Radon spectroscopy of inter-packet delay

Andre Broido, Ryan King, Evi Nemeth, kc claffy
CAIDA, San Diego Supercomputer Center,
University of California, San Diego
E-mail: {broido, 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.**

The paper consists of two parts. First part presents an algorithm that evaluates 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. The 2D distribution is then converted by a coarse-grained Radon transform to a family of 1D marginals. Each marginal has semantics of 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 Radon transform technique, we determine the target cell time for the rate limiting performed by the university commodity ISP to provide 20 Mbps connection over 155 Mbps link. This allows to verify an underfulfilment of the contract. The 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 time domain: a spike with a width of two cell times that corresponds to the rate limiter's fluctuations around target rate, and "true" link idle time, whose integral (ccdf) closely follows a Weibull curve, while individual values are subject to a fine-grained delay quantization. We also find that the link's high load makes the packet arrival process to be very dissimilar with Poisson. Combined with the rate limiter's long-term memory, this deviation makes the byte counting process to be strictly non-Gaussian over a very wide range of aggregation intervals (up to 1 sec).

In the second part we analyze bitrates and other properties of broadband mass-market connections. We determine 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 interpacket delay distributions is available.

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.

I. 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) infrastrictures.

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 us. 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 electrooptical converters, switching fabrics, input and 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.

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 [10] 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 based on interarrival delay quanta.

In 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 II) 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 III we analyze our university data sets, determining the rate-limiting constraints and in particular the provisioned bandwidth. In sections IV and V we present analyses of bitrates and interarrival time distributions for mass market broadband connections. Section VII summarizes and outlines future work.

II. 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 of 40 ns, but without external synchronization. This timestamp is associated with the first ATM cell in the packet.

We study inbound, i.e., provider to university direction since it is more congested. The average bitrate is 16.9 Mbps (2535*8/1200, Table II); utilization is 85%. The detailed description of this dataset can be found in [11].

The timestamps produced by Dag cards are of superior quality since they are synchronized to a GPS clock [12] [13] and since they are specifically designed with timing precision as a goal.[1] We use data from Dag cards 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.

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 IAT.

III. University link

We find that at the level of congestion observed at the university link, the bulk of packet interarrival times is determined by the durations associated with each packet's ATM cell count c_a. This quantity is a coarse discretization of packet sizes (Fig.1) of which the most prominent are

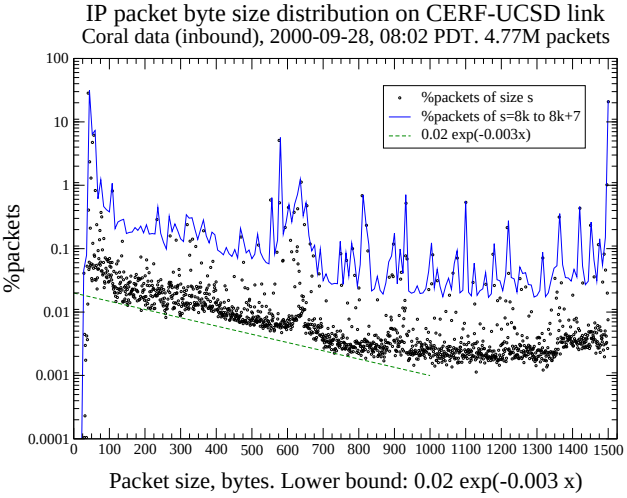

Fig. 1.

40 and 1500 bytes (Table I). 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 per each 230 bytes, flattening between 800-1400 bytes and rising again between 1400-1500 bytes.

Table II lists protocol, packet and byte mix.

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

TABLE I

MOST COMMON PACKET SIZES IN THE SEP.2000 UNIVERSITY DATASET

Rank	Size	Percent	Accum.
1	40	28.42	28.42
2	1500	20.78	49.20
3	56	6.14	55.34
4	576	5.10	60.44
5	52	4.74	65.18
6	44	2.34	67.52

TABLE II

UNIVERSITY INBOUND TRAFFIC BY PROTOCOL, SEPT.2000 DATASET

Proto	#	Bytes	%Bytes	Packets	%Pkt	Av.sz
	0	785520	0.03	19638	0.41	40
ICMP	1	19.2 M	0.76	311558	6.53	62
TCP	6	2337 M	92.17	4023780	84.34	581
UDP	17	178 M	7.03	413088	8.66	431
	54	204	0.00	3	0.00	68
	57	394	0.00	2	0.00	197
	93	331030	0.01	2871	0.06	115
	Total	2535 M	100	4770939	100	531

We start our analysis of packet delay with a 2D scatterplot (Fig.2) of packet size vs. interarrival time. The ob-

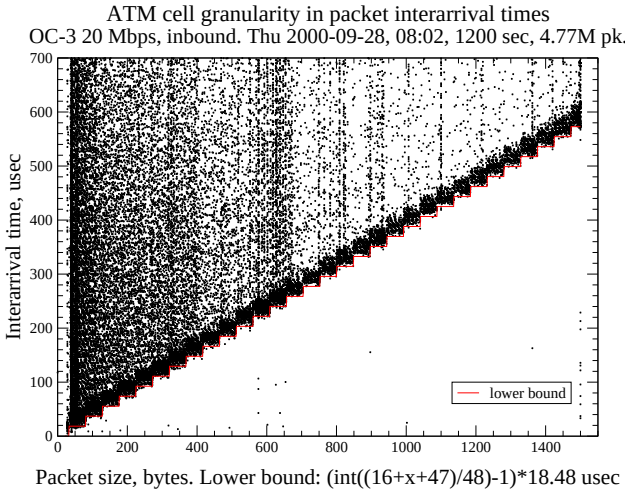

Fig. 2.

servations 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 Radon transform as described below.[2] 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.[3]

Closer inspection of the plot shows that interarrival times tend to align on equispaced points separated by 2.24 us, or $4 \cdot 560$ ns, which gives rise to 4.48 us spacing of preferred delay values (see below.) Another instance of this 14-ticks (560 ns) granularity is ATM cell time 18.48 us $= 33 \cdot 560$ 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.

