which we can infer that (t_1, t_2) , (t_3, t_4) and (t_5, t_6) are three busy periods in $W(t)$, whereas (t_2, t_3) , (t_4, t_5) , and (t_6, t_7) are three idle periods in $W(t)$. Time instance t_1 is the arrival time of the probing packet, whereas t_7 marks the end point of the intrusive range².

When $W(t)$ is probed by a packet-pair $\langle a_1, \delta, s \rangle$, we are interested in the left-hand limit of $W_d(t)$ at time a_2 , denoted

as $R_\delta(a_1)$

$$\begin{aligned} R_\delta(a_1) &= \lim_{t \rightarrow (a_1 + \delta)^-} W_d(t) = W_d(a_2^-) \\ &= \tilde{W}(a_2^-) - W(a_2^-) = \tilde{W}(a_2^-) - W(a_2). \end{aligned}$$

The last equality is due to the simple arrival assumption. Since there is no cross-traffic packet arrival at time a_2 when p_2 arrives, we have $W(a_2^-) = W(a_2)$.

The term $R_\delta(a_1)$ is the intrusion residual at time a_2 caused by the probing packet p_1 and “experienced” by the packet p_2 . In other words, the queuing delay of p_2 in the hop is given by

$$\tilde{W}(a_2^-) = W(a_1 + \delta) + R_\delta(a_1). \quad (14)$$

As a direct result of the observation illustrated by Fig. 2, $R_\delta(a_1)$ can be computed as follows³

$$R_\delta(a_1) = \left(\frac{s}{C} - I_\delta(a_1) \right)^+ = \left(\frac{s - B_\delta(a_1)\delta}{C} \right)^+. \quad (15)$$

We are also interested in computing $\tilde{I}_\delta(a_1)$ when the hop is probed by packet-pair $\langle a_1, \delta, s \rangle$, which, from the intrusion behavior described in Fig. 2, can be expressed as following

$$\tilde{I}_\delta(a_1) = \left(I_\delta(a_1) - \frac{s}{C} \right)^+ = \left(\frac{B_\delta(a_1)\delta - s}{C} \right)^+. \quad (16)$$

Notice that between the two terms $R_\delta(a_1)$ and $\tilde{I}_\delta(a_1)$, there is at most one positive term for any given a_1 . When $R_\delta(a_1) > 0$, the two packets in the pair share the same hop busy period and $\tilde{I}_\delta(a_1) = 0$. When $\tilde{I}_\delta(a_1) > 0$, the two packets fall into different hop busy period and $R_\delta(a_1) = 0$. Hence, the positiveness of the two terms corresponds to JQ and DQ conditions respectively.

We are now ready to derive the relation between the input spacing δ and the output spacing δ' for any individual packet-pair. This relation is a milestone of our packet-pair analysis.

C. Output Packet-Pair Dispersion

We first present a corollary, which is due to the work-conserving assumption.

Corollary 1: For any packet arriving into the hop at time t_1 and departing from the hop at time t_2 , the time interval $[t_1, t_2]$ is a hop busy period.

This corollary immediately leads to the following lemma

Lemma 1: When a hop is probed by a packet-pair $\langle a_1, \delta, s \rangle$, we have $\tilde{I}(d_1, d_2) = \tilde{I}(a_1, a_2)$.

Proof: First, due to corollary 1, we have

$$\tilde{I}(a_1, d_1) = \tilde{I}(a_2, d_2) = 0. \quad (17)$$

Further, notice that $\tilde{I}(a_1, d_2)$ can be expressed in the following two ways

$$\tilde{I}(a_1, d_2) = \tilde{I}(a_1, d_1) + \tilde{I}(d_1, d_2) = \tilde{I}(d_1, d_2) \quad (18)$$

$$\tilde{I}(a_1, d_2) = \tilde{I}(a_1, a_2) + \tilde{I}(a_2, d_2) = \tilde{I}(a_1, a_2) \quad (19)$$

Combining (18) and (19), we have $\tilde{I}(d_1, d_2) = \tilde{I}(a_1, a_2)$. ■

³It is customary to denote $\max(X, 0)$ using X^+ and call it the “positive part” of X .

²Note that the probing packet departs before t_7 .

Our next theorem expresses the output spacing of a packet-pair from two different angles.

Theorem 2: When $W(t)$ is probed by a packet pair $\langle a_1, \delta, s \rangle$, the output spacing δ' can be expressed as

$$\begin{aligned} \delta' &= \frac{Y_\delta(a_1)\delta}{C} + \frac{s}{C} + \left(\frac{B_\delta(a_1)\delta - s}{C} \right)^+ \\ &= \delta + D_\delta(a_1) + \left(\frac{s - B_\delta(a_1)\delta}{C} \right)^+. \end{aligned} \quad (20)$$

Proof: We examine the hop activity with respect to $\tilde{W}(t)$ within the time interval $[d_1, d_2]$. Notice that s/C time units are spent on serving probing packet p_2 and that

$$\frac{V(a_2) - V(a_1)}{C} = \frac{Y_\delta(a_1)\delta}{C} \quad (21)$$

time units are spent on serving the cross-traffic that has arrived to the hop during the time interval (a_1, a_2) . Thus, the total hop working time in $[d_1, d_2]$ is given by

$$\frac{Y_\delta(a_1)\delta}{C} + \frac{s}{C}. \quad (22)$$

Also notice that $\tilde{I}(d_1, d_2)$ is the total idle time of the hop during this time interval. Since the sum of the hop working time in (22) and hop idle time must be equal to $d_2 - d_1$, we immediately have the following

$$\delta' = d_2 - d_1 = \frac{Y_\delta(a_1)\delta}{C} + \frac{s}{C} + \tilde{I}(d_1, d_2). \quad (23)$$

Further, due to Lemma 1 and (16), we get

$$\tilde{I}(d_1, d_2) = \tilde{I}(a_1, a_2) = \tilde{I}_\delta(a_1) = \left(\frac{B_\delta(a_1)\delta - s}{C} \right)^+. \quad (24)$$

Substituting (24) back to (23), we proved the first equality in (20). For the second part of (20), first notice that the total delays of p_1 and p_2 at the hop are given by

$$d_1 - a_1 = W(a_1) + \frac{s}{C} \quad (25)$$

$$d_2 - a_2 = R_\delta(a_1) + W(a_2) + \frac{s}{C}. \quad (26)$$

Subtracting (25) from (26), we get

$$\delta' = \delta + R_\delta(a_1) + D_\delta(a_1). \quad (27)$$

Substituting (15) into (27), we get the second half of (20). ■

The most salient feature of Theorem 2 is that the result is almost unconditional, in the sense that it neither relies on any assumption on cross-traffic arrival pattern nor imposes any restriction on the input signal δ . In addition, this result enforces such a conceptual idea that packet-pair probing can be viewed as a “sampling-and-constructing” procedure as illustrated in Fig. 3. The packet-pair $\langle a_1, \delta, s \rangle$ is essentially sampling the three sample-paths $Y_\delta(t)$, $D_\delta(t)$, and $B_\delta(t)$ at the time point a_1 and then constructing the output signal δ' using the three samples based on (20). Although (20) shows two different ways of constructing the output signal, they both produce the same result. We surely can take advantage of Theorem 1 and rewrite (20) in a form involving only two processes (e.g., $Y_\delta(t)$ and $D_\delta(t)$). However, the present version is more intuitive and makes later analysis easier. Our characterization already sheds

