Radon spectroscopy of packet delay

Andre Broido, Ryan King, Evi Nemeth, kc claffy
CAIDA, SDSC, University of California, San Diego
E-mail: {broido,evi,kc}@caida.org *

Abstract We demonstrate the feasibility of Internet spectroscopy techniques for analysis of rate limiting, packet interarrival delay and passive bitrate estimation of cell- or slot-based broadband connections. Working with highly diverse packet trace data, we find that delay's quantization in micro- and millisecond range is ubiquitous in today's Internet and that different providers have strong preferences for specific delay quanta in their infrastructures.

1. Introduction

In [1] we defined network spectroscopy as a branch of Internet science that deals with object identification on the basis of delay, period and frequency spectra. The primary goal of Internet spectroscopy is to identify qualitative features that are impossible to determine with available IP-level measurements, e.g., determining which Layer 2 technologies, switch types and levels of congestion were present in the packet's path. Several recent research studies have proposed techniques for inference of link and path characteristics [2] [3] [4]. Spectroscopy is different from previous approaches in that it emphasizes extracting information from: (1) packet timing jitters, which most other techniques interpret as noise; or (2) fine-grained delay quantizations, such as e.g. cell or slot times in TDM (time-division multiplexed) infrastructures.

Thousands of periodic processes coexist within the Internet system: DNS updates have periods of 24 hours, 1 hour and 75 minutes [1]; BGP updates are sent with 30 second periods; round-trip times (approximate periods) of most TCP connections are between 10 ms and 1 second [5]; SONET frames follow each other every 125 μs. It is natural to try to identify these phenomena by their periods.

As packets travel over the network they encounter various obstacles in the form of electro-optical converters, switching fabrics, input/output buffers, and forwarding engines, each of which affects their timing in predictable ways. Packet timings absorb quanta of various delays and realignments, e.g., SONET overheads or scheduled ATM cell times. The goal of network spectroscopy is to derive this information from observed timings, solving an analog of the inverse scattering problem. To that end, we use existing inverse scattering techniques, specifically the Radon transform previously applied in geophysics [6] and computer tomography [7], and entropy minimization, a technique similar to those used in image processing [8].

Different types of equipment or different bandwidth provisioning schemes normally do not span a continuous spectrum but come in distinct categories: operating system version, chassis model, processor type, link rate; all of these quantities assume only limited sets of values. Spectroscopy suggests a shift in methods from continuous to discrete, from dealing with real numbers to investigating properties of integers, and is characterized by a flavor of numerology. Its archetypal task is 'extracting the summands from the sum', hopeless on the surface of things in spite of the fact that cashiers solve such a problem many times a day when they provide us change in cash and coin.

This study tests techniques of Internet spectroscopy on the problems of interpreting packet interarrival delay distributions and passively estimating configured bitrate for broadband Internet links.

*Support for this work is provided by the Defense Advanced Research Project Agency (DARPA) NMS (N66001-01-1-8909) program and DISA's National Communiations System Organization. CAIDA is a collaborative organization supporting cooperative efforts among the commercial, government and research communities aimed at promoting a scalable, robust Internet infrastructure. CAIDA is based at the University of California's San Diego Supercomputer Center (SDSC). www.caida.org. Ryan King's work was done during his 2002 summer internship with CAIDA, supported by Rice University's EE department.

Knowledge of realistic packet interarrival delay distributions is necessary for verifying statistical theories and models of Internet traffic. These models in turn frame requirements directly related to running IP networks, such as provisioning router buffer memory and link rates [9]. They also restrict the choice of multiplexing schemes and connection topology. Our contribution involves demonstrating the utility of the discrete Radon transform in identifying both density shapes and singular points of probability measures associated with delay summands.

The problem of bitrate estimation arises in the context of SLA (service-level agreements) verification whereby the monitoring party needs to find out whether an ISP provides the contractually promised rate and maximal packet delay jitter to a customer. Bitrate estimation could also arise in security contexts such as IP traceback, when attackers use spoofed source addresses so that conventional source lookup in `whois` databases becomes infeasible. In such a case, information regarding provisioned bandwidth of upstream links and associated Layer 2 technologies may be used to narrow the search. Our contribution includes a technique for estimation of provisioned bitrate for cell-based links using interarrival delay quanta.

In the wider perspective, we demonstrate that the *modus operandi* of Internet spectroscopy can be successfully applied to a number of practically important and intellectually rewarding inverse problems.

Our data sources (Section 1.1) include passive traces taken at a major research university and at a Tier 1 ISP's backbone 2.5 Gbps link, as well as active measurements that corroborate our passive data analysis.

The rest of the paper is organized as follows. In section 2 we analyze our university data sets, using a course grained 2-dimensional Radon transform to determine the rate-limiting constraints and in particular the provisioned bandwidth. In sections 3.1 and 3.3 we present analyses of bitrates and interarrival time distributions for mass market broadband connections. Section 4 summarizes and outlines future work. Due to space limitations, some material has been omitted; this material can be found in [10].

1.1. Data description

Our two primary sources of data are our university's connection to the commodity Internet and a Tier 1 ISP's backbone link between two major metropolitan areas on the US West Coast. The Tier 1 ISP has rich infrastructure in Asia and Europe so that the traffic on this measured backbone link multiplexes a huge variety of diverse sources.

The university trace was collected on an ATM circuit provisioned over optic fiber running SONET OC-3 (155.5 Mbps) but rate limited to about 20 Mbps by the ISP. The link is often congested when school is in session, as it was when we collected the trace. This dataset contains 4.77 M inbound and 4 M outbound packets observed over 20 minutes between 8:00 and 8:20 on September 28, 2000. The capture is done by a Fore/Marconi FATM card with nominal timestamp resolution (*tick*) of 40 ns, but without external synchronization. This timestamp is associated with the first ATM cell in the packet.