[2]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].

[3]Plots in Fig. 2 and 3 are binned at nominal resolution of 40 ns.

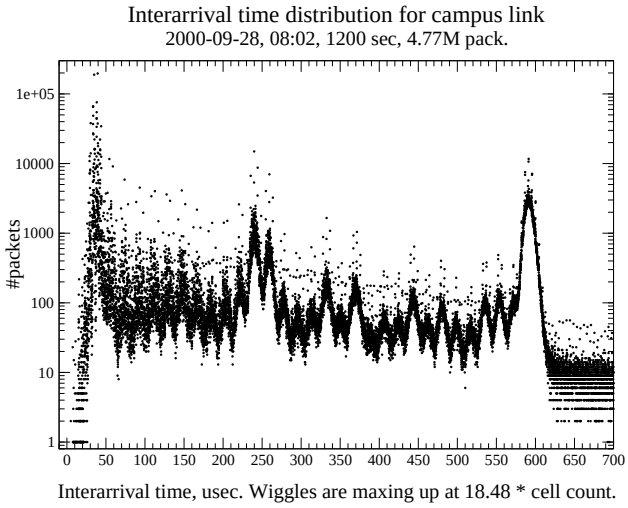

Fig. 3.

Fig.3 projects this ATM 'staircase' from the vertical delay axis of Fig.2 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 (similar to the distribution of ATM cell counts in Fig.9). These spikes are augmented by vertical outliers at preferred interarrival times and a sequence of higher packet counts stretching diagonally across the plot.

Fig.4 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 cell counts disribution and rate limiter "window", which makes it very far from exponential, and the process far from being Poisson.

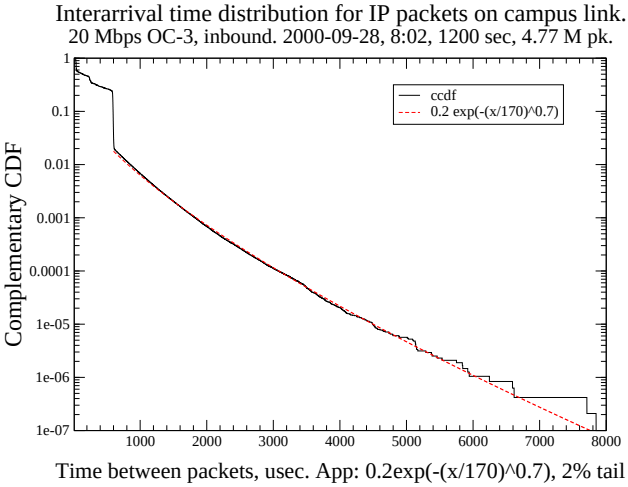

Fig. 4.

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 (recall Fig.1) is a compactly supported bathtub-like curve with maximums at extremes (40 and 1500 bytes) and spikes at other preferred packet sizes. It is definitely not an exponential, neither its mappings to PDUs are exponentials.

A. Radon transform

We split the distribution of the interarrival time d on the university's link into a sum (cf.[15]) of two delays, $d = d_a + d_r$ with d_a proportional to packet's ATM cell count and d_r being residual (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 equals 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 *slant stack* in geophysics [6] [18].

The Radon transform or slant stack works as follows. Let $p(b, d)$ be the normalized count (probability) of observed IAT d for packet size b bytes (Fig.2, shows where $p(b, d)$ is non-zero.) Let s be a guess for t_{cell}. Take all points in (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.7) corresponding to the observed staircase of Fig.2 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].

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,[4] $c_a(b) = \text{ceil}\left(\frac{b+16}{48}\right)$ 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 [20]

$$H(s) = -\sum_d p_R(d, s) \log p_R(d, s)$$

which 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., probablility 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.5; 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.[5]

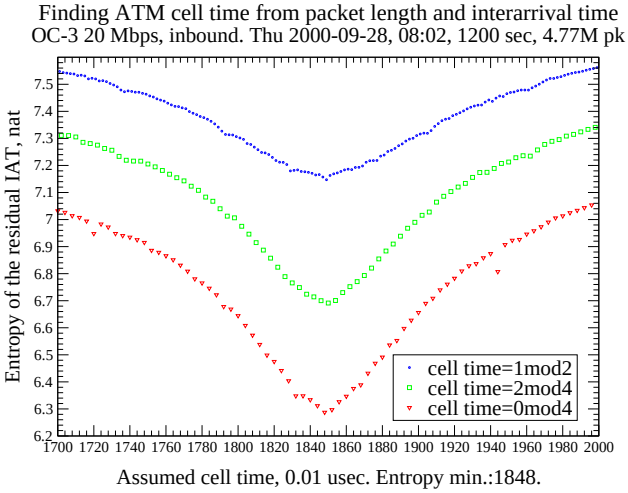

Fig. 5.