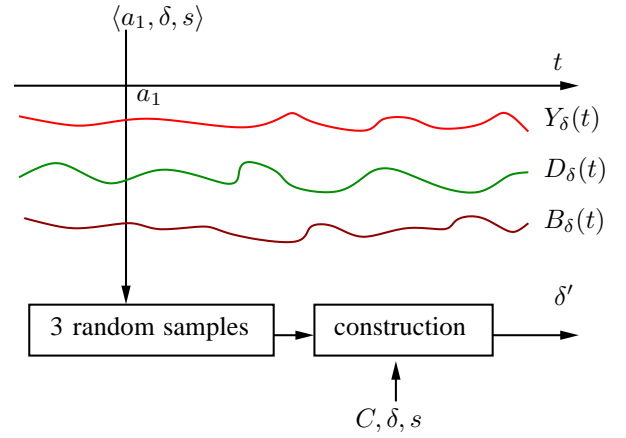


Fig. 3. The “sampling-and-constructing” nature of packet pair probing.

light on what the encoded probing signals are and how they are encoded. It also allows investigation of their statistical nature from an analytical angle rather than experimental observation.

IV. THE STATISTICS OF ENCODED PROBING SIGNALS

We use $\langle \{T_n, 1 \leq n < \infty\}, \delta, s \rangle$ to denote an infinite sequence of packet-pair probings driven by a point process $\Lambda(t) = \max\{n \geq 0 : T_n \leq t\}$. We use δ'_n to denote the output spacing of the n -th packet-pair $\langle T_n, \delta, s \rangle$ in the probing sequence. Adjacent packet-pairs are sufficiently separated, meaning that we neglect the cases where a packet-pair falls into the intrusive range of the preceding pairs. Consequently, the “sampling and constructing” model holds for all pairs in the probing sequence. This is a practically valid simplification because measurement tools [18], [26] all devise the inter-probing delays much larger than δ so as to keep the average probing traffic intensity small. Given this discussion, the packet-pair sequence will generate three discrete-time sample-paths of *encoded probing signal* (EPS) samples: $Y_\delta(T_n)$, $D_\delta(T_n)$, and $B_\delta(T_n)$. We now investigate the statistical properties of these EPS sample-paths.

A. Basics

We introduce a concept similar to probability distribution called *frequency distribution* to characterize sample-path statistics. For the details of this concept, please refer to [19, pages 46-50].

Definition 8: For continuous-time sample-path $X(t)$, define indicator function $\Psi(x, t)$

$$\Psi(x, t) = \begin{cases} 1 & X(t) \leq x \\ 0 & X(t) > x \end{cases}. \quad (28)$$

The frequency distribution function $P(x)$ of $X(t)$ is defined as follows (assuming the limit exists for $\forall x \in \mathbb{R}$)

$$P(x) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \Psi(x, t) dt. \quad (29)$$

For discrete-time sample-path X_n , define indicator function $\Psi(x, n)$

$$\Psi(x, n) = \begin{cases} 1 & X_n \leq x \\ 0 & X_n > x \end{cases}. \quad (30)$$

The frequency distribution function $P(x)$ of X_n is defined similarly

$$P(x) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=1}^k \Psi(x, n). \quad (31)$$

In the spirit of the probabilistic mean, we have continuous-time and discrete-time sample-path means as follows ⁴

$$\begin{aligned} E[X(t)] &= \int_{-\infty}^{\infty} x dP(x) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} X(t) dt, \\ E[X_n] &= \int_{-\infty}^{\infty} x dP(x) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k X_i. \end{aligned}$$

To examine the statistics of EPS sample-paths, we impose the another ASTA (Arrivals Sees Time Averages) property on the sampling process $\{T_n\}$. ASTA guarantees the equality of the statistics (or frequency distribution) in sampled series to the corresponding statistics in the continuous-time sample-path being sampled [15]. With non-negligible ASTA bias, sampling-based estimation usually fails to reach the measurement target and not much is left for discussion. The fact that quite a few current measurement techniques perform well without factoring ASTA into their design suggests that ASTA bias is either not present or is negligible in the studied Internet environments.

B. Cross-traffic Assumptions and Implications

Under the ASTA assumption, our focus is now shifted to the sample-path statistics of $Y_\delta(t)$, $D_\delta(t)$, and $B_\delta(t)$, which again are dependent on the probabilistic properties of the underlying cross-traffic arrival process. We make an ergodic stationarity assumption on cross-traffic arrival and investigate its implications on these sample-paths.

Assumption 1: The cumulative traffic arrival process $\{V(t)\}$ has ergodic stationary increments, i.e., for any positive δ , the process $\{Y_\delta(t)\}$ is an ergodic stationary process with ensemble mean λ , which is less than the hop capacity C .

This assumption imposes two restrictions on the process $\{Y_\delta(t)\}$. First, the stationarity assumption implies that $Y_\delta(t)$ has identical distribution for all t . Consequently, $\{Y_\delta(t)\}$ is also a wide sense stationary process ⁵. Second, the ergodicity assumption implies that, at any time instance t , the variance of $Y_\delta(t)$ decays to 0 as δ increases.

Due to Szczotka [27], [28], the hop workload process $\{W(t)\}$ will “inherit” the ergodic stationarity property from the traffic arrival process. Also because of the definition of workload-difference process and Theorem 1, this property is further carried over to process $\{D_\delta(t)\}$ and $\{B_\delta(t)\}$, whose ensemble means are 0 and $C - \lambda$ respectively.

The following lemma states the implications of Assumption 1 on the three sample-paths under investigation. They are all immediate consequences of ergodic stationarity.

Lemma 2: For any positive δ , the sample-path means of $Y_\delta(t)$, $D_\delta(t)$, and $B_\delta(t)$ are given by

$$E[Y_\delta(t)] = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} Y_\delta(t) dt = \lambda \quad (32)$$

$$E[D_\delta(t)] = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} D_\delta(t) dt = 0 \quad (33)$$

$$E[B_\delta(t)] = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} B_\delta(t) dt = C - \lambda. \quad (34)$$

Finally, ergodicity also implies that the variance of $B_\delta(t)$ decays when the observation interval δ becomes large and that this decaying variance will be reflected on the sample-path frequency distribution. Hence, we have the following lemma, which is intuitive and we skip the formal proof.