We study the inbound (provider to university) direction since it is more congested. Analysis of packet traces is mostly done on properly provisioned links where the load is $< 50\%$. Our data is a rare glimpse of an overloaded link that is neither under attack nor experimental. The average bitrate is 16.9 Mbps (2535*8/1200, utilization is 85%.

We use data from Dag cards [11] [12] in the second half of the paper where we analyze a 39-minute packet trace collected on March 05, 2002 14:00-14:39 at a Tier 1 ISP's backbone link. The timestamps produced by Dag cards are of superior quality since they are synchronized to a GPS clock [13] [12] and since they are specifically designed with timing precision as a goal.[2]

Assuming all timestamps are taken at the beginning of packets, we define *interarrival time* (IAT) or packet delay as difference between adjacent timestamps, $d(k) = t_{k+1} - t_k$, and associate it with the k-th packet. When the portion of IAT attributable to packet length is known, we define *inter-packet delay* as the residual interarrival time, obtained by subtracting size-dependent delay from the IAT.

2. University link

We find that at the level of congestion observed on the university link, most packet interarrival times are determined by the durations associated with each packet's ATM cell count c_a. This quantity is a coarse discretization of packet sizes of which the most prominent are 40 and 1500

[2] Our Dag 4.11 cards have timing resolution of 15 ns.

bytes. Various flavors of SYN and ACK packets with TCP options and default MTU size of 576 bytes are also noticeable, as well as spikes that can be caused by specific preferences of some applications or individual sources. The 'continuous spectrum' of packet size follows an exponential baseline decreasing by a factor of 2 for each 230 bytes, flattening between 800-1400 bytes and rising again between 1400-1500 bytes.

We start our analysis of packet delay with a 2D scatterplot (Fig. 1a) of packet size vs. inter-arrival time.

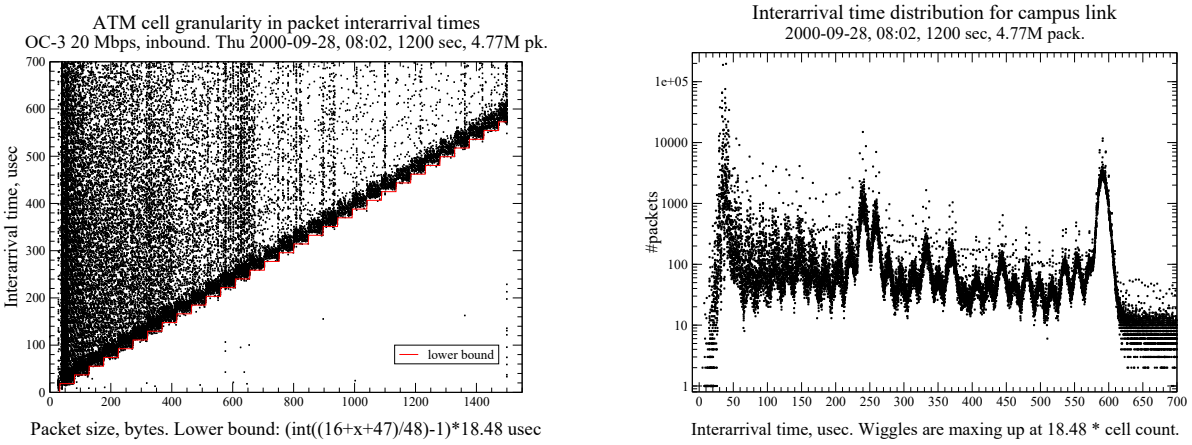

Figure 1.

The observations cluster in the staircase-shaped diagonal portion of the plot, where delay is likely to be close to link rate. The horizontal step is 48 bytes, i.e. ATM cell's IP payload, and the vertical step is one cell time that we find by the Radon transform described below.[3] The vertical width of the staircase is two cell times, which makes rectangles overlap on adjacent sides. The staircase contains 93% of all observations, with only a few outliers below it.[4]

Closer inspection of the plot shows that interarrival times tend to align on equi-spaced points separated by 2.24 μs, or $4 \cdot 560$ ns, which gives rise to 4.48 μs spacing of preferred delay values. Another instance of this 14-ticks (560 ns) granularity is ATM cell time 18.48 μs $= 33 \cdot 560$ ns that we will find by Radon transform. The Dag card data also has distinctive albeit different quantization patterns [16]. Identifying origins and sources of delay quantization is part of our proposed agenda for network spectroscopy.

Fig. 1b projects this ATM 'staircase' from the vertical delay axis of Fig. 1a onto the x-axis, creating a pattern of triangular spikes with Laplace-like ($e^{-|d-d_0|}$) shapes centered on integer multiples of ATM cell delay. These spikes are augmented by vertical outliers at preferred interarrival times and a sequence of higher packet counts stretching diagonally across the plot.

Fig. 2a plots a ccdf of this same data; the tail of the interarrival time histogram closely follows a Weibull distribution with shape parameter 0.7 and thus falls off more slowly than exponential (cf. [17].) However, only 2% of the interarrival times belong to this tail; the rest is attributed to the cell count distribution and rate limiter "window", which makes it very far from exponential, and the process far from being Poisson.

In fact, the packet arrival process on any highly loaded link cannot be Poisson. A continuous flow of packets has its packet interarrival time distribution determined by packet sizes and the underlying Layer 2 technology that maps packets into cells, slots, frames and other protocol data units (PDU). The packet size distribution is a compactly supported bathtub-like curve with maximums at the extremes (40 and 1500 bytes) and spikes at other preferred packet sizes. It is definitely not an exponential.