The ATM cell time for the link in question is evaluated as $t_{cell} = 18.48$ us. 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 [19], the effective IP bandwidth is around 19.3 Mbps, which is lower than the contractual 20 Mbps.

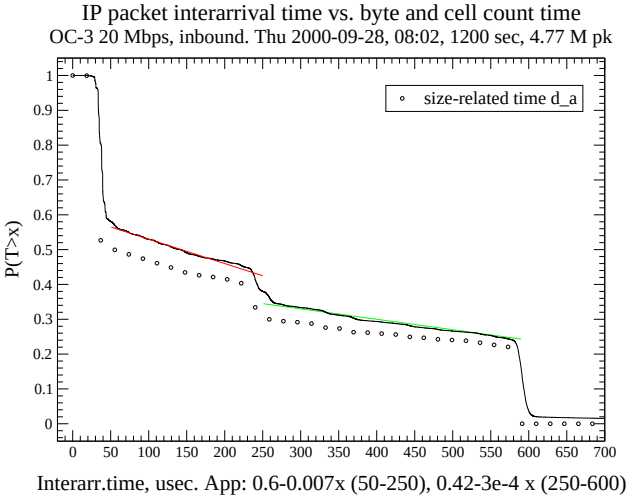

Fig. 6.

[4]The extra 16 bytes, the header and trailer of AAL5 encapsulation [19] render almost every IP packet to consist of two or more cells.

[5]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.5.

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.6.) 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.7, 8.) 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 and $e^{-0.3d}$ on the positive side. This reflects the rate limiter's inclination to err on the positive side of delay, i.e., to provide rather less than more bandwidth.

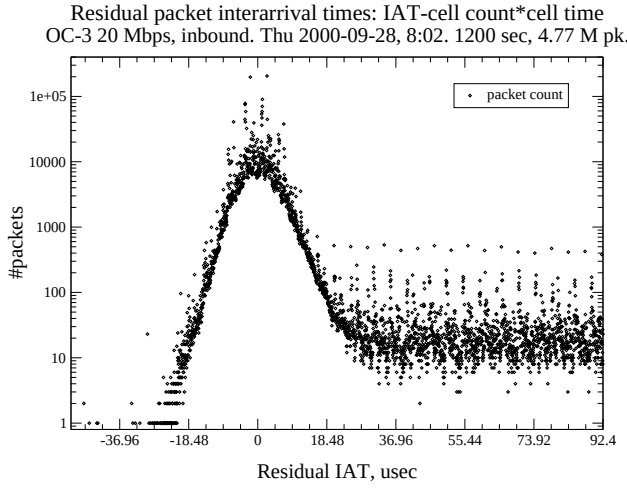

Fig. 7.

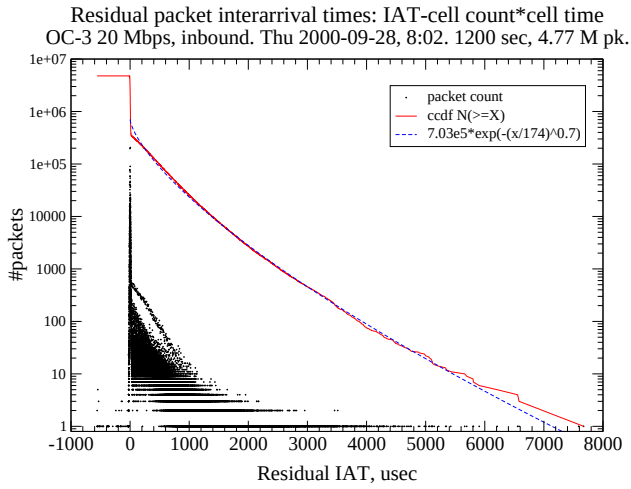

Fig. 8.

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.[6]

[6]The Weibull curve $7.03\exp(-(x/173)^{0.7})$ has 18.9% relative error for 5.8% of packets with $118 \le dt \le 5400$ us, 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].

The overall distribution is thus a sum of two probablility 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.

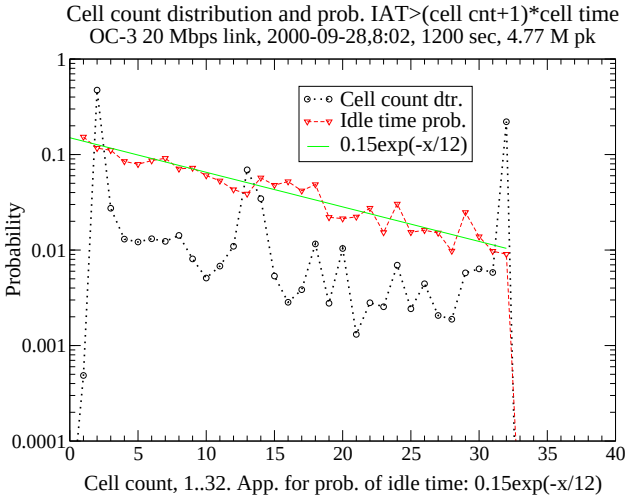

Fig. 9.

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 us ($4 \cdot 560$ ns), $4.48k + 2.48$ us, $k = -2, -1, 0, 1, \ldots$ 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.7.) 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 us.

The prevalence of features due to cell counts, rather than idle times (gaps) in the interarrival times distribution can be also directly observed in properties of aggregation for counting processes associated with cells, packets and bytes. Over aggregation intervals between 1 and 1024 ms, only packet counts converge to a bell-shaped distribution (Fig.10) 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 of Fig.11, similar to one observed in [13], and do not converge to a Gaussian [22] [17].