Lemma 3: Under the assumptions of the paper, the frequency distribution function $P_\delta(x)$ of sample-path $B_\delta(t)$ approaches the following step function as δ gets large

$$P_\infty(x) = \begin{cases} 0 & x < C - \lambda \\ 1 & x \geq C - \lambda \end{cases}. \quad (35)$$

We point out that the ergodic stationarity assumption is not a necessary condition for these sample-path properties. We conjecture that traffic with asymptotic stationarity also exhibits the same properties (including traffic driven by regenerative on-off arrival processes). However, the main goal of the paper is not to identify the weakest traffic condition that gives us the desirable sample-path properties, but rather to make a realistic traffic assumption and study its implications on bandwidth measurement techniques. A recent study showed that Internet traffic can be well modeled as a stationary process on the timescale of hours [29]. Therefore, we started by assuming traffic stationarity; however, later results in this paper are all derived based on the sample-path properties. They are expected to have broader applicability than those limited to stationary cross-traffic.

C. Unifying Signal Model Revisited

Thanks to ASTA, the three discrete-time sample-paths generated by packet-pair sampling have the following limiting time-averages

$$\begin{cases} E[Y_\delta(T_n)] = \lambda \\ E[D_\delta(T_n)] = 0 \\ E[B_\delta(T_n)] = C - \lambda \end{cases}. \quad (36)$$

Also recall that a large observation interval δ reduces the spread of the frequency distributions for $Y_\delta(T_n)$ and $B_\delta(T_n)$ ⁶. We now revisit the unifying model presented in Section II and explain why it is not valid.

Recall that when the input signal δ is very large, the distribution of $B_\delta(T_n)$ is concentrated around $C - \lambda$. The positive-part term $R_\delta(T_n)$ in the second equality of model (20) becomes almost constantly 0. Hence, the additive zero-mean term $D_\delta(T_n)$ becomes easily detected from the output signal samples δ'_n . However, this does not mean that there is

⁴We use $E[\cdot]$ to denote the sample-path mean, or limiting time-average, instead of the ensemble mean.

⁵Assuming second order moments exist.

⁶Unlike $B_\delta(T_n)$ and $Y_\delta(T_n)$, $D_\delta(T_n)$ is not a moving average process by nature. Consequently, the frequency distribution of $D_\delta(T_n)$ gradually becomes insensitive to the increase of δ .

no cross-traffic intensity information captured in δ'_n . The first equality of (20) shows that $Y_\delta(T_n)$ is always sampled by the packet-pair. It is only because of the strong noise term $\tilde{I}_\delta(T_n)$ that the desired signal is undetectable based on the observation of δ'_n .

When input signal $\delta < s/C$, the positive-part term $\tilde{I}_\delta(T_n)$ in the first equality of model (20) is 0. Hence, additive signal s/C and multiplicative λ -mean signal $Y_\delta(T_n)$ can be easily detected from the statistics of δ'_n . However, the additive zero-mean signal $D_\delta(T_n)$ does not escape packet-pair sampling. It only becomes undetectable from the statistics of δ'_n due to the strong noise term $R_\delta(T_n)$.

When δ is neither large nor small, both positive-part terms become prominent noise. Neither the additive zero-mean signal nor the multiplicative λ -mean signal can be easily detected. However, they are always captured by probing sampling and encoded in the output spacing δ' . This is exactly the subtlety that escaped the unifying model (6). The unifying model shows that δ'_n carries a sample of the zero-mean signal only under DQ condition (when $\tilde{I}(T_n) > 0$). However, the truth is that the output spacing always carries a sample of $D_\delta(T_n)$ regardless of the condition it is under. It is the *whole* set of samples $D_\delta(T_n)$ that exhibits additive zero-mean nature, *not* the set of samples under the DQ condition. Similarly, it is the whole set of samples $Y_\delta(T_n)$ that has a mean equal to λ , not just the set of samples under the JQ condition. This explains why the unifying signal model is not accurate.

V. ASYMPTOTICS IN PACKET-PAIR SAMPLING

One important application of our packet-pair probing model (20) is that it allows immediate derivation of the asymptotic average of probing output spacing δ' as a closed-form function of the input spacing δ . This functional relation, which we call the *gap response curve*, serves as a theoretical foundation for bandwidth measurement tools such as TOPP [18] and Spruce [26]. Previous derivations assumed constant-rate fluid cross-traffic and the results given in (2)-(3). In this section, we revisit this problem under the condition of ergodic stationary cross-traffic arrival, which is clearly more realistic than fluid traffic.

A. Asymptotic Average of δ'

Note that the output spacings also form a discrete-time sample-path δ'_n . The asymptotic average of δ' is just the sample-path mean of δ'_n . That is,

$$E[\delta'_n] = \lim_{n \rightarrow \infty} \sum_{i=1}^n \delta'_i. \quad (37)$$

Our next theorem provides a closed-form expression for $E[\delta'_n]$.

Theorem 3: When a hop is probed by ASTA packet-pair sequence $\{\{T_n\}, \delta, s\}$ and assuming that δ -interval available bandwidth sample-path $B_\delta(t)$ has frequency distribution function $P_\delta(x)$, then the following holds

$$\begin{aligned} E[\delta'_n] &= \frac{\delta\lambda + s}{C} + \int_{s/\delta}^C \frac{x\delta - s}{C} dP_\delta(x) \\ &= \delta + \int_0^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x). \end{aligned} \quad (38)$$

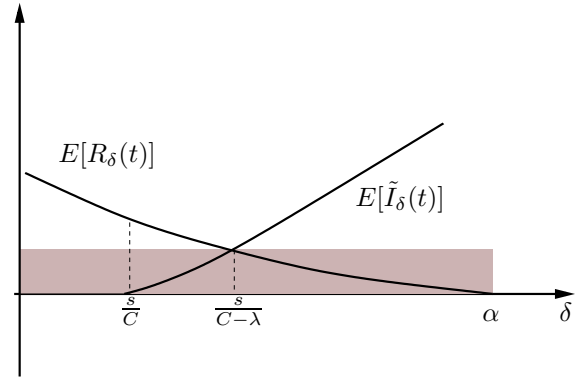


Fig. 4. The evolving trend of $E[\tilde{I}_\delta(t)]$ and $E[R_\delta(t)]$ with respect to δ while keeping s constant.

Proof: Recall Theorem 2 and notice that due to ASTA property, the frequency distribution of $B_\delta(T_n)$ is also $P_\delta(x)$. Also note that $B_\delta(T_n)$ is only distributed in $[0, C]$. Therefore,

$$\begin{aligned} E\left[\left(\frac{B_\delta(T_n)\delta - s}{C}\right)^+\right] &= \int_{s/\delta}^C \frac{x\delta - s}{C} dP_\delta(x) \\ E\left[\left(\frac{s - B_\delta(T_n)\delta}{C}\right)^+\right] &= \int_0^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x) \end{aligned}$$

Combining these results with (20) and (36), we have the statement of the theorem. ■

To help understand this result, we examine the two integral terms $E[\tilde{I}_\delta(t)]$ and $E[R_\delta(t)]$ in (38). Fig. 4 plots the evolving trend of the two terms with respect to δ while keeping s constant. As seen in the figure, $E[R_\delta(t)]$ shows a monotonically decreasing trend as δ increases and at some point α it becomes 0 or practically negligible. The other term $E[\tilde{I}_\delta(t)]$ remains 0 for $\delta \leq s/C$ and then shows a monotonically increasing trend, asymptotically approaching a linear function of δ . The two curves intersect at the point $\delta = s/(C - \lambda)$. These results can be summarized into the following two set of formulas

$$\begin{cases} E[\tilde{I}_\delta(t)] = 0 & \delta < \frac{s}{C} \\ E[\tilde{I}_\delta(t)] = E[R_\delta(t)] & \delta = \frac{s}{C - \lambda} \\ E[R_\delta(t)] = \frac{s}{C} - \frac{C - \lambda}{C}\delta & \delta < \frac{s}{C} \end{cases} \quad (39)$$

$$\begin{cases} \lim_{\delta \rightarrow \infty} E[\tilde{I}_\delta(t)] = \frac{C - \lambda}{C}\delta - \frac{s}{C} \\ \lim_{\delta \rightarrow \infty} E[R_\delta(t)] = 0 \\ \lim_{\delta \rightarrow s/C} E[\tilde{I}_\delta(t)] = 0 \end{cases} \quad (40)$$

All these properties are provable based on the lemmas we had in the previous section. We next prove one of them and leave the verification of the others to the reader.