2.1. Radon transform

We split the distribution of interarrival times d on the university's link into a sum (cf.[15]) of two delays, $d = d_a + d_r$ with d_a proportional to a packet's ATM cell count and d_r being residual

[3]The 48-byte staircase is specific to ATM-induced delay. Firewall delay analyzed in [14] for Ethernet connections does not show this pattern, nor does router delay in [15].

[4]Plots in Fig. 1a and 1b are binned at a nominal resolution of 40 ns.

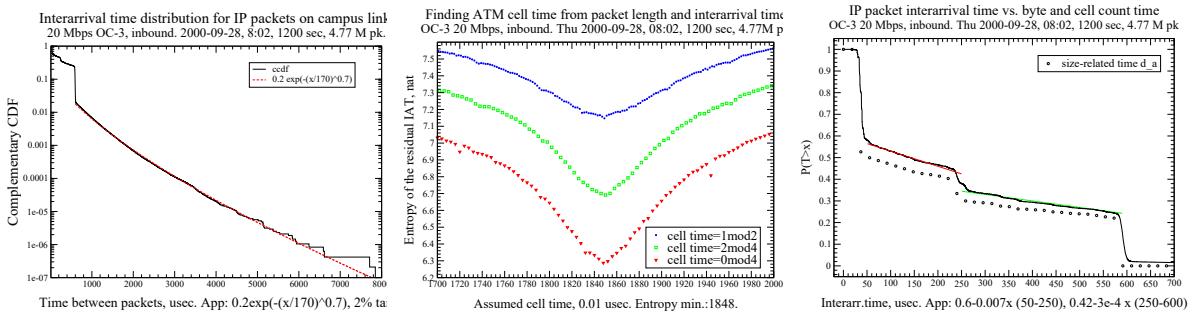

Figure 2.

(inter-packet) delay. The crux of our analysis is in determining the rate-limited cell time t_{cell} that relates ATM cell count c_a to d_a. This factor is the rate limiter's 'equilibrium' cell time. It translates directly to the university's provisioned bitrate for this connection. The search for this factor involves the discrete Radon transform [7] known as the *slant stack* in geophysics [6] [18].

The Radon transform works as follows. Let $p(b,d)$ be the normalized count (probability) of observed IAT d for packet size b bytes (Fig. 1a, shows where $p(b,d)$ is non-zero.) Let s be a guess for t_{cell}. Take all points in the (b,d) plane which lie on the staircase with step s seconds every 48 bytes, and sum the histogram $p(b,d)$ over these points. The result depends on the staircase's vertical shift, i.e. it is a histogram in d.

When the choice of s is right, we expect the summed histogram to have a single spike (Fig. 3a) corresponding to the observed staircase of Fig. 1a. We can recognize it by comparing some measure of spread for histograms with different s; for the right s the spread must be smallest. We use the standard Shannon-Kolmogorov entropy to estimate the spread, although this choice is by no means unique [2].

The Radon transform is given by the formula $p_R(d,s) = \sum_{b=b_{\min}}^{b_{\max}} p(b, d + s \cdot c_a(b))$ where $c_a(b)$ is the ATM cell count for a b-byte packet,[5] $c_a(b) = \text{ceil}\left(\frac{b+16}{48}\right)$ and the slope s is time per cell. The sum is taken from $b_{min} = 20$ to $b_{max} = 1500$ bytes. We find the cell time by looking for a slope that minimizes the entropy of the Radon-transformed distribution. Entropy is a measure of unpredictability [19] $H(s) = -\sum_d p_R(d,s) \log p_R(d,s)$ that is highest when all outcomes (in our case, delay values) are present with equal probability, i.e., there are no spikes. Since we would expect the entropy to be lower when the spikes in $p(b,d)$ for different b's line up (i.e., probability is concentrated in a fewer number of spikes after the transform), we expect that the entropy will be at a minimum at the true period of the data. We search for s_{me}, a slope that minimizes entropy of the transformed distribution $p_R(d,s)$ while performing the Radon transform over an interval of slopes s: $s_{me} = \text{argmin}(H(s),\ s_{\min} \le s \le s_{\max})$. We identify s_{me} with t_{cell}, the rate limiter's target time for transmission of an individual cell. We plot the entropy in Fig. 2b; it has a downward-looking cusp, rather than a parabolic shape at the minimum. This allows us to determine the cell time with full precision.[6]

The ATM cell time for this link is evaluated as $t_{cell} = 18.48\ \mu$s. The corresponding ATM payload bandwidth is 20.8 Mbps (20.2 with AAL5 correction). Since each packet is sent in an integer number of cells without reuse of the padding bytes before the trailer in the last cell [20], the effective IP bandwidth is around 19.3 Mbps, which is lower than the contractual 20 Mbps.

The delay distribution $p(d)$ is closely approximated by ATM cell counts derived from packet sizes in bytes, $p(d_a)$ with $d_a = t_{cell} c_a(b)$ (Fig. 2c.) The maximum absolute difference between two ccdfs (of cell counts scaled by t_{cell} and of interarrival times) is under 7%.

We now analyze the distribution of the residual delay d_r (Fig. 3a, 3b.) This distribution has two components of qualitatively different shape. Between plus or minus one cell time the curve is interpolating between Gaussian-like (center) and Laplace-like $\exp(-c|d|)$ (tail) curves with large (order of magnitude higher) singular points around the central rate. The exponential decay is not precisely symmetric though; it is close to $e^{0.42d}$ on the negative side and $e^{-0.3d}$ on the positive side. This reflects the rate limiter's inclination to err on the positive side of delay,

[5]The extra 16 bytes, the header and trailer of AAL5 encapsulation, cause almost every IP packet to consist of two or more cells.