IV. Broadband connections: DSL

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

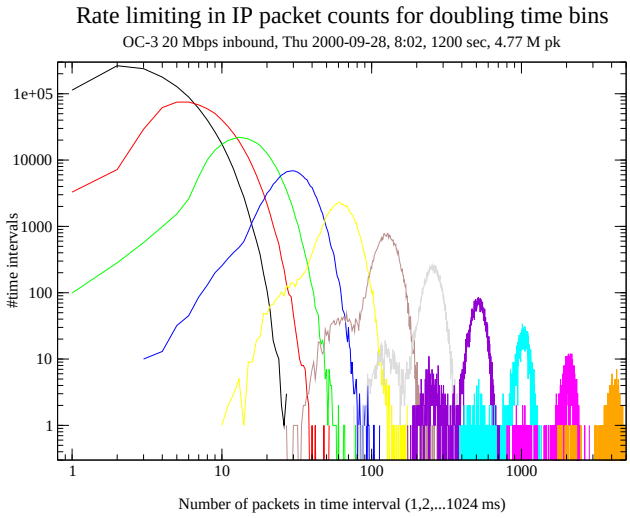

Fig. 10.

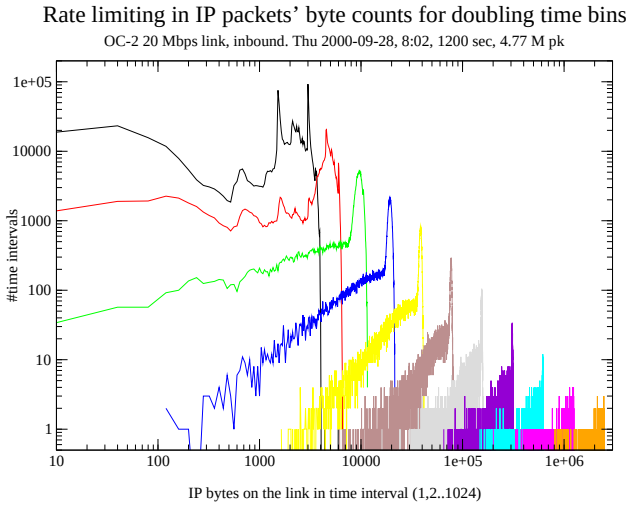

Fig. 11.

ATM has significant detectable effects on the traffic traversing the link. While our experiments have analyzed DSL upload characteristics, it is likely that we would 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.

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 in this analysis 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 in a table produced by CoralReef suite [23]. We then merged statistics for several IP addresses that met four criteria: (1) same domain (e.g. `pacbell.com`); (2) same city code (e.g. `snfc`); (3) an indication of being used for broadband access (e.g. `dsl`); and (4) same routed IP prefix. For packet traces we used 1, 10 or 100 usec bins in the interarrival histograms[7].

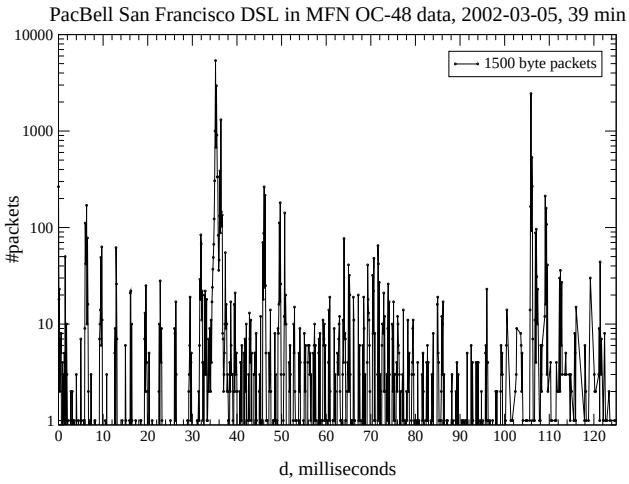

Fig. 12. San Francisco DSL IAT histogram

The interarrival delay histogram for a set of DSL hosts in San Francisco (Fig.12) is highly periodic, with 3.313 ms between successive spikes. This periodicity is due to ATM cell framing over the DSL links [24]. 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 offered in [25] although it does not specify which encapsulation achieves that rate.

The interarrival histograms for 1500 byte packets[8] (Fig.12) have two large spikes, one at 35 ms and another at 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 384 Kbps link upstream from a constellation of 128 Kbps sources. This alternative is less likely though, since multiplexing devices normally have much higher bitrates than individual users.

To corroborate this analysis we took a set of DSL hosts with the same provider as before, but located 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. This trace 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 by taking

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

[8]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 [12], thus the delay is attributed to the first packet.

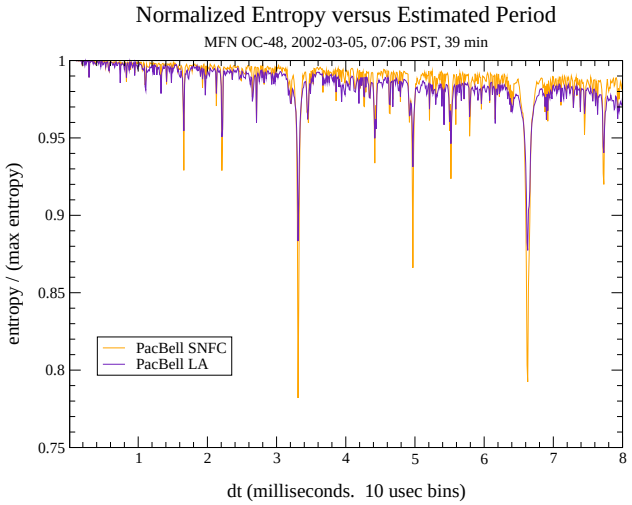

Fig. 13.

