Measurement and Analysis of Intraflow Performance Characteristics of Wireless Traffic

Dimitrios P. Pezaros, Manolis Sifalakis, and David Hutchison

Computing Department, Infolab21, Lancaster University, Lancaster, UK, LA1 4WA
{dp, mjs, dh}@comp.lancs.ac.uk

Abstract. It is by now widely accepted that the arrival process of aggregate network traffic exhibits self-similar characteristics which result in the preservation of traffic burstiness (high variability) over a wide range of timescales. This behaviour has been structurally linked to the presence of heavy-tailed, infinite variance phenomena at the level of individual network connections, file sizes, transfer durations, and packet inter-arrival times. In this paper, we have examined the presence of fractal and heavy-tailed behaviour in a number of performance aspects of individual IPv6 microflows as routed over wireless local and wide area network topologies. Our analysis sheds light on several questions regarding flow-level traffic behaviour: whether burstiness preservation is mainly observed at traffic aggregates or is it also evident at individual microflows; whether it is influenced by the end-to-end transport control mechanisms as well as by the network-level traffic multiplexing; whether high variability is independent from diverse link-level technologies, and whether burstiness is preserved in end-to-end performance metrics such as packet delay as well as in the traffic arrival process. Our findings suggest that traffic and packet delay exhibit closely-related Long-Range Dependence (LRD) at the level of individual microflows, with marginal to moderate intensity. Bulk TCP data and UDP flows produce higher Hurst exponent estimates than the acknowledgment flows that consist of minimum-sized packets. Wireless access technologies seem to also influence LRD intensity. At the same time, the distributions of intraflow packet inter-arrival times do not exhibit infinite variance characteristics.

Keywords. LRD, Hurst exponent, heavy-tailed distribution, ACF

1. Introduction

Seminal measurement studies during the last fifteen years have demonstrated that data communications networks' traffic is self-similar (statistically fractal) in nature remaining bursty over a wide range of timescales. These findings advocated that statistical properties of the (aggregate) network traffic, when viewed as time series data, remain similar irrespective of the time scale of observation, and were in sharp contrast with the up till then commonly employed Poisson and Markovian models which were based on exponential assumptions about the traffic arrival process. A characteristic of self-similar processes is that they often exhibit long memory, or

Long-Range Dependence (LRD), signifying that their current state has significant influence on their subsequent states far into the future. Consequently, values at a particular time are related not just to immediately preceding values, but also to fluctuations in the remote past. Hence, high variability in the behaviour of self-similar processes is preserved over multiple time scales. Pioneering work has focused on the measurement and characterisation of LAN [13], WAN [17], and transport/application-specific traffic [6], having the traffic arrival process at a single network (edge) point as the common primary quantity of interest. The revealed LRD properties of aggregate network traffic have been subsequently linked to heavy-tailed, infinite variance phenomena at the level of individual source-destination pairs, represented by ON/OFF sources and packet trains models whereby a source alternates between active (ON-period) and idle (OFF-period) states [22][23]. These are in turn attributed to distributed systems' objects and file sizes being heavy-tailed, a property that has been shown to be preserved by the protocol stack and mapped to approximate heavy-tailed busy periods at the network layer [15][16]. However, the existence (or otherwise) of self-similar and/or heavy-tailed behaviour within performance facets of individual microflows has not received much attention to date. This is partly due to the majority of individual connections in the Internet being short-lived, and therefore their behaviour over multiple time scales can be studied only through aggregation. At the same time, the vast majority of bytes over the Internet are carried within a relatively small number of very large flows which are sufficiently long-lived for the temporal evolution of their intraflow properties to be investigated [4]. It is henceforth feasible to examine whether the self-similar characteristics of aggregate traffic resulted from the superimposition of numerous ON/OFF sources are also manifested at the level of long-lived individual traffic flows. Characterising the long-term flow behaviour and revealing possible burstiness preservation properties therein can prove instrumental for network resource management and accountability purposes, since end-to-end flows are the primary entity subjected to open and closed-loop network control. Similar to proposals advocating small buffer capacity/large bandwidth resource provisioning when input traffic is self-similar, relevant characterisation of end-to-end flow behaviour and performance aspects therein can form the basis for designing adaptive flow control algorithms to operate at multiple timescales and take into consideration potential correlation between distant events/values throughout the lifetime of a flow. From a measurement point of view, the comparative analysis of temporal flow behaviour is also important since it can reveal certain idiosyncrasies potentially linked to the operation of the different transport control algorithms. Differences in the long-term behaviour of diverse traffic flows can hint to additional causality relationships between burstiness intensity and flow control, as well as the levels of traffic multiplexing in the network.