[6]We scan values of s with 10 ns spacing; since the timestamps d are multiples of 40 ns, we get different entropy values for $s \equiv 10, 30$; $s \equiv 20$; and $s \equiv 0 \bmod 40$ ns, which align on the three curves in Fig. 2b.

i.e., to provide less rather than more bandwidth.

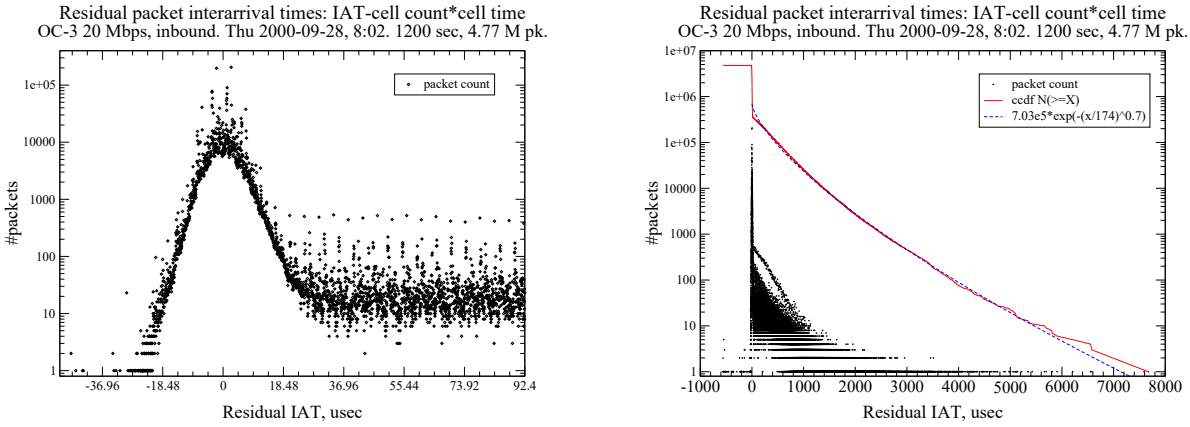

Figure 3.

Beyond the residual delay of one cell time, the distribution crosses over to what can be identified as link idle times, with a corresponding ccdf close to Weibull in the far tail.[7] The overall distribution is thus a sum of two probability measures: $p_{rl}(d_r)$ which is caused by rate limiter; $p_i(d_r)$ which contains idle times, although in the crossover region it is hard to distinguish them.

The description of the distributions as exponential, Weibull, etc. is only valid as a macroscopic (mean-field, average) approximation. In microscopic terms, the distribution is dominated by delay quantization, specifically the preference for residual delay d_r spaced with the step 4.48 μs (4·560 ns), $4.48k$+2.48 μs, $k = -2, -1, 0, 1, \ldots$ This quantization results in selected values having a factor of 3 higher probabilities than adjacent packet counts and an order of magnitude over the baseline where most counts are concentrated (Fig. 3a.) In particular, the two largest packet counts (at around 200 K packets) correspond to $k = -1$ or 0. We observe these progressions for original (non-residual) delay values where the progression's shift depends on the number of cells, e.g., for packets of 2 cells (33-80 bytes) it is 3.6 μs.

The prevalence of features due to cell counts, rather than idle times (gaps) in the interarrival times distribution can also be directly observed in properties of the aggregation of counts associated with cells, packets and bytes. Over aggregation intervals between 1 and 1024 ms, only packet counts converge to a bell-shaped distribution albeit a bimodal one. This is related to the fact that the rate limiter does not limit the number of packets per se. Cell counts and byte counts follow a characteristic pattern similar to one observed in [12], and do not converge to a Gaussian [22] [17].

We conclude that the well-known phenomenon of long-range dependence may have rate-limited sources as one of its contributing causes. Indeed, a never-ending stream of packets with a bitrate that is close to constant (like one generated by our connection) will necessarily create a correlation between byte counts with large time lags.

3. Backbone data

We observed exponential interarrival distributions for our backbone packet traces and selected two three-minute intervals, one in each direction, from the trace for detailed analysis. The intervals contained about 20 M packets each, yielding a packet rate of about 100 Kpps.

We observed a feature similar to that of the university link, namely a drop of the ccdf at several values at the beginning of the trace that correspond to "tailgaiting" packets with a few favorite sizes. However, the drop was much less than the drop in the university data, since the number of back-to-back packets was much smaller (link utilization under 25%.)

The remaining portion of the distribution was well-approximated by the exponential, that corresponds to Poisson interarrivals. We chose an exponent with average interarrival time 10 μs,

[7] The Weibull curve $7.03\exp(-(x/173)^{0.7})$ has 18.9% relative error for 5.8% of packets with $118 \le dt \le 5400$ μs, which covers over 4 orders of magnitude on a probability scale. We assume that the approximation is good when its relative error is less than 20% [21].

$\exp(-d/10)$ and found that the exponential interarrival distribution approximates southbound traffic on the interval $[6, 130]$ μs with relative accuracy under 20%[8]. We chose a lower bound of 6 μs since the back-to-back distribution ends at around 5 μs for an OC-48 link. The upper bound at 130 μs cuts off 51 values of d (each with sample count of 1).

The northbound traffic (Fig. 4a) is slightly more intense (about 21 M packets in 3 minutes). We found that the approximation by $\exp(-d/10)*0.7$ had relative error under 19% on $[6, 100]$ μs. The fact that we needed the correction factor is likely due to the larger number of back-to-back packets in this sample.