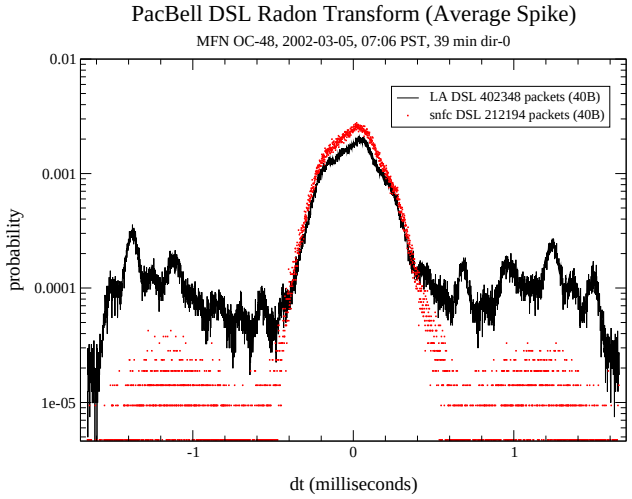

Fig. 14.

the Radon transform[9] of the data over different periods s,

$$p(d,s) = \sum_{k=0}^{\infty} p(d+ks), \qquad 0 \le d < s$$

and computing the entropy for each s as in Sect.III. 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 very well for our data. Both the Los Angeles and San Francisco data sets have sharp downward spikes at $t_{cell} = 3.313$ ms (Fig.13.)

After selecting $t_{cell} = 3.313$ ms as the delay quantum corresponding to an ATM cell, we investigated the inter-packet delay histograms for SF and LA data sets (Fig. 14). We found that they share similar traits, which is expected, since the ISP likely uses similiar infrastructure design and equipment in both cities.

[9]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.

A. *Active measurements*

To better determine what types of effects were causing the periodicity in the backbone trace, we made active measurement 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 which were received on a 1 Ghz dual processor FreeBSD machine at the university with a 100 Mbps Ethernet access link.[10]

The data gathered shows similiar 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.15) 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 (Sec.III.) Our active measurement lasted over 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, with each burst containing packets of the same size.[11] We recorded interarrival times for each burst using Pentium `rdtsc` (read timestamp counter) instruction [26] with nominal resolution of 1 processor cycle, i.e. 1 ns.

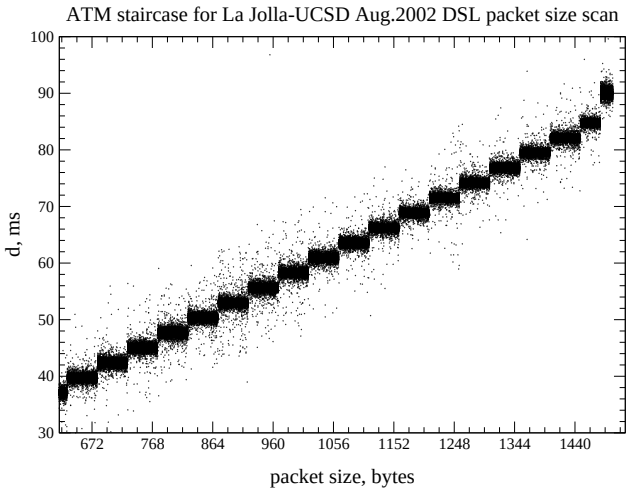

Fig. 15.

Fig.15 shows the staircase with 48-byte stairs in size-delay space (similar to that of Fig.2), 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.[12] We can model the ATM staircase with the formula:

$$d_a(b) = \text{ceil}((b+6+18+16)/48) \cdot 53 \cdot 8/160\text{ms}$$

[10]`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.

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

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

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.[13]

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 througput is around 130 Kbps.

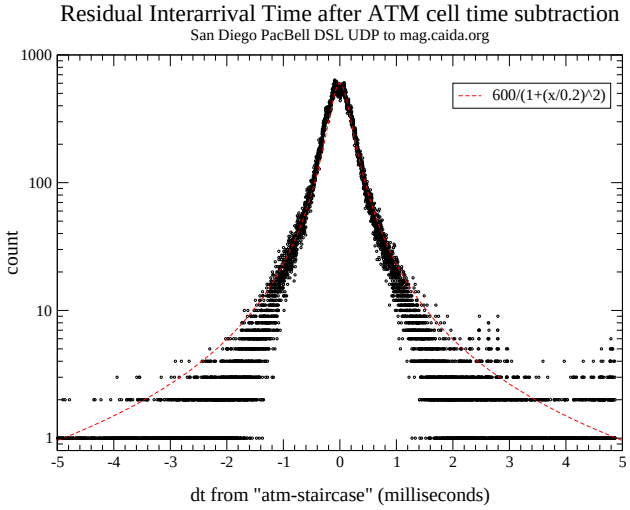

Fig. 16.

The shape of the residual delay distribution $d_r = d - d_a$ for this DSL experiment (Fig.16) is symmetric with decay slower than exponential as we move away from the center. It is reminiscent of Levy-stable distributions [27].[14]

We also tried to analyze several other IP addresses which had `dsl` in their 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 as observed for the studied Californian DSL provider. We leave the study of international DSL for future work.