Theorem 4: The term $E[R_\delta(t)]$ is a continuous and monotonically decreasing function of δ in the range $(0, \infty)$. It converges to 0 as δ increases.

Proof: First, note that for any $\Delta > 0$ and t , we have

$$0 \leq R_\delta(t) - R_{\delta+\Delta}(t) \leq \Delta. \quad (41)$$

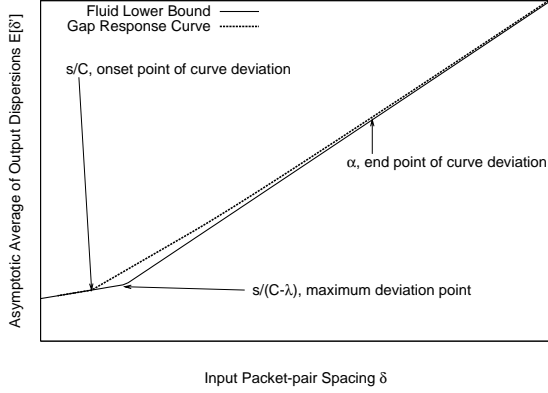


Fig. 5. Illustration of the gap response curve.

This difference defines a new sample-path, whose time-average can be computed as following

$$0 \leq E[R_\delta(t) - R_{\delta+\Delta}(t)] \leq \Delta, \quad (42)$$

which can be rewritten as

$$0 \leq E[R_\delta(t)] - E[R_{\delta+\Delta}(t)] \leq \Delta. \quad (43)$$

This shows that $E[R_\delta(t)]$ is a monotonic decreasing function of δ . Further, by taking the limit of (43) when $\Delta \rightarrow 0$, we have

$$\lim_{\Delta \rightarrow 0} (E[R_\delta(t)] - E[R_{\delta+\Delta}(t)]) = 0. \quad (44)$$

This proves the continuity of $E[R_\delta(t)]$ with respect to δ in the range $(0, \infty)$. Next, we show its convergence to 0 as δ increases.

First, note that

$$\begin{aligned} \lim_{\delta \rightarrow \infty} E[R_\delta(t)] &= \lim_{\delta \rightarrow \infty} \int_0^{s/\delta} \frac{s-x\delta}{C} dP_\delta(x) \\ &= \lim_{\delta \rightarrow \infty} \int_0^{s/\delta} \frac{s}{C} dP_\delta(x) - \lim_{\delta \rightarrow \infty} \int_0^{s/\delta} \frac{x\delta}{C} dP_\delta(x). \end{aligned} \quad (45)$$

Note that the first item in (45) is zero⁷

$$\lim_{\delta \rightarrow \infty} \int_0^{s/\delta} \frac{s}{C} dP_\delta(x) = \lim_{\delta \rightarrow \infty} \frac{s}{C} P_\delta\left(\frac{s}{\delta}\right) = \frac{s}{C} P_\infty(0) = 0, \quad (46)$$

and the second item in (45) is also zero

$$\begin{aligned} 0 &\leq \lim_{\delta \rightarrow \infty} \int_0^{s/\delta} \frac{x\delta}{C} dP_\delta(x) < \lim_{\delta \rightarrow \infty} \int_0^{s/\delta} \frac{s/\delta}{C} dP_\delta(x) \\ &= \lim_{\delta \rightarrow \infty} \frac{s}{C} P_\delta\left(\frac{s}{\delta}\right) = \frac{s}{C} P_\infty(0) = 0. \end{aligned} \quad (47)$$

Hence, the limit of $E[R_\delta(t)]$ when $\delta \rightarrow \infty$ is zero. ■

B. Deviation From Fluid Model

Note that the real gap response curve (38) is different from the fluid model (3). In fact, as schematically showed in Fig. 5, the fluid model is a lower bound of the real curve. In the input dispersion range $(s/C, \alpha)$, the real curve positively deviates

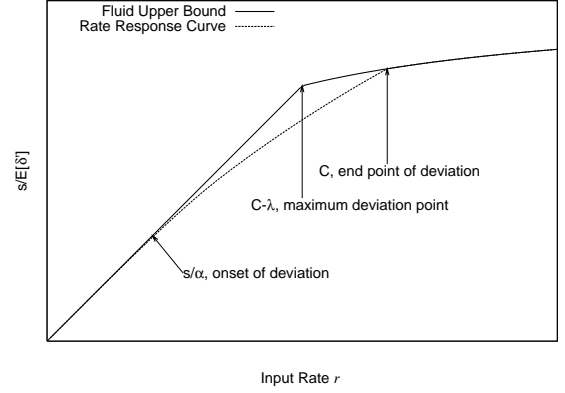


Fig. 6. Illustration of the rate response curve.

from this lower bound and reaches the maximum deviation at the point $\delta = s/(C - \lambda)$, where the input rate is equal to the available bandwidth. This deviation is also illustrated by the curve in the shadow area of Fig. 4 and can be expressed as

$$\begin{cases} \int_{s/\delta}^C \frac{x\delta - s}{C} dP_\delta(x) & \delta \leq \frac{s}{C - \lambda} \\ \int_0^{s/\delta} \frac{s - x\delta}{C} dP_\delta(x) & \delta \geq \frac{s}{C - \lambda} \end{cases}. \quad (48)$$

In contrast, note that using the unifying signal model (6), we can not foresee such a deviation phenomenon. Especially at the point $\delta = s/(C - \lambda)$ where the amount of deviation is maximized, the unifying model would, however, predict no deviation at all.

It helps to identify the exact value of α , which represents the end point of the deviation range. Note that α is the minimum input spacing that causes a zero intrusion residual

$$\begin{aligned} \alpha &= \inf\{\delta : E[R_\delta(t)] = 0\} \\ &= \inf\{\delta : P_\delta\left(\frac{s}{\delta}\right) = 0\}. \end{aligned} \quad (49)$$

This requires that $B_\delta(t)$ be greater than the input probing rate s/δ for almost every t along the time axis. In other words, the input rate s/δ must be smaller than the distribution lower bound of δ -interval available bandwidth. It is often not possible to satisfy such a condition exactly⁸, since the convergence of $P_\delta(x)$ to the step function in (35) might only be asymptotic and $P_\delta(x)$ may remain positive for all $x \in (0, C]$ regardless of the observation interval δ . In that case, $\alpha = \infty$ and we can only mark it approximately as a point where the deviation becomes practically negligible.