3.1. DSL broadband connections

The vast majority of DSL modems today use ATM as a transport layer protocol. Often modems will use PPP over ATM or even PPP over Ethernet over ATM. The use of ATM has significant detectable effects on traffic traversing the link. While our experiments have analyzed DSL upload characteristics, we would likely see similar characteristics in DSL download traffic. Since delay quantization in such traffic can only be measured at the modem, we do not yet have any DSL download data.

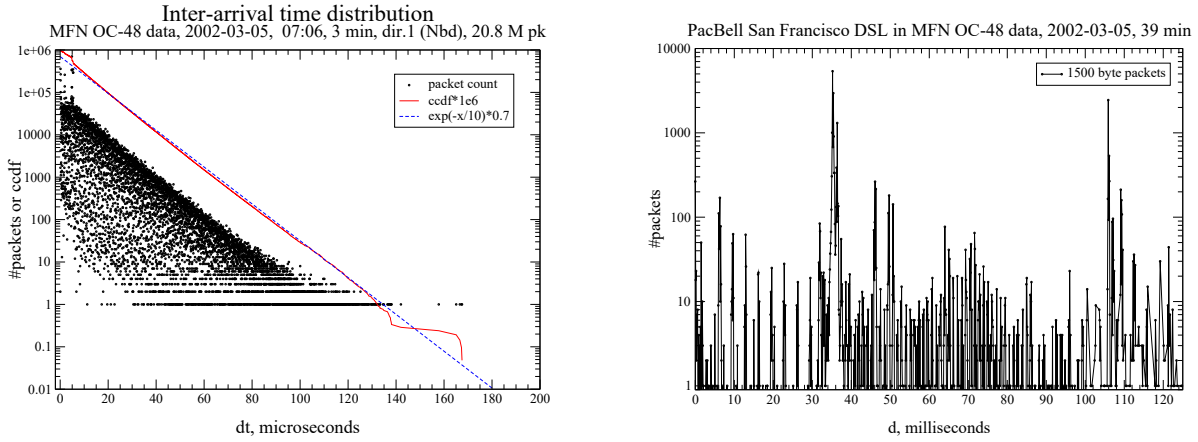

Figure 4. Interarrival time distributions: a) MFN backbone, b)San Francisco DSL provider

We have analyzed both active and passively gathered data. The active measurements were collected by sending UDP packets and timestamping them at a receiving computer. The passive data used was gathered from a traffic monitor on an OC48 link.

Unless otherwise specified, we measured packet interarrival times between packets from the same source IP address selected among flows of more than 1 MB as computed by CoralReef [23]. We then merged statistics for several IP addresses whose DNS names met four criteria: (1) same domain (e.g. `pacbell.com`); (2) same city code (e.g. `sfo`); (3) an indication of broadband access (e.g. `dsl`); and (4) same routed IP prefix. For packet traces we used 1, 10 or 100 μs bins in the interarrival histograms[9].

The interarrival delay histogram for a set of DSL hosts in San Francisco (Fig. 4b) is highly periodic, with 3.313 ms between successive spikes. This periodicity is due to ATM cell framing over the DSL links. Assuming ATM cell time of 3.313 ms, ATM's raw bitrate is $53 \cdot 8$ bit / 3.313 ms = 128 Kbps, which matches advertised DSL offerings. For example, upload bandwidths of 128 and 384 Kbps are available from PacBell.

The interarrival histograms for 1500 byte packets[10] (Fig. 4b) have two large spikes at 35 ms and 106 ms. The 106 ms spike in the 1500 byte traffic corresponds to 128 Kbps ($32 \cdot 53 \cdot 8$ bits / 106 ms) and the 35 ms spike to 384 Kbps. This pattern indicates that some of the DSL modems we are analyzing have a higher level of service. Another possibility is queueing at a

[8] We argue in [21] that the approximation can be viewed as good when its accuracy is 20% or less.

[9] The rationale for using 1 μs or larger bins is to avoid 0.5 μs uncertainty caused by SONET transport overheads [13] [12].

[10] The packet to which delay is attributed is 1500 bytes long. Dag timestamping is done at the packet's first HDLC byte, i.e., 4 bytes before the IP header [13], thus the delay is attributed to the first packet.

384 Kbps link upstream from a constellation of 128 Kbps sources. This alternative is less likely since multiplexing devices normally have much higher bitrates than individual users.

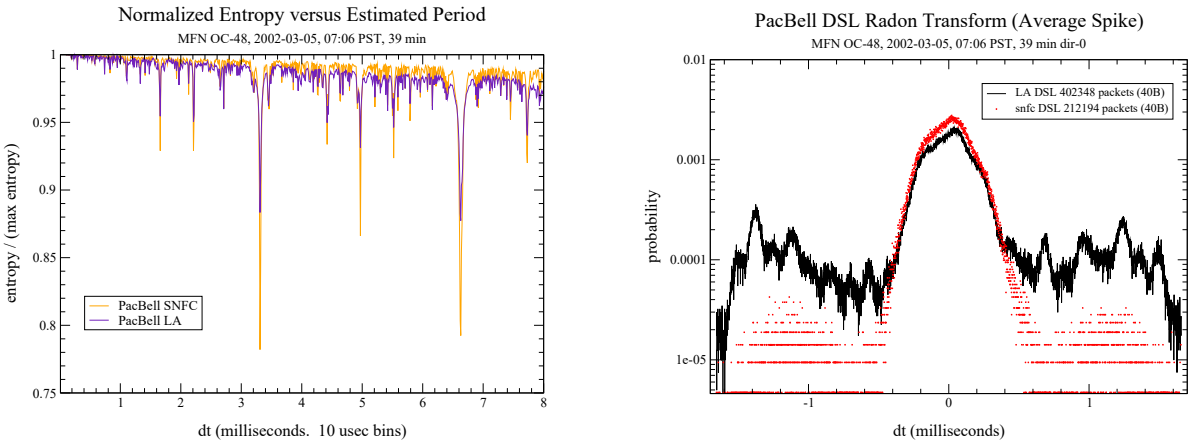

Figure 5.