Likewise, packet delay is one of the most commonly employed metrics to assess the service quality levels experienced by an end-to-end flow. In contrast to traffic arrivals, delay indicates how traffic is routed between two network nodes and is among the primary performance aspects whose absolute value and temporal variations (in either the unidirectional or the round-trip representation) are attempted to be controlled by transport mechanisms and kept within certain ranges depending on individual applications' requirements. Links between network traffic self-similarity and the temporal intraflow delay behaviour can identify the degree of penetration of high variability to different facets of network performance, and the relative level of causality between performance, end-to-end flow control, and traffic multiplexing in the network. A few studies on fractal analysis of packet delay have reported non-stationary LRD in round-trip delay of synthetic UDP traffic [3], and in aggregate NTP/UDP flows [14], yet the burstiness preservation relationships between the different transport mechanisms and the unidirectional contributors of the end-to-end delay of individual flows have not been investigated.

In this paper, we have quantified the burstiness preservation properties of a set of end-to-end unidirectional IPv6 traffic flows routed over IEEE 802.11 and GPRS service networks, two media which themselves exhibit highly variable performance characteristics [10][5]. We have comparatively examined the presence of Long-Range Dependence (LRD) in the intraflow traffic arrival process and in per-packet unidirectional delay, and we have investigated the extent to which transport control, packetisation, and wireless access mechanisms influence its intensity. In addition, we have analysed the tail behaviour of packet inter-arrival times at the individual flow level, and revealed the absence of infinite variance phenomena therein. In section 2 we provide the definition and mathematical formulation of self-similarity and LRD, their interrelation, and a brief discussion on the estimators used to identify and quantify LRD on empirical time series. Section 3 includes an outline our measurement methodology and a description of the wireless experimental infrastructure over which measurements were conducted. In section 4 we present and discuss in detail the results of measurement analysis, and we provide possible explanations and interpretations of our findings. We comment on the similarity between LRD in the per-flow traffic and unidirectional delay, and on the different levels of LRD exhibited by diverse application flows and over the different wireless topologies. We also compare and contrast the tail behaviour of per-flow packet inter-arrival times to heavy-tailed distributions. Section 5 concludes the paper.

2. Self-Similarity and Long-Range Dependence: Definitions and Estimation

A stochastic process or time series *Y(t)* in continuous time $t \in \mathbb{R}$ is self-similar with self-similarity (Hurst) parameter $0 < H < 1$, if for all $\alpha > 0$ and $t \geq 0$,

$$Y(t) =_d a^{-H} Y(\alpha t).$$

Self-similarity describes the phenomenon of a time series and its time-scaled version following the same distribution after normalizing by α^{-H}. It is relatively straightforward to show [2] that this implies that the autocorrelation function (ACF) of the stationary increment process $X(t) = Y(t) - Y(t-1)$ at *lag k* is given by

$$\rho(k) = \frac{1}{2}((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}),\ k \geq 1.$$

In addition, for $0 < H < 1,\ H \neq \frac{1}{2}$, it can be shown that $\lim_{k \to \infty} \rho(k) = H(2H-1)k^{2H-2}$,

and in particular for the case $0.5 < H < 1$, *ρ(k)* asymptotically behaves as $ck^{-\beta}$ for