It is often more informative to look at the rate version of the response curve rather than the spacing version, because the rate response curve has a direct association with cross-traffic intensity and available bandwidth. Plotting $s/E[\delta'_n]$ with respect to $r = s/\delta$ and comparing it with the fluid model (3), we get Fig. 6. The fluid rate model becomes an upper bound of the real response curve. The input rate range $(s/\alpha, C)$ becomes the area where the real curve negatively deviates from the fluid model.

⁷Recall that P_∞ is the step function given in (35).

⁸For instance, when cross-traffic arrival is Poisson.

C. Impact of Probing Packet Size

How does the packet size s affect the response curves or the amount of deviation from fluid models? We first consider the case when input rate $r < C - \lambda$. Notice that the second expression in (48) can be simplified to the following form

$$\begin{aligned} \int_0^{s/\delta} \frac{s-x\delta}{C} dP_\delta(x) &= \int_0^r \frac{r\delta-x\delta}{C} dP_\delta(x) \\ &= \frac{\delta}{C} \left(r \int_0^r dP_\delta(x) - \int_0^r x dP_\delta(x) \right). \end{aligned} \quad (50)$$

Applying integration by parts, we get

$$\int_0^r x dP_\delta(x) = rP_\delta(r) - \int_0^r P_\delta(x) dx. \quad (51)$$

Substituting (51) back to (50), we get

$$\int_0^{s/\delta} \frac{s-x\delta}{C} dP_\delta(x) = \frac{\delta}{C} \int_0^r P_\delta(x) dx. \quad (52)$$

When $s \rightarrow \infty$ while keeping r constant, the sampling interval $\delta = s/r$ approaches infinity proportionally. Hence, we have

$$\lim_{s \rightarrow \infty} \int_0^{s/\delta} \frac{s-x\delta}{C} dP_\delta(x) = \lim_{\delta \rightarrow \infty} \frac{\delta}{C} \int_0^r P_\delta(x) dx. \quad (53)$$

Dropping the constant $1/C$, we get a sufficient and necessary condition for the amount of deviation at input rate $r < C - \lambda$ to vanish when $s \rightarrow \infty$

$$\lim_{\delta \rightarrow \infty} \delta \int_0^r P_\delta(x) dx = 0. \quad (54)$$

Similarly, for any input rate $r > A$, a sufficient and necessary condition for the deviation to vanish is

$$\lim_{\delta \rightarrow \infty} \delta \left(C - r - \int_r^C P_\delta(x) dx \right) = 0. \quad (55)$$

These conditions require the sample-path distribution of $B_\delta(t)$ not only to exhibit decaying variance when the observation interval δ becomes large, but also to show sufficient decaying speed. Our experimental observation so far shows that these properties are usually satisfied. The problem of identifying cross-traffic types in which (54) or (55) is violated remains open.

Even though large probing packet size usually implies less deviation of the real response curves from the fluid models, we point out the existence of such interesting cases where certain probing packet size can lead to a response curve with no derivation at all. This happens only when cross-traffic exhibits periodic arrival pattern. Assuming the workload sample-path $W(t)$ is a periodic function that repeats itself every T time units, then setting probing packet size to $T(C - \lambda)$ causes the real response curves to coincide with the fluid models.

In practice, cross-traffic is rarely periodic and we have to rely on large packet size to reduce the amount of response curve deviation. However, due to the limit of path MTU and the concern of packet fragmentation, probing packet size s can not be made arbitrarily large. The question becomes whether the commonly used MTU of 1500 bytes is enough to reduce the amount of curve deviation to such an extent that its impact

on bandwidth measurement accuracy becomes insignificant? In what follows, we show that even for relatively ‘‘smooth’’ (e.g., Poisson) cross-traffic, the deviation phenomenon still can have significant adverse impact on bandwidth estimation.

VI. THE IMPACT OF RESPONSE CURVE DEVIATION ON BANDWIDTH MEASUREMENT

A. Computing Response Curves

We devise an off-line algorithm to compute the single-hop response curves based on cross-traffic packet arrival trace, the probing packet size s , and the hop capacity C . The trace file provides information regarding the arrival time and packet size for every cross-traffic packet. Given a trace file with sufficiently long time interval recorded, the frequency distributions of the associated sample-paths (such as $Y_\delta(t)$, $D_\delta(t)$, and $B_\delta(t)$) in that finite time interval become good approximations of their limiting frequency distributions. Our off-line algorithm approximates the sample-path mean of δ'_n for any given input spacing δ . Next, we briefly explain the spirit of this algorithm.

We use Υ_0^τ to denote a cross-traffic trace in the time interval $[0, \tau]$. Given Υ_0^τ and hop capacity C , the hop workload sample-path $W(t)$ in the interval $[0, \tau]$, denoted as $W(t)_0^\tau$, can be computed. The following corollary states the function properties of workload sample-path.

Corollary 2: Hop workload sample-path consists of alternating busy periods and idle periods. Any busy period comprises piece-wise linear segments with slope -1 .

Taking advantage of these functional properties and using a proper data structure, we can represent $W(t)_0^\tau$ without losing any information about the original process. From $W(t)_0^\tau$, we are able to retrieve $Y_\delta(t)$, $D_\delta(t)$ and $B_\delta(t)$ for any t in $[0, \tau - \delta]$. In other words, we keep the full information about $Y_\delta(t)_0^{\tau-\delta}$, $D_\delta(t)_0^{\tau-\delta}$, and $B_\delta(t)_0^{\tau-\delta}$ in the data structure of $W(t)_0^\tau$.

Instead of approximating $E[\delta'_n]$ using a finite number of output dispersion samples, we approximate $E[\delta'(t)]$, the corresponding continuous-time sample-path mean, using the time average of $\delta'(t)$ in a finite interval. Note that due to the ASTA assumption, $E[\delta'_n] = E[\delta'(t)]$. Hence, a good approximation of the latter sample-path mean also serves as a good approximation for the former. The continuous-time sample path $\delta'(t)$ also has certain ‘‘nice’’ properties as we state in the next theorem. The proof is in constructive terms, which provides a concrete idea of how our off-line algorithm is designed.

Definition 9: *Event-points* are the time instances at which the workload sample path switches from a busy period to an idle period or undergoes a sudden increment due to packet arrival. An *epoch* is a time interval between two adjacent event-points.

Theorem 5: The sample-path $\delta'(t)$ consists of piece-wise linear segments with possible slopes 0, 1 and -1 . For any two time instances $0 < t_1 < t_2$, $\delta'(t)$ is continuous in the interval (t_1, t_2) given that t_1 and t_2 fall into the same epoch of $W(t)$, and that $t_1 + \delta$ and $t_2 + \delta$ also fall into the same epoch of $W(t)$.