To corroborate this analysis we took a set of DSL hosts with the same provider in Los Angeles. We found again a 3.313 ms period in histograms and a 106 ms spike for 1500-byte packets. This set did not have a 35 ms (384 Kbps) spike for 1500 byte packets, the most likely cause being that the higher bandwidth was not offered in that area. The Los Angeles data also had some finer splitting of spectral lines (about 1 ms spacing) for 40-byte packets around 25 and 75 ms.

We determined the exact period of the spikes using the Radon transform[11] of the data over different periods s, $p(d,s) = \sum_{k=0}^{\infty} p(d+ks),\ 0 \le d < s$ and computed the entropy for each s as in Sec. 2. We expect the entropy to be lower when the spikes in the distribution line up, and at a minimum at the true period of the data. This technique worked remarkably well for our data. Both the Los Angeles and San Francisco data sets have sharp downward spikes at $t_{cell} = 3.313$ ms (Fig. 5a.)

After selecting $t_{cell} = 3.313$ ms as the delay quantum corresponding to an ATM cell, we investigated the inter-packet delay histograms for the SF and LA data sets (Fig. 5b). We found that they share similar traits, which is expected, since the ISP is likely to use similar infrastructure design and equipment in both cities.

3.2. Active measurements

To better determine what types of effects were causing the periodicity in the backbone trace, we made active measurements over a DSL modem in San Diego, CA. We wanted to to verify that the periodic pattern was imposed by the DSL modem, rather than by TCP or a specific application. We wrote WinSock software to send UDP packets that were received on a 1 Ghz dual processor FreeBSD machine at the university with a 100 Mbps Ethernet access link.[12]

The data gathered shows similar characteristics to the passive data set. Both have a distinct period. It is interesting to note that the period of the data from the San Diego DSL modem (Fig. 6a) is approximately 2.65 milliseconds whereas the data gathered passively from modems in San Francisco and Los Angeles had periods of 3.313 ms.

The joint distribution of interarrival time vs. packet size for the DSL modem is similar to that of the university ATM link (Sect. 2.) Our active measurement lasted nearly 24-hours on August 10th 2002. The UDP packets of size 40-1500 bytes were sent back-to-back in bursts of 11, with a 2-second gap between bursts; each burst contained packets of the same size.[13] We recorded interarrival times for each burst using the Pentium `rdtsc` (read timestamp counter) instruction with a nominal resolution of 1 ns.

[11] These are one-dimensional Radon transforms, taken over delay histograms with fixed packet size, 40 bytes in this case, as opposed to 2D transforms for determining ATM cell timing which we obtained by summation over the range of packet sizes.

[12] `tracert` shows 14 hops in 4 ASes between our source and receiver (250 miles via Los Angeles); surface distance is about 1 mile. RTT is 39 ms, of which 19 ms is taken by the first hop.

[13] This makes interarrival time independent of where in the packet the timestamp is taken. Packet filters timestamp the end of the packet [12].

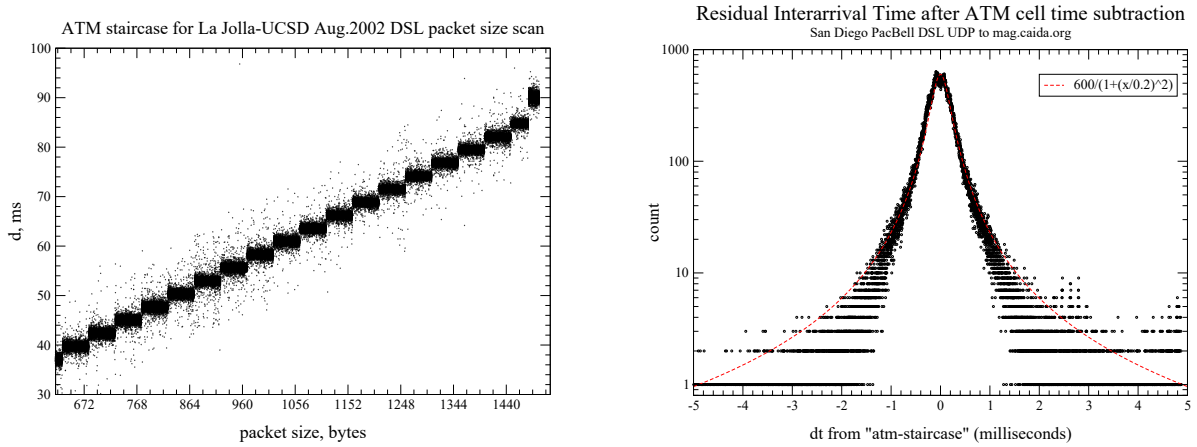

Figure 6.

Fig. 6a shows the staircase with 48-byte stairs in size-delay space (similar to that of Fig. 1a), and confirms the use of ATM as the link layer protocol for the DSL modem.

We can derive the number of additional encapsulation bytes from the offset of the stairs in the packet delay-vs-size 2D histogram; it turns out to be 40.[14] We can model the ATM staircase with the formula: $d_a(b) = \text{ceil}((b + 6 + 18 + 16)/48) \cdot 53 \cdot 8/160\text{ms}$ with 6 bytes for PPP, 18 bytes of Ethernet headers, and 16 bytes of AAL5. This number suggests that the DSL modem is using PPP over Ethernet over ATM.[15]

The contract with the provider specified that the modem would operate at 1.5 Mbps download and 128 Kbps upload. However it is clear that ATM bytes are actually being uploaded over the DSL link at 160 Kbps. Since the data is being encapsulated by Ethernet, PPP and ATM, the actual IP throughput is around 130 Kbps.