$0 < \beta < 1$, where $c > 0$ is a constant, $\beta = 2 - 2H$ [2][16]. This implies that the correlation structure of the time series is asymptotically preserved irrespective of time aggregation, and the autocorrelation function decays hyperbolically which is the essential property that constitutes it not summable:

$$\sum_{k=1}^{\infty} \rho(k) = \infty.$$

When such condition holds, the corresponding stationary process *X(t)* is said to be *Long-Range Dependent (LRD)*. Intuitively, this property implies that the process has infinite memory for $0.5 < H < 1$, meaning that the individually small high-lag correlations have an important cumulative effect. This is in contrast to conventional short-range dependent processes which are characterised by an exponential decay of the autocorrelations resulting in a summable autocorrelation function. LRD causes high variability to be preserved over multiple time scales and is one of the manifestations of self-similar processes alongside non-summable spectral density for frequencies close to the origin and slowly decaying variances. This latter characteristic of self-similar and LRD processes can be disastrous for classical tests and prediction of confidence intervals. The variance of the arithmetic mean decreases more slowly than the reciprocal of the sample size, behaving like $n^{-\beta}$ for $0 < \beta < 1$, instead of like n^{-1} which is the case for processes whose aggregated series converge to second-order pure noise [2][13]. Therefore, usual standard errors derived for conventional models are wrong by a factor that tends to infinity as the sample size increases. Two theoretical models that have been used to simulate LRD is the fractional Gaussian noise (fGn) which is the stationary increment process of fractional Brownian motion (fBm), and fractional ARIMA processes that can simultaneously model the short and long term behaviour of a time series.

A number of estimators [21][13] have been extensively used in multidisciplinary literature for LRD detection and quantification by estimating the value of the Hurst exponent; as $H \rightarrow 1$ the dependence is stronger. They are classified in time-domain and frequency-domain estimators, depending on the methodology they employ to estimate *H*. Time-domain estimators are based on heuristics to investigate the evolution of a statistical property of the time series at different time-aggregation levels. They are hence mainly used for LRD detection, rather than the exact quantification of the phenomenon. Frequency-domain estimators focus on the behaviour of the power spectral density of the time series. In this paper, we have employed two time-domain estimators (the *aggregated variance* and the *rescaled adjusted range* methods) to detect whether our measured quantities exhibit LRD characteristics ($H > 0.5$), and we have subsequently focused on the more robust frequency-domain *Whittle* estimator for the exact LRD quantification. Whittle is a maximum likelihood type estimate which is applied to the periodogram of the time series and it provides an asymptotically exact estimate of *H* and a confidence interval. However, it presupposes that the empirical series is consistent with a specific process (e.g. fGn) whose underlying form must be provided, hence its use with time series that have already been shown to be LRD (by other means) is strongly advisable. Wavelet-based LRD estimation [1] has also been developed whose accuracy is comparable to Whittle, yet it has been lately shown to consistently overestimate the Hurst exponent on synthesized LRD series and hence Whittle was preferred for this study [12].

3. Measurement Methodology and Experimental Environment

In-line measurement [18] has been employed to instrument a representative set of IPv6 traffic flows as these were routed over diverse wireless topologies during one week, in November 2005 [19]. The technique exploits extensibility features of IPv6, to piggyback measurement data in Type-Length-Value (TLV) structures which are then encapsulated within an IPv6 *destination options* extension header and carried between a source and a destination. Being encoded as a native part of the network-layer header, inline measurement is potentially applicable to all traffic types carried over the IPv6 Internet infrastructure. Destination options extension headers in particular, are created at the source and are only processed at the (ultimate) destination node identified in the destination address field of the main IPv6 header. Hence, their presence does not negatively impact the option-bearing datagrams at the intermediate forwarding nodes [18]. For the purposes of this study, 32-bit Linux kernel timestamps were encoded in an appropriate TLV structure to record time T_{DEP} immediately before a packet is serialised at the NIC of the source IPv6 node, and time T_{ARR} as soon as the packet arrives in the destination IPv6 node's OS kernel. The intraflow traffic arrival process has been calculated as the number of packets (or bytes) arriving at the destination within disjoint subintervals throughout the flow duration. The end-to-end unidirectional delay for a given packet P is calculated as $D = T^P{}_{ARR} - T^P{}_{DEP}$. Moreover, the inter-arrival time between two successive packets P_i and P_{i+1} is computed as $IA = T_{ARR}^{P_{i+1}} - T_{ARR}^{P_i}$. For the purposes of the delay measurement the two end-systems synchronised using the Network Time Protocol (NTP) with a common stratum 1 server through additional high-speed wired network interfaces, in order to avoid NTP messages competing with the instrumented traffic over the bottleneck wireless links. The NTP daemon was allowed sufficient time to synchronise prior to the experiments until it reached a large polling interval. The offset reported by NTP was always at least one order of magnitude smaller with respect to the minimum one-way delay observed. All the delay traces were empirically examined against negative values as well as against linear alterations (trend) of the minimum delay over time. None of these offset/skew-related phenomena were experienced.