V. Broadband connections: Cable

In this section we analyze traffic from cable modems. The data was gathered from the same monitor as above, on March 5, 2002. Packet interarrival times are measured as in Sect.IV. Only flows of over 1 MB are considered.

[13]The prominent jump at the upper right end of Fig.15 reflects packet fragmentation caused by host MTU of 1480 bytes. With IP payload $b > 1480$ bytes, encapsulated first fragment contains 32 cells (40+1480 bytes) and 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, resp. 2.65 ms in Fig.16.

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

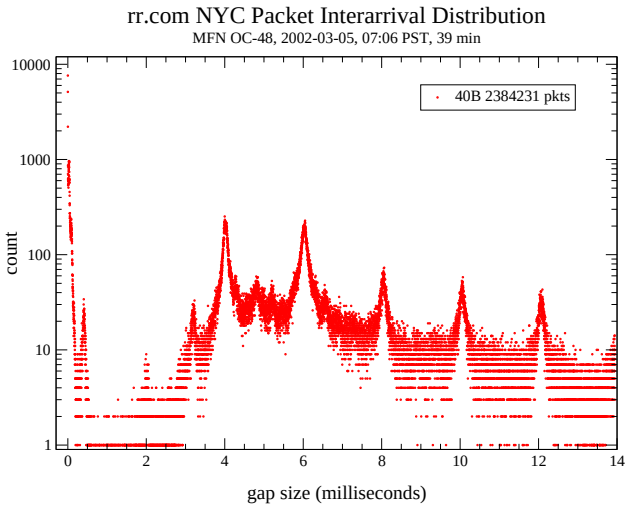

Fig. 17.

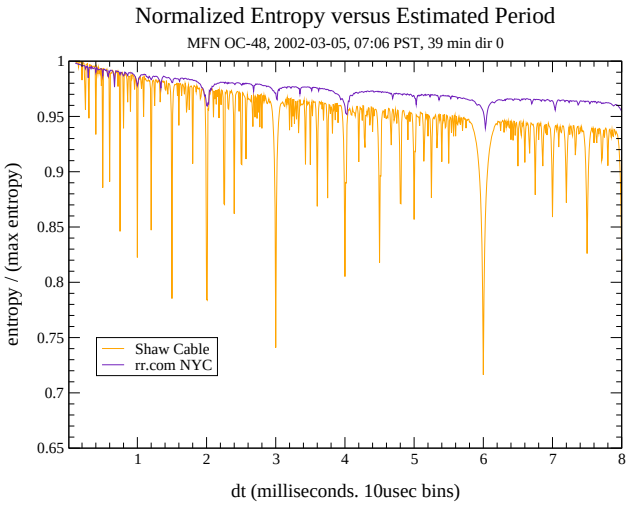

Fig. 18.

The packet interarrival histogram for a set of hosts on the `rr.com` network in NYC (Fig.17) is periodic, but the period is different from the DSL traffic; it is close to 2 ms.

The primary cause of delay quantization is the time division multiplexing scheme used in cable modems which allows for upstream transmissions only in predetermined slot times [28], [29]. These slot times must be reserved through negotiation with the cable modem termination system (CMTS) using a system of reserved slots with collision detection in a manner similar to Ethernet. Detailed descriptions can be found in the DOCSIS specifications [30].

We took the Radon transform of the data over a wide range of periods. We then computed the entropy of the results. As in Sect.III and IV, when the spikes of the distribution are lined up we would expect the entropy to be lower. The minimum in the `rr.com` data was around 2 ms. The Shaw Cable data had spikes at 6, 3, 2, and 1.5 ms.[15] This is a rather computationally expensive algorithm, so we used 10 usec bins.

[15]A similar pattern with a 2 ms period and stronger spikes every 6 ms is observed in ATT Broadband data in Fig.21.

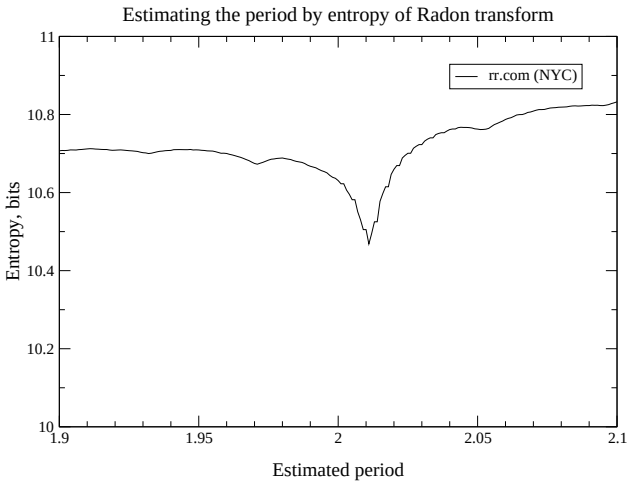

Fig. 19.

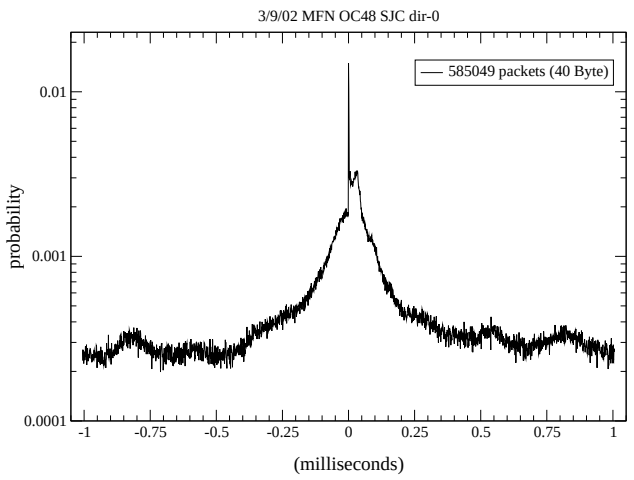

Fig. 20.

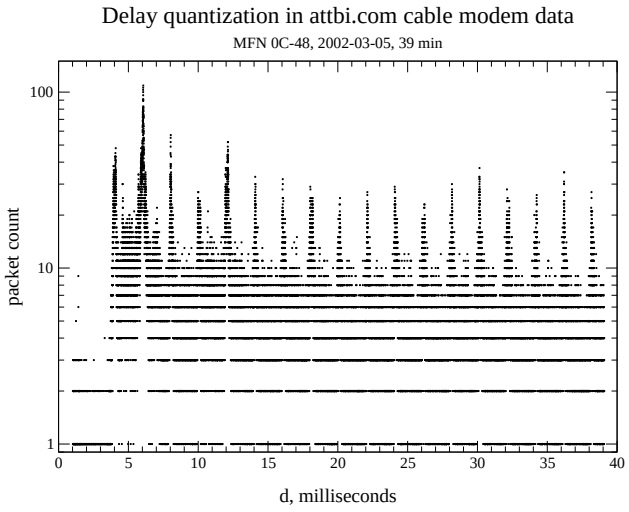

Fig. 21.

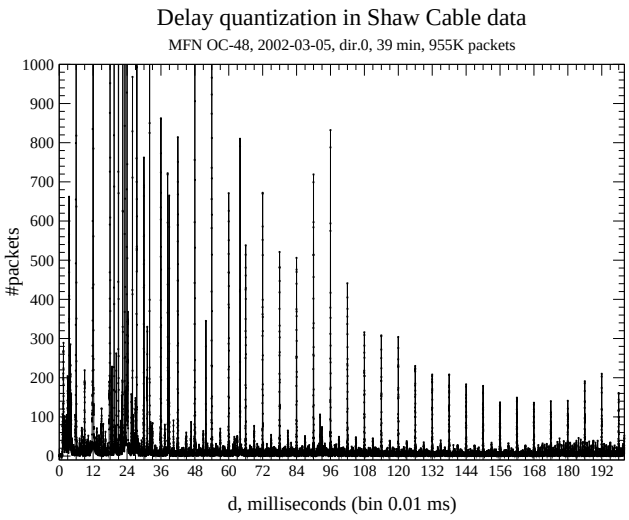

Fig. 22.

The plots for normalized entropy for Shaw Cable (WA) and Road Runner (NYC) networks have drops at multiples of 1 ms. The Shaw Cable has multiple spikes associated with rational numbers (p/q where p, q are integers) of milliseconds with higher spikes for smaller values of q.

To determine the exact period, we made a calculation using 1 usec bins (Fig.19) over a range close to the period of 2 ms estimated from Fig.18. The minimum-entropy period of the `rr.com` data is 2.011 ms. Once again (as in Sect.III) entropy has a 'beak' with a downward-looking cusp that makes the resulting estimate quite sharp. We plan to investigate in future how close these estimates match characteristics of broadband networks.

By taking the slant stack of the data with the determined period, we get an 'average spike' in the delay distribution. Fine structure (splitting of the central spike in two) present in some of the peaks becomes apparent. This stucture is likely due to the effects of cross traffic at other links. It may be possible to infer details of the path from this structure. (The distance between the two thin peaks is about 50 usec.)

The interarrival times histogram from sources in the Shaw Cable network in WA (Fig.22, 23) has strong preference for values of $d = 6k$ ms. The remaining spikes tend to line up at multiples of 1.5 and 2 ms.

The Radon transform for the Shaw Cable data is similarly the result of the superposition of 2 distinct patterns (Fig.24 shows only central spike since it has linear scale.). It is possible that the Shaw Cable network is composed of two sets of sources whose interarrival distributions have a period of 1.5 ms, resp. 2 ms. Thus they would line up every 6 ms, explaining the strength of the spike at exactly a period of 6 ms. We have yet to determine what is causing two series of spikes. Other than that, Shaw Cable's min-entropy Radon transform in Fig.24 closely resembles the rate-limiting type (triangular in logarithmic Y-coordinate, i.e. close to Laplace distribution although with multiple spikes of similar shape) that we saw in Fig.3.

Comparison of min-entropy transforms in Fig.24 also shows that the cable modems' average spikes decrease faster than DSL spikes, i.e. delay quantization in cable is more pronounced.

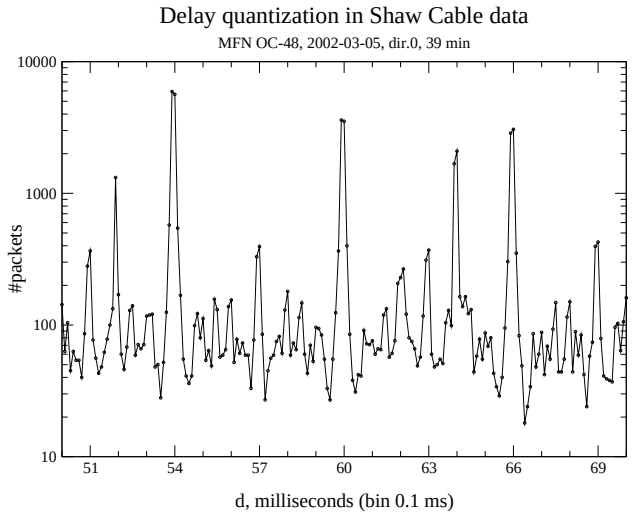

Fig. 23.

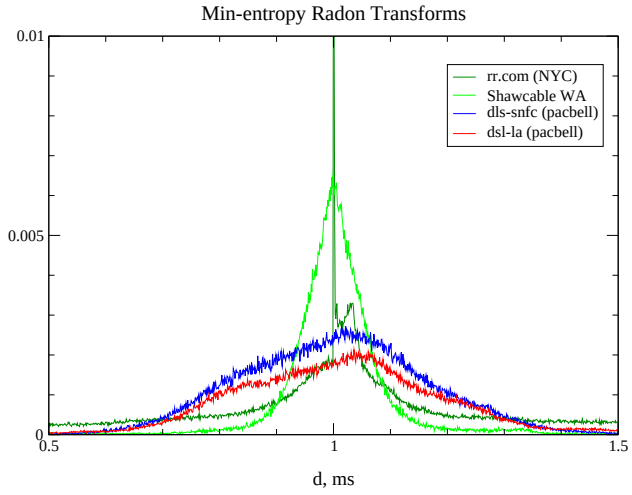

Fig. 24.

VI. Discussion

We found that entropy minimization-based algorithms for idenifying delay quanta need to take into account two issues: smooth variability in baseline entropy and the presence of "resonances", rational multiples of 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 Fig.13, 18. However, this division can result in closer values for minimums of normalized entropy so that the choice between periods becomes non-unique (e.g. Fig.18 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 data (e.g. in Fig.17 at the smallest possible values of d) can make the average spike corresponding to the entropy minimum s_{me} depend mostly upon the contribution of one largest spike rather than the totality of all spikes in the data. The needle in the center of Fig. Fig.20 reflects only the shape of the first spike. Conclusions based on the average spike may only 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.

In [16] we show how one can mitigate the disparity in a special case of exponential decay in the spikes' envelope by dividing out the exponential factor. 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.

VII. 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.

To promote network spectroscopy from a computational art to a routine occupation, it will be neccessary to enable automated analysis of measurement data by the algorithms presented here. We intend to develop these means in the future. To that end, we plan 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 Fourier-based identification method which we expect to be more sensitive to the quantitative details of data clustering than binned entropy. We also plan to extend the analysis presented here to other connection technologies, including major brands of wireless (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 the potential for measurement artifacts to distort the data and affect the results of interpretation. We plan to evaluate timing precision of available monitoring equipment and to cross-check timestamping done by products from different vendors as one of our next goals.

VIII. Acknowledgements

The idea of network spectroscopy crystallized during the first author's visit in March-June 2002 to UCLA's Institure for Pure and Applied Mathematics Large-scale communication Networks Program. Thanks to 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 Ken Keys of CAIDA for his help with Coral software suite, to Stephen Donnelly of University of Waikato/Endace for discussions of SONET timing, and to Thomas Karagiannis of Riverside for related discussions and for reading preliminary drafts of this paper and follow-up work. The data analysis presented here would have never been possible were it not for the effort of those who supplied us with the packet traces: David Moore of CAIDA who provided the university trace, and Joerg Micheel of University of Waikato/NLANR who provided backbone measurement setup and collected several traces used in this and other papers.

References

[1] A. Broido, E.Nemeth, and kc claffy, "Spectroscopy of DNS update traffic," Submitted for publication, Nov.2002.

[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] APNIC, ARIN, and RIPE, "Whois service," www.ripe.net/ripencc/pub-services/db/whois/whois.html.

[11] A. Broido, E.Nemeth, and kc claffy, "Packet arrivals on rate-limited Internet links," CAIDA, Nov.2000, www.caida.org/∼broido/coral/packarr.sigm20001102.html.

[12] Joerg Micheel, Stephen Donnelly, and Ian Graham, "Precision timestamping of network packets," Nov 2001, Proceedings of ACM Sigcomm Internet Measurement Workshop.

[13] Stephen Donnelly, "High precision timing in passive measurements of data networks," June 2002, Ph.D. thesis, University of Waikato, Hamilton, New Zealand.

[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," December 2002, 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] Juha Heinanen, "Multiprotocol Encapsulation over ATM Adaptation Layer 5, RFC1483," July 1993.

[20] Claude Shannon, *The mathematical theory of communication*, 1949, Urbana, Univ.Illinois Press.

[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] John Navas, "Navas Cable Modem/DSL Tuning Guide," 2002, `navasgrp.home.att.net/index.htm`.

[25] "Sbc/pacbell dsl networking plans," www.lanlogic.net/connectivity/dsl/sbc-dslplans.asp.

[26] "Using the RDTSC instruction for performance monitoring," August 1997, Pentium II processor application notes, cedar.intel.com/software/idap/media/pdf/rdtscpm1.pdf.

[27] Paul Lévy, "Theorie de l'addition des variables aleatoires," Gauthier-Villiers, Paris, 1937.

[28] Rolf V. Ostergaard, "What is a Cable Modem? (Tutorial)," 1999, `www.cable-modems.org`.

[29] Robin D. Walker, "Cable Modem Troubleshooting Tips," 2002, `homepage.ntlworld.com/robin.d.h.walker/cmtips/index.html`.

[30] "DOCSIS," 2002, www.cablemodem.com/specifications/.