The shape of the residual delay distribution $d_r = d - d_a$ for this DSL experiment (Fig. 6b) is symmetric with decay slower than exponential as we move away from the center. It is reminiscent of Levy-stable distributions [24].[16]

We also tried to analyze several other IP addresses with "`dsl`" in their FQDN hostname. One address in San Jose has a delay quantum (spike period) of 8.3 ms, which corresponds to 51 Kbps of ATM bitrate. We do not know whether the technology used is actually ATM, although it may include ATM with some kind of rate limiting or sharing. All overseas hosts with `dsl` in their domain names had different preferences for certain interarrival times yet none of them had a series of sharp spikes with the strict periodicity observed for the studied Californian DSL provider. We leave the study of international DSL for future work.

3.3. Cable Modems

We analyzed traffic from three ISPs whose service uses cable modems. We found that cable modems have their delay quantized by integer number of milliseconds. This is related to the standard frequency for permissions to send data (maps), typically 500 maps/sec. Due to space limitations, this section is presented in our preprint [10].

3.4. Discussion

Entropy minimization-based algorithms for identifying delay quanta need to take into account two issues: smooth variability in baseline entropy and the presence of *resonances*, rational multiples of the optimal period with spike heights comparable to those at the true value.

Scanning over large ranges of periods (e.g., when $s_{\max}/s_{\min} > 2$) involves growth in the baseline entropy (of the associated uniform distribution) which is equal to the logarithm of the number of data bins in the period, $\log(s/\Delta d)$. We found it useful to divide (normalize) entropy by this number, and this is what we plot in Figure 5a However, this division can result in

[14] For example, the jump to 15 cell times (40 ms) in Fig. 6a occurs not at 673 bytes (14*53+1) but at 633 bytes.

[15] The prominent jump at the upper right end of Fig. 6a reflects packet fragmentation caused by host MTU of 1480 bytes. With IP payload $b > 1480$ bytes, the encapsulated first fragment contains 32 cells (40+1480 bytes) and the smaller fragment contains 2 cells (40+28+1...40+28+20 bytes.) This fragmentation causes an extra two cells for IP packets between 1480-1496 bytes, and one extra cell for sizes 1497-1500, resulting in smaller spikes at 5.3 and 2.65 ms respectively in Fig. 6b.

[16] We show Cauchy density $600/(1 + (x/0.2)^2)$ in Fig. 6b. The non-residual density of 1500 byte packets also admits a Cauchy fit. This 1500 byte data has very fine structure in delay spectrum with maximums spaced at 164 and 41 μs, the origin of which is unknown.

closer values for minimums of normalized entropy so that the choice between periods becomes non-unique suggests both 3 and 6 ms as likely candidates for periods.) This result is expected since if the data has a period of D, it also has a period of $2D$. However, using unnormalized entropy in the algorithm would result in the smaller period.

Another issue is that large disparity between spikes in the data can make the average spike corresponding to the entropy minimum s_{me} depend mostly upon the contribution of the one largest spike rather than the totality of all spikes. Conclusions based on the average spike may only be valid for a narrow range of d that comprise the largest summand. This is a special case of the problem of averaging data with outliers.

Our experience with outliers suggests that it may be possible to obtain better results by averaging logarithms of the data, i.e. taking geometric averages. Whether this approach applies to the problem of creating representative generic shapes for quantized delay distributions requires further investigation.

4. Conclusions and future work

We have shown that the Radon transform of packet interarrival time distributions, coupled with entropy minimization, can be used for estimation of provisioned bandwidth and for identification of Layer 2 technologies such as ATM, rate-limited ATM, DSL, PPP, Ethernet and cable modems in IP traffic. More generally we have demonstrated the utility of the Internet spectroscopy approach for solving a variety of identification problems.

We presented an algorithm that evaluates the interarrival delay quantum (cell time) for rate-limited cell-based links. The algorithm takes a joint 2D distribution of packet sizes and interarrival times as its input, converts it by a coarse-grained Radon transform to a family of 1D marginals. Each marginal has the semantics of an inter- and intra-packet delay (i.e. link idle time) histogram that corresponds to an assumed value of cell time. Our estimate of cell time is the value that minimizes entropy of such a marginal, i.e. makes it closest to a delta function.

As an application of the Radon transform technique, we determined the target cell time for the rate limiting performed by the university commodity ISP to provide a 20 Mbps connection over 155 Mbps link. This allows us to verify an under-fulfillment of the service contract. Knowledge of cell time enables us to compute the distribution of inter-packet delay. We find that it consists of two separate components overlapping in the time domain: a spike with a width of two cell times that corresponds to the rate limiter's fluctuations around the target rate, and the 'true' link idle time. The true idle time's integral (ccdf) closely follows a Weibull curve, while individual values are subject to fine-grained delay quantization. We also find that the link's high load renders the packet arrival process quite dissimilar to Poisson. Combined with the rate limiter's long-term memory, this deviation makes the byte counting process strictly non-Gaussian over a wide range of aggregation intervals (up to 1 sec).

We analyzed bitrates and other properties of broadband mass-market connections and determined interarrival times for DSL and cable modem sources by a simplified 1D version of the min-entropy Radon algorithm applied to packets of fixed size (40 or 1500 bytes). We find that delay quantization in broadband access is dependent upon providers, technologies and markets, even though the number of choices appears to be rather limited. This suggests that network spectroscopy has a potential for source recognition, if a library of interarrival quanta and inter-packet delay distributions is available.