Instrumented traffic consisted of moderate-size bulk TCP transfers and CBR UDP video streaming flows. Measurements have been conducted end-to-end over two diverse wireless service networks over the Mobile IPv6 Systems Research Laboratory (MSRL) infrastructure. MSRL includes a wireless cellular network as well as a combination of 802.11 technologies and it comprises a real service infrastructure, as shown in **Fig. 1**. The measurements were carried out between a host machine connected to MSRL's wired backbone network (Linux 2.4.18; Intel 100BaseT adapter) and a host machine with multiple wireless interfaces (Linux 2.4.19; NOKIA D211 combo PCMCIA 802.11b/GPRS adapter), connected through the 802.11b/g campus-wide network and through the GPRS/GSM W-WAN network. The W-LAN infrastructure is part of Lancaster University campus wireless network, and includes 802.11b and 802.11g. Although the nominal speed for 802.11b is 11Mb/s, it has been observed that due to interference with other appliances operating at the same frequency band (2.4 GHz), the cards often fallback to 5.5, 2, and 1 Mb/s.

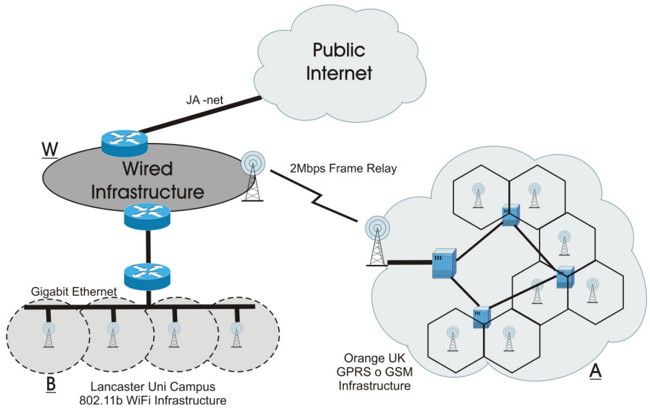

Fig. 1. Wireless experimental environment (MSRL infrastructure)

The W-WAN network is the Orange UK GPRS/GSM service network, practically allowing for speeds of up to 20/50Kb/s (up/downlink), due to asymmetric slot allocation. Connectivity between Orange UK and the MSRL backbone is served by a 2Mb/s wireless Frame Relay point-to-point link.

4. Measurement Analysis and Results

Packet-level measurements were taken upon arrival of each packet to its destination, hence at irregular time instants. In order to covert the traces to time series data, they were discretised into equally-sized bins based on normalised packet arrival time. Unidirectional delays and inter-arrival times of multiple packets arriving within each bin were averaged, and the mean values were considered for the particular bin. Although this process inevitably smooths out short-term variations, bin size was carefully selected for each flow to contain as few packets as possible, while at the same time avoiding empty bins. The time series' lengths varied from 2^9 to 2^{12} which has been reported sufficient for Hurst exponent estimates with less than 0.05 bias and standard deviation [7]. A number of sanity tests and calibration measures have been employed to tune the LRD estimation process by eliminating time series effects such as periodicity, trend and short-range correlations which are known to constitute LRD estimation error prone. Trends and non-stationarities have been checked against during pre-processing, while periodicity and short-term correlations have been neutralised using the randomised buckets method to perform internal randomisation to the signal [9]. Time-domain (detection) estimators which operate on the aggregated time series can suffer from the limited number of samples available at high aggregation levels. We have employed oversampling with digital low-pass filtering to increase the sampling rate of the time series and hence refine their estimates. For further details regarding techniques for tuning the LRD estimation process and their effects, the interested reader is referred to [12][20].

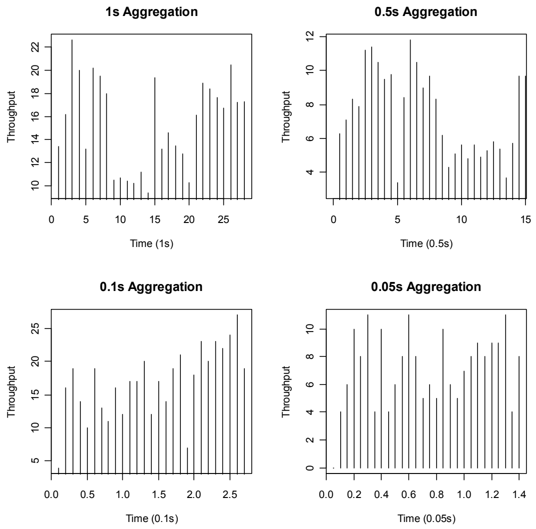

Fig. 2. Stochastic Self-Similarity – burstiness preservation across time scales of 50 ms, 100 ms, 500 ms, and 1000 ms for the bulk TCP data path over the W-LAN network

LRD behaviour in traffic arrivals and unidirectional end-to-end delay

The two intraflow properties analysed for LRD behaviour were the traffic arrival process and the unidirectional end-to-end packet delay. These were comparatively examined for TCP data, TCP reverse, and CBR UDP flows routed over the two diverse wireless topologies. We have used the LRD detection (time-domain) estimators on all time series data which all reported that traffic arrivals and one-way delay exhibit LRD at various intensities, and produced Hurst exponent estimates $0.5 < H < 1$. We subsequently applied the Whittle estimator on the data to quantify the levels of LRD with high confidence. **Fig. 2** indicatively shows the burstiness preservation of the data path of a bulk TCP flow over the W-LAN topology at varying time scales, demonstrating the absence of a characteristic size of a burst. Traffic throughput is shown in packets which are all MSS-sized and hence byte-throughput is linearly identical. The upper left plot shows a complete presentation of the time series using one-second bins. Then, the bottom left plot shows the same time series whose first 3-second interval is "blown up" by a factor of ten, and the truncated time series has time granularity of 100 ms. Likewise, the rightmost plots show parts of the time series with an equivalent number of samples for time granularities of 500 and 50 ms, respectively. Overall, the plots show significant bursts of traffic at different levels spanning almost three orders of magnitude. **Table 1** and **Table 2** show the Whittle estimate (and 95% confidence interval) of the Hurst exponent for traffic arrivals and unidirectional delay, respectively, for all the different flows routed over the W-LAN and W-WAN networks. Their comparative examination yields some very interesting observations.

Table 1. Hurst exponent estimates – Whittle method: Traffic arrivals

Microflow	Whittle Estimator
	H est. & 95% C.I.
TCP data [W-LAN]	0.714 ± 0.005
TCP data [W-WAN]	0.554 ± 3.47e-05
TCP reverse [W-LAN]	0.584 ± 0.004
TCP reverse [W-WAN]	0.534 ± 0.001
UDP [W-LAN]	0.534 ± 3.37e-05
UDP [W-WAN]	0.697 ± 0.001

Table 2. Hurst exponent estimates – Whittle method: Unidirectional packet delay

Microflow	Whittle Estimator
	H est. & 95% C.I.
TCP data [W-LAN]	0.739 ± 0.025
TCP data [W-WAN]	0.599 ± 0.15
TCP reverse [W-LAN]	0.552 ± 0.007
TCP reverse [W-WAN]	0.528 ± 0.014
UDP [W-LAN]	0.687 ± 0.003
UDP [W-WAN]	0.742 ± 0.20

It is evident in both phenomena (tables) that the majority of the unidirectional flows over the two media independently, show marginal to moderate LRD intensity with Hurst exponent values for some of them close to those of short-range dependent processes, differing by less than 0.1. This fact reinforces the argument of LRD being intensified by the aggregation of traffic. However, there are cases of individual flows which suggest dependency between LRD, traffic type and wireless medium. For bulk TCP data over W-LAN, both traffic arrival (whose burstiness preservation was shown in **Fig. 2**) and end-to-end delay exhibit considerable LRD manifested by Hurst values greater than 0.71. This is in contrast to the same type of traffic routed over W-WAN for which both properties assume marginal intensity values less than 0.6. The opposite behaviour with respect to the two wireless media is observed for constant bit rate UDP traffic. Over W-LAN, UDP traffic does not assume considerable LRD, whereas moderate intensity is suggested for UPD over W-WAN with an estimated Hurst exponent close to 0.7. The same relationship holds for the packet delay behaviour of the UDP flows over the two media, although the absolute Hurst estimates are in both cases larger than those of the traffic process. The acknowledgment path of bulk TCP connections over both media is characterised by smaller intensity than the corresponding data path, and overall marginal LRD.

When comparing the Hurst estimates of the per-flow traffic behaviour with those of the corresponding unidirectional delay, it is apparent that there is a considerably close agreement between their LRD intensities. This implies that although traffic arrival process and unidirectional delay are metrics describing different aspects of network dynamics (the former describes how traffic is delivered at a single network node while the latter describes how traffic is routed between two nodes), they are both influenced by common sets of parameters. Hurst estimates of the two processes for each flow/medium combination lie within a range which is smaller than the wider 95% confidence interval of the two. The only exception is the TCP acknowledgment path over the W-LAN topology, where the Hurst estimates for traffic and packet delay differ by 0.032 while the wider 95% C.I. of the two (delay) is 0.007. Overall, Whittle performs better on the traffic arrival process by producing narrow 95% C.I.s for the

Hurst estimates on the order of 10^{-3} or better. For unidirectional packet delay the corresponding C.I.s are on the order of 10^{-2} or better, however in two cases, the width of the 95% C.I.s can put even the existence of marginal LRD under doubt.

It is well accepted that traffic self-similarity arises through the aggregation of multiple traffic sources since its causality relationship with the heavy-tailed property of the ON/OFF sources model is based on limit theorems for large number of sources and large temporal intervals [22][23]. However, the per-flow indications of LRD for traffic as well as for the packet delay raise some interesting issues. The two wireless configurations over which the measurement was conducted can be safely assumed to operate as access networks where clients are mainly downloading content and not uploading, hence the download path is more congested than the reverse direction. At the same time, the W-LAN network was lightly utilised during the time of the experiments, as this was indicated by APs' client association logs. Therefore, the stronger LRD intensity exhibited by the TCP data flow over W-LAN yields some dependency between LRD and the packetisation/congestion control algorithm operating on the flow, irrespective of traffic aggregation over the medium. The fact that similar LRD levels are not seen for the TCP data path over W-WAN hints towards relationship between traffic behaviour and link-local delivery mechanisms. The dense protocol stack of GPRSoGSM which to some extent replicates TCP's reliable transmission seems to neutralise the effect that transport-layer congestion control has on the long-range behaviour of traffic. The higher Hurst estimates produced for both TCP data and UDP flows, as opposed to the (lower) estimates of the TCP reverse flows over both wireless media signifies dependency between LRD behaviour and packet size. It is worth noting that TCP data and UDP flows consisted of constant-size packets of 1440 and 544 bytes, respectively, while TCP reverse flows consisted of (mainly) 56 and (fewer) 84-byte packets. All sizes exclude network and link-layer headers. Apart from the Hurst exponent estimates, accompanying indisputable evidence regarding long-memory behaviour of per-flow data are computed sample statistics such as ACF which exhibit nontrivial correlations at large lags. **Fig. 3** indicatively shows the correlation structures of traffic arrivals from a selection of flows with different LRD intensity whose estimates appear in **Table 1**. The ACFs demonstrate the different correlation structures between time series with considerably different Hurst estimates. The plots also show the effect of the randomised buckets with internal randomisation which were employed to neutralise short-range correlations and periodicities. It is evident that ACFs follow asymptotic decay at various levels, which is unsurprisingly more intense for flows with larger Hurst estimates. It is interesting to note that even for the TCP reverse flow whose estimated Hurst value is considerably small (<0.6), correlations seem non-degenerate since they mostly remain above the 95% C.I. of zero.

Tail behaviour of intraflow packet inter-arrival times

The heavy-tailed property of the transmission (or the idle) times of individual sources has been characterised as the main ingredient needed to obtain LRD characteristics ($H>0.5$) in aggregate traffic [22][23]. Heavy-tailness has since then been reported in a number of facets of traffic behaviour, including (web) file sizes, transfer times, burst lengths and inter-arrival times [6][8][17].

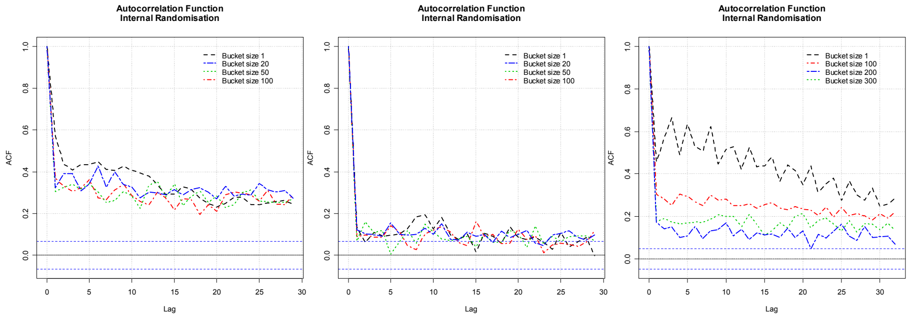

Fig. 3. Autocorrelation function (ACF) of TCP data [*H=0.714*] and TCP reverse [*H=0.584*] traffic over W-LAN, and UDP traffic [*H=0.697*] over W-WAN

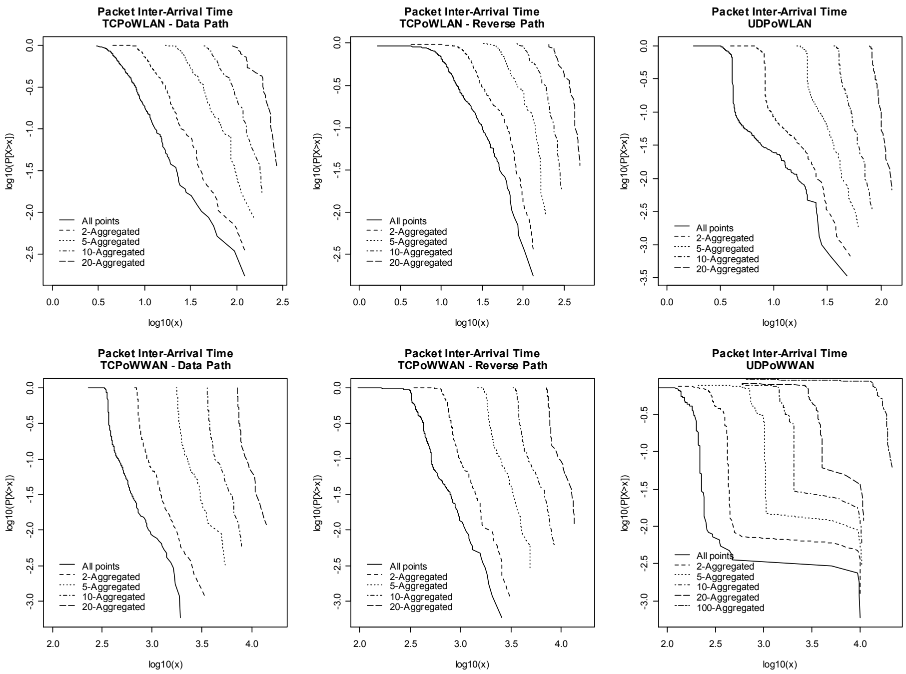

Fig. 4. LLCDs of intraflow packet inter-arrival times for TCP data, TCP reverse, and UDP traffic over W-LAN and W-WAN configurations

A distribution is heavy-tailed if $P[X > x] \sim x^{-\alpha}$, as $x \to \infty$, $0 < \alpha < 2$. That is, regardless of its behaviour for small values of the random variable, its asymptotic shape is hyperbolic. For $1 \leq \alpha < 2$, the distribution has infinite variance, whereas for $0 < \alpha < 1$, it has infinite mean as well. The tail index α can be empirically estimated using the log-log complimentary distribution (LLCD) plot and calculating its slope through least-squares regression for large values of the random variable above which the plot appears to be linear. We have used a test (also used in [6]) based on the

Central Limit Theorem (CLT) to examine whether the packet inter-arrival times within our measured traffic flows exhibit infinite variance, and hence heavy-tailed characteristics. According to CLT, the sum of a large number of i.i.d. samples from any distribution with finite variance will tend to be normally distributed. Hence, for a distribution with finite variance, the slope of the LLCD plot of the *m-aggregated* dataset will increasingly decline as *m* increases, reflecting the distribution's approximation of a normal distribution. On the contrary, if the distribution exhibits infinite variance the slope will remain roughly constant with increasing aggregation levels (*m*). **Fig. 4** shows the CLT test for various aggregation levels applied to the packet inter-arrival times of all flows over the two wireless topologies. It is evident that for increasing aggregation levels the slope of the LLCD plots of packet inter-arrival times decreases, arguing against the infinite variance characteristics of heavy-tailed distributions like e.g. Pareto. Usually, higher aggregation levels are used for the CLT test however these could not be achieved given the length of the instrumented flows. Yet, the slope decrease becomes apparent even for slightly increasing aggregation levels and absence of infinite variance can be safely assumed. Indeed, the least-squares regression we performed on the LLCD plots yielded values $\alpha > 2$, and their overall shape was better described by (light-tailed) log-normal distributions. This finding is in accordance to other studies suggesting that log-normal distributions give better fit to many duration and inter-arrival distributions observed over the Internet, than heavy-tailed Pareto distributions do [8]. It has also been shown that this observed light-tailness is not contradictory to LRD of aggregate traffic [11].

5. Conclusion

We have examined the temporal behaviour of traffic performance characteristics at the level of individual, sufficiently long-lived flows, routed over diverse wireless networks. We have provided evidence of similar LRD intensity between the intraflow traffic arrival process and the unidirectional packet delay, demonstrating that LRD behaviour of aggregate traffic penetrates other measurable quantities at finer granularities, albeit in lesser intensities than the ~0.8 levels reported for traffic aggregates [13][6]. However, even for relatively small LRD intensity (Hurst) values, the ACFs indicate non-obviously-degenerate correlations at large lags. Through the comparative examination of refined Hurst estimates of intraflow traffic properties we identified the possibility of other network and traffic idiosyncrasies, such as transport control mechanisms, wireless access technologies and packet sizes, influencing the intensity of LRD behaviour. At the same time, we have shown the absence of infinite variance phenomena at the distributions of intraflow packet inter-arrival times. Although this study focused on IPv6 flows, similar behaviour is expected for IPv4 traffic, since both protocols assume the same packetisation mechanisms at their higher layers. Whether in practice intermediate