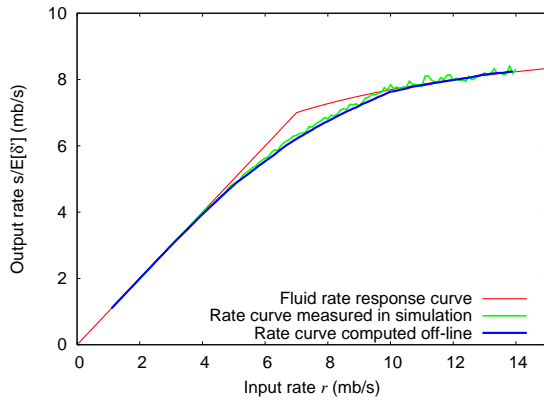


Fig. 7. A comparison between the fluid curve of prior work, the curve computed offline, and the one measured in simulations.

Proof: See the Appendix. ■

Our algorithm computes $\delta'(t)_0^{\tau-\delta}$, the sample-path $\delta'(t)$ in time interval $[0, \tau - \delta]$, based on Υ_0^{τ} , C , s and δ . Taking advantage of Theorem 5, we can represent the sample-path information of $\delta'(t)_0^{\tau-\delta}$ to its full precision in a proper data structure. We then compute the following⁹

$$\overline{\delta'(t)_0^{\tau-\delta}} = \frac{1}{\tau - \delta} \int_0^{\tau-\delta} \delta'(t) dt, \quad (56)$$

and use it as an approximation to

$$E[\delta'(t)] = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} \delta'(u) du. \quad (57)$$

It is clear that the precision of this approximation is mainly decided by τ . Thus, we can pick a large τ so that further increase of τ would make little difference. It could be sometimes impractical to have such a long trace. However, note that even when (56) is not a good approximation of $E[\delta'(t)]$, it still represents the correct result in a hypothetical periodic cross-traffic that repeats itself after every τ time units. This is due to the fact that in periodic cross-traffic, sample-path $\delta'(t)$ has a limiting time-average equal to its time average in one period.

In our experiment, we use a single-hop path with capacity $C = 10$ mb/s. We use Poisson cross-traffic with average arrival rate of 500 packets per second. The packet size is 750 bytes. Hence, $\lambda = 3.0$ mb/s. The probing packet size is chosen to be 1500 bytes. We compute $E[\delta'(t)]$ for 130 input values of δ in $[0.86 \text{ ms}, 12 \text{ ms}]$, which corresponds to 130 equally spaced input rates in the range of $[1.0 \text{ mb/s}, 14 \text{ mb/s}]$. We generate a 100-second cross-traffic trace and find that the time average of $\delta'(t)$ in time interval $[0, 30\text{s}]$ has sufficiently converged. Even for the smallest input $\delta = 0.86$ ms, doubling the trace interval to 60 s does not produce more than 1% difference.

To validate our algorithm, we also use the same setting in NS2 simulation to measure the response curve. For every input rate, the sender transmits 1000 packet-pairs. The inter-probing delay is controlled by an exponentially distributed random variable to meet the ASTA sampling condition. The average

⁹Theorem 5 also allows an efficient computation of (56) with high accuracy.

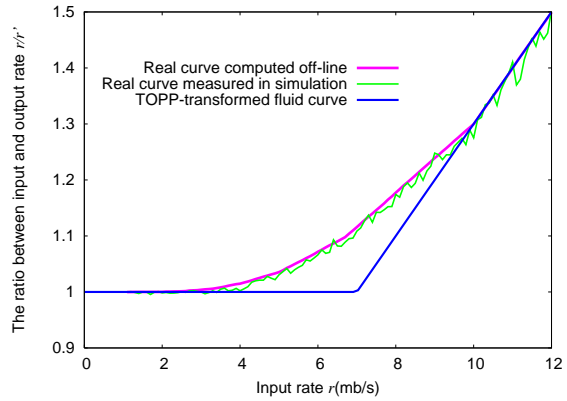


Fig. 8. Transformed curves used by TOPP.

inter-probing delay is chosen to be 10δ . We use the average of the 1000 output dispersions to approximate $E[\delta'_n]$.

Fig. 7 shows both the rate curve computed off-line and the rate curve measured in NS2 simulation. The figure shows that the two agree and non-negligibly deviate from the fluid upper bound. The deviation is very clear even though we use the largest probing packet size. The figure also shows that we can obtain a much smoother curve using off-line computation than using NS2 simulation.

Next, we discuss the implications of our findings on two packet-pair bandwidth measurement techniques: TOPP and Spruce.

B. TOPP

We first consider a single-hop path. TOPP uses the fluid rate model (2) as its measurement rationale. It first collects the output rates $\bar{r}' = s/\bar{\delta}'$ ¹⁰ for a series of equally spaced input rates in some interval $[r^{min}, r^{max}]$. TOPP then transforms the measured rate response curve to a function between r/\bar{r}' and r , which admits the following piece-wise linear relation in fluid models:

$$r/\bar{r}' = \begin{cases} 1 & r < C - \lambda \\ \frac{r + \lambda}{C} & r \geq C - \lambda \end{cases}. \quad (58)$$

TOPP identifies the second segment in the curve and applies linear regression to calculate the capacity C and cross-traffic intensity λ .

We transform both the computed and the measured rate response curves in Fig. 7 to the form of (58). As showed in Fig. 8, the deviation ranges appear as the second segments. When applying linear regression on these segments, we get estimation results in Table II. The table shows that even if TOPP could manage to get the asymptotic rate curve, it would not achieve an accurate measurement due to its unawareness of the deviation phenomenon.

In a single-hop case, note that the real curve agrees with fluid model when $r \geq C$. Therefore, linear regression can be applied to this curve portion to extract the capacity C

¹⁰We use $\bar{\delta}'$ to denote a measurement of $E[\delta']$ using finite number of samples of δ' .

TABLE II
TOPP ESTIMATION RESULTS (IN MB/S)

	C	λ	$C - \lambda$
Real values	10	3.0	7.0
TOPP in ns2 simulation	35.97	32.33	3.64
TOPP in off-line computation	35.81	32.38	3.43

and cross-traffic intensity λ . In a multi-hop path, unlike what can be easily derived from the fluid model as showed in (4), the real curve becomes very complex and is not amenable to closed-form analysis. However, based on the single-hop results, we can predict that one necessary condition for the curve portion that contains the information of λ and C to survive against multi-hop interference is that the available bandwidth for all pre-tight links *must be significantly higher than the tight link capacity*. Otherwise, when the input rate is greater than C , it falls into the deviation range of some pre-tight link and the probing rate gets reduced on average before the packet-pairs arrive at the tight hop.

C. Spruce

Spruce assumes a single bottleneck link whose capacity C can be estimated beforehand. Spruce sends probing pairs with intra-pair spacing δ set to s/C , which is the bottleneck link transmission delay of the probing packet. Inter-pair delay is controlled by an exponentially distributed random variable to meet ASTA property. Each probing pair generates an available bandwidth estimate by computing

$$A_n = C \left(1 - \frac{\delta'_n - \delta}{\delta} \right). \quad (59)$$

Spruce averages the last 100 samples to obtain a final estimate for the path available bandwidth. Our next theorem states the unbiasedness of Spruce estimator

Theorem 6: Under the assumptions of this paper, in a single-hop path, we have

$$E \left[C \left(1 - \frac{\delta'_n - \delta}{\delta} \right) \right] = C - \lambda. \quad (60)$$

Proof: Combining (38) and the first equality in (39), this theorem is easily proved. ■

In a multi-hop path, the pre-bottleneck links introduce noise to the input signal δ , which shifts it from a constant to a random variable. The post-bottleneck links introduce noise to δ' . This impact can be minimized if both noise signals are zero-mean, which happens when the non-bottleneck hops have available bandwidth significant higher than the bottleneck capacity. A detailed analysis of spruce's robustness to multi-hop queueing interference requires future work, for which the insights of this paper will be helpful.

VII. CONCLUDING REMARKS

In this paper, we identified three important sample-paths related to cross-traffic arrival and established the "sampling-and-constructing" model to characterize packet-pair probing. This approach uncovers the full picture of encoded probing signals and leads to a closed-form solution to the asymptotic

average of output packet-pair dispersions in a single-hop path, which extends previous fluid models and serves as a theoretical foundation for packet-pair bandwidth estimation. In our future work, we will apply these results to analyze the robustness of current available bandwidth measurement techniques to multi-hop queueing interference.

REFERENCES

- [1] K. G. Anagnostakis, M. B. Greenwald, and R. S. Ryger, "cing: Measuring Network-Internal Delays using only Existing Infrastructure," *IEEE INFOCOM*, April 2003.
- [2] S. Banerjee and A. Agrawala, "Estimating Available Capacity of a Network Connection," *IEEE International Conference on Networks*, Sept. 2001.
- [3] J. Bolot, "Characterizing end-to-end packet delay and loss in the internet," *ACM SIGCOMM*, 1993.
- [4] R. Carter and M. Crovella, "Measuring Bottleneck Link Speed in Packet-Switched Networks," *Internation Journal on Performance Evaluation*, 2728, 1996.
- [5] C. Dovrolis, P. Ramanathan, and D. Moore, "What Do Packet Dispersion Techniques Measure?," *IEEE INFOCOM*, April 2001.
- [6] A. B. Downey, "Using Pathchar to Estimate Internet Link Characteristics," *ACM SIGCOMM*, 1999.
- [7] K. Harfoush, A. Bestavros, and J. Byers, "Measuring Bottleneck Bandwidth of Targeted Path Segments," *INFOCOM*, March 2003.
- [8] G. He and J. Hou, "On Exploiting Long Range Dependence of Network Traffic in Measuring Cross Traffic on an End-to-end Basis," *IEEE INFOCOM*, March 2003.
- [9] N. Hu and P. Steenkiste, "Evaluation and Characterization of Available Bandwidth Probing Techniques," *IEEE JSAC*, August 2003.
- [10] V. Jacobson, "Congestion Avoidance and Control," *ACM SIGCOMM*, 1988.
- [11] V. Jacobson, "Pathchar – A Tool to Infer Characteristics of Internet Paths," 1997. Slides available from ftp://ftp.ee.bl.gov/pathchar/msri-talk.pdf.
- [12] M. Jain and C. Dovrolis, "End-to-end available bandwidth: measurement methodology, dynamics, and relation with TCP throughput," *ACM SIGCOMM*, August 2002.
- [13] K. Lai and M. Baker, "Measuring Bandwidth," *IEEE INFOCOM*, 1999.
- [14] K. Lai and M. Baker, "Measuring link bandwidths using a deterministic model of packet delay," *ACM SIGCOMM*, 2000.
- [15] B. Melamed and D. Yao, "The ASTA property," *Advances in Queueing: Theory, Methods and Open Problems*, pages 195–224, 1995.
- [16] B. Melander, M. Bjorkman, and P. Gunningberg, "A New End-to-End Probing and Analysis Method for Estimating Bandwidth Bottlenecks," *IEEE Globecom Global Internet Symposium*, November 2000.
- [17] B. Melander, M. Bjorkman, and P. Gunningberg, "First-Come-First-Served Packet Dispersion and Implications for TCP," *IEEE GLOBECOM*, 2002.
- [18] B. Melander, M. Bjorkman, and P. Gunningberg, "Regression-Based Available Bandwidth Measurements," *SPECTS*, July 2002.
- [19] E. Muhammad and S. Stidham. *Sample-path Analysis of Queueing Systems*. Kluwer Academic Publishers, 1999.
- [20] A. Pasztor and D. Veitch, "Active Probing using Packet Quartets," *ACM SIGCOMM IMW*, 2002.
- [21] A. Pasztor and D. Veitch, "On the scope of end-to-end probing methods," *IEEE Communication Letters*, 2002.
- [22] A. Pasztor and D. Veitch, "The packet size Dependence of packet pair like methods," *In Proceedings IWQoS02*, 2002.
- [23] V. Paxson, "End-to-end Internet packet dynamics," *IEEE/ACM Transactions on Networking*, 7(3):277–292, 1999.
- [24] R. S. Prasad, C. Dovrolis, and B. A. Mah, "The effect of layer-2 store-and-forward devices on per-hop capacity estimation," *IEEE INFOCOM*, 2003.
- [25] V. Ribeiro, M. Coates, R. Riedi, S. Sarvotham, B. Hendricks, and R. Baraniuk, "Multifractal Cross Traffic Estimation," *ITC Specialist Seminar on IP Traffic Measurement*, September 2000.
- [26] J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," *ACM IMC*, 2003.
- [27] W. Szczotka, "Stationary representation of queues. I.," *Advance in Applied Probability*, 18:815–848, 1986.
- [28] W. Szczotka, "Stationary representation of queues. II.," *Advance in Applied Probability*, 18:849–859, 1986.

[29] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker., "On the constancy of Internet Path Properties," *ACM SIGCOMM Internet Measurement Workshop*, November 2001.

APPENDIX I
PROOF OF THEOREM 5

Proof: We approach the proof in three steps. In the first step, we prove the second statement in the theorem. That is, we show that $\delta'(t)$ is continuous in the time interval (t_1, t_2) . Recalling Theorem 2, we have

$$\delta'(t) = \frac{Y_\delta(t)\delta}{C} + \frac{s}{C} + \tilde{I}_\delta(t). \quad (61)$$

The first two terms on the right hand side of (61) are constants with respect to t because of the following

$$\begin{aligned} Y_\delta(t)\delta &= V(t + \delta) - V(t) \\ &= Y_\delta(t_1)\delta - (V(t) - V(t_1)) + (V(t + \delta) - V(t_1 + \delta)). \end{aligned}$$

Since $(t_1, t]$ is within a workload epoch, there is no cross-traffic arrival during this interval. Hence $V(t) - V(t_1) = 0$. The same is also true in the interval $(t_1 + \delta, t + \delta]$ and we have $V(t + \delta) - V(t_1 + \delta) = 0$. Therefore, $Y_\delta(t) = Y_\delta(t_1)$ is a constant with respect to t . The third term $\tilde{I}_\delta(t)$ on the right hand side of (61) is obviously a continuous function of t . Hence $\delta'(t)$ is continuous in (t_1, t_2) .

In the second step, we show that $\delta'(t)_0^{\tau-\delta}$ can be split into a series of consecutive continuous segments. That is,

$$\delta'(t)_0^{\tau-\delta} = \bigcup_{i=1}^n \delta'(t)_{t_i}^{t_{i+1}}, \quad (62)$$

where $t_1 = 0$, $t_{n+1} = \tau - \delta$, and $t_i < t_{i+1}$ for all $1 \leq i \leq n$. We now prove this result by constructing such a partition.

Let e_1 be the first event-point in $W(t)$ after t_1 and e_2 be the the first event-point in $W(t)$ after $t_1 + \delta$. Then $\delta'(t)$ is continuous in the interval $(t_1, \min(e_1, e_2 - \delta))$ due to the result proved in the first step. Suppose that $\tau - \delta \leq \min(e_1, e_2 - \delta)$, then setting $n = 1$ and $t_{n+1} = t_2 = \tau - \delta$, the partition is accomplished. Otherwise, let $t_2 = \min(e_1, e_2 - \delta)$ and proceed with the partition of $\delta'(t)_{t_2}^{\tau-\delta}$ recursively. This eventually splits $\delta'(t)_0^{\tau-\delta}$ into the form of (62).

In the third step, we show that in any continuous segment $\delta'(t)_{t_i}^{t_{i+1}}$, $\delta'(t)$ is a piece-wise linear curve with possible slopes 0, 1, and -1 . We prove this result by recursively identifying all the linear segments in the interval (t_i, t_{i+1}) . We denote by e_1 and e_2 the first event-points in $W(t)$ after t_i and $t_i + \delta$ respectively. In what follows, we discuss the identification procedure in six possible cases.

In the first case when $W(t_i) > 0$ and $W(t_i + \delta) > 0$, let $t' = \min(e_1, e_2 - \delta, t_{i+1})$. Notice that for any $t_i < t < t'$, $I_\delta(t_i) = I_\delta(t)$ and consequently $\tilde{I}_\delta(t_i) = \tilde{I}_\delta(t)$. Combining the proof in the first step, we have $\delta'(t_i) = \delta'(t)$, which means that $\delta'(t)_{t_i}^{t'}$ is a linear segment with slope 0.

In the second case when $W(t_i) = 0$ and $W(t_i + \delta) = 0$, let $t' = \min(e_1, e_2 - \delta, t_{i+1})$. Using a similar argument, it is easy to see that $\delta'(t)_{t_i}^{t'}$ is also a linear segment with slope 0.

In the third case when $W(t_i) > 0$, $W(t_i + \delta) = 0$, and $I_\delta(t_i) \geq s/C$, let $t' = \min(e_1, e_2 - \delta, t_{i+1})$. Notice that for any $t_i < t < t'$, we have

$$I_\delta(t) = I_\delta(t_i) + (t - t_i). \quad (63)$$

Since $I_\delta(t_i) \geq s/C$, we get $\tilde{I}_\delta(t_i) \geq 0$ and

$$\tilde{I}_\delta(t) = \tilde{I}_\delta(t_i) + (t - t_i). \quad (64)$$

Combining (64) and the proof in the first step, which shows the other two terms on the right hand side of (61) are constants, we have $\delta'(t) = \delta'(t_i) + (t - t_i)$, which means that $\delta'(t)_{t_i}^{t'}$ is a linear segment with slope 1.

In the fourth case when $W(t_i) > 0$, $W(t_i + \delta) = 0$, and $I_\delta(t_i) < s/C$, let $t' = \min(e_1, e_2 - \delta, t_{i+1}, t_i + s/C - I_\delta(t_i))$. Notice that for any $t_i < t < t'$, (63) still holds, and we also have $\tilde{I}_\delta(t_i) = 0$ due to the case condition

$$\tilde{I}_\delta(t_i) = \left(I_\delta(t_i) - \frac{s}{C} \right)^+ = 0. \quad (65)$$

Further recall that $t < t' \leq t_i + s/C - I_\delta(t_i)$, and consequently we also have

$$0 \leq \tilde{I}_\delta(t) = \left(I_\delta(t_i) + (t - t_i) - \frac{s}{C} \right)^+ \leq 0. \quad (66)$$

Therefore, $\tilde{I}_\delta(t_i) = \tilde{I}_\delta(t) = 0$. Combining this result with (61), we have $\delta'(t) = \delta'(t_i)$, which means that $\delta'(t)_{t_i}^{t'}$ is a linear segment with slope 0.

In the fifth case when $W(t_i) = 0$, $W(t_i + \delta) > 0$, and $I_\delta(t_i) > s/C$, let $t' = \min(e_1, e_2 - \delta, t_{i+1}, t_i + I_\delta(t_i) - s/C)$. Notice that for any $t_i < t < t'$, we have

$$I_\delta(t) = I_\delta(t_i) - (t - t_i) \quad (67)$$

$$\tilde{I}_\delta(t_i) = \left(I_\delta(t_i) - \frac{s}{C} \right)^+ = I_\delta(t_i) - \frac{s}{C}. \quad (68)$$

Again recall $t < t' \leq t_i + I_\delta(t_i) - s/C$, we have

$$\begin{aligned} \tilde{I}_\delta(t) &= \left(I_\delta(t) - \frac{s}{C} \right)^+ = \left(I_\delta(t_i) - (t - t_i) - \frac{s}{C} \right)^+ \\ &= I_\delta(t_i) - (t - t_i) - \frac{s}{C} = \tilde{I}_\delta(t_i) - (t - t_i). \end{aligned} \quad (69)$$

Combing (61) and (69), we have $\delta'(t) = \delta'(t_i) - (t - t_i)$, which means that $\delta'(t)_{t_i}^{t'}$ is a linear segment with slope -1 .

In the sixth case when $W(t_i) = 0$, $W(t_i + \delta) > 0$, and $I_\delta(t_i) \leq s/C$, let $t' = \min(e_1, e_2 - \delta, t_{i+1})$. Notice that for any $t_i < t < t'$, we have

$$I_\delta(t) = I_\delta(t_i) - (t - t_i) \quad (70)$$

$$\tilde{I}_\delta(t) = \left(I_\delta(t) - \frac{s}{C} \right)^+ = 0. \quad (71)$$

Consequently, we also have

$$\tilde{I}_\delta(t) = \tilde{I}_\delta(t_i) = 0. \quad (72)$$

Combing (61) and (72), we have $\delta'(t) = \delta'(t_i)$, which means that $\delta'(t)_{t_i}^{t'}$ is a linear segment with slope 0.

Finally, if $t' = t_{i+1}$, we finished identifying all linear segments. Otherwise, we recursively identify the remaining linear segments in $\delta'(t)_{t'}^{t_{i+1}}$. Notice that the six cases we discussed above cover all possibilities. Hence, $\delta'(t)_{t'}^{t_{i+1}}$ can only contain piece-wise linear segments with possible slopes 1, 0, and -1 .

Combining all three steps, we proved the theorem. \blacksquare