To promote network spectroscopy from a computational art to routine technology, it will be necessary to enable automated analysis of measurement data by the algorithms presented here. We intend to create a library of delay spectra corresponding to known devices and link types. This library will allow recognition of variations in standards' implementation specific to different markets and providers. We are working on a Fourier-based identification method which we expect to be more sensitive than binned entropy is to quantitative details of data clustering. We also plan to extend the analysis presented here to other connection technologies, including major brands of wireless networking (802.11, fixed broadband and mobile), and to associate delay quanta that we found in the last section with settings that reflect customer provisioning and rate-limiting policies of some ISPs.

Yet another important venue of research is assessment of the accuracy of spectroscopy-based inference and potential for measurement artifacts that distort the data and affect results of interpretation. We plan to evaluate timing precision of available monitoring equipment and to cross-check timestamping by products from different vendors.

5. Acknowledgments

The idea of network spectroscopy crystallized during the first author's visit in March-June 2002 to UCLA's Institute for Pure and Applied Mathematics Large-scale communication Networks Program. Thanks to professor David Donoho of Stanford, IPAM's director Mark Green and friendly staff for creating an environment conducive to this research. We greatly benefited from discussions with Robert Nowak of Rice University and with participants at CAIDA's Bandwidth Estimation workshop in June 2002, especially Constantinos Dovrolis. Thanks to our colleagues Ken Keys of CAIDA for his help with the Coral software suite and to David Moore for the university trace. Thanks to Stephen Donnelly of University of Waikato/Endace for discussions of SONET timing. The data analysis presented here would never have been possible were it not for the efforts of Joerg Micheel of University of Waikato/NLANR who created the measurement infrastructure and collected traces used in this and many other studies.

REFERENCES

1. A. Broido, E.Nemeth, and kc claffy, "Spectroscopy of DNS update traffic," ACM SIGMETRICS 2003, San Diego, June 2003.
2. Dina Katabi and Charles Blake, "Inferring congestion sharing and link characteristics from packet interarrival times," MIT LCS Technical Report, MIT-LCSTR -828, December 2001.
3. Mark Coates, Alfred Hero, Robert Nowak, and Bin Yu, "Internet tomography," May 2002, IEEE Signal Processing Magazine v.19.
4. Ravi S. Prasad, Constantinos Dovrolis, and Bruce A. Mah, "The effect of layer-2 switches on pathchar-like tools," in *Proceedings of the IMW 2002. Marseille, France*, Nov 2002.
5. Andre Broido, "Invariance of Internet RTT spectrum," October 2002, ISMA, Oct.2002, Leiden. `www.caida.org/outreach/isma/0210/ISMAagenda.xml`.
6. Jon Claerbout, *Imaging the Earth's Interior*, Blackwell, 1985.
7. Frank Natterer, *The Mathematics of Computerized Tomography*, Wiley, 1985.
8. "Bayesian inference and maximum entropy methods in science and engineering : 21st International Workshop," 2002, Edited by Robert L. Fry.
9. Charles Fraileigh, "Provisioning IP backbone networks to support delay sensitive traffic," Ph.D.Thesis, EE Department, Stanford University, 2002.
10. Andre Broido, Ryan King, Evi Nemeth, and kc claffy, "Radon spectroscopy of inter-packet delay," 2003, Proceedings of High Speed Networking (HSN) Workshop at Infocom.
11. Ian D Graham, Murray Pearson, Jed Martens, and Stephen Donnelly, "Dag - A cell capture board for ATM measurement systems," www.cs.waikato.ac.nz /Pub/Html/ATMDag/dag.html.
12. Stephen Donnelly, "High precision timing in passive measurements of data networks," June 2002, Ph.D. thesis, University of Waikato, Hamilton, New Zealand.
13. Joerg Micheel, Stephen Donnelly, and Ian Graham, "Precision timestamping of network packets," Nov 2001, Proceedings of ACM Sigcomm Internet Measurement Workshop.
14. Klaus Mochalski, Joerg Micheel, and Stephen Donnelly, "Packet delay and loss at the Auckland Internet access path," PAM 2002.
15. K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, F. Tobagi, and C. Diot, "Analysis of measured single-hop delay from an operational backbone network," Proceedings of INFOCOM 2002, June 2002, New York.
16. Andre Broido and kc claffy, "Spectral analysis of DAG and SONET clocks," in preparation.
17. J. Cao, W. Cleveland, D. Lin, and D. Sun, "Internet traffic tends to Poisson and independent as the load increases," 2001.
18. O.G.Kutina and A.B.Kutin, *Seismic Interface Tracing*, Nedra, Moscow, 1993.
19. Claude Shannon, *The mathematical theory of communication*, 1949, Univ.Illinois Press.
20. J.Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5, RFC1483," 1993.
21. Andre Broido and k claffy, "Internet Topology: connectivity of IP graphs," in *SPIE conference on Scalability and Traffic Control in IP Networks*, Aug 2001, `http://spie.org/Conferences/Programs/01/itcom/confs/4526.html`.
22. Jorma Kilpi and Ilkka Norros, "Testing the gaussian approximation of aggregate traffic," in *Proceedings of the IMW 2002. Marseille, France*, Nov 2002.
23. K. Keys, D. Moore, R. Koga, E. Lagache, M. Tesch, and k. claffy, "The architecture of the CoralReef: Internet Traffic monitoring software suite," in *PAM 2001*.
24. Paul Lévy, "Theorie de l'addition des variables aleatoires," Gauthier-Villiers, Paris, 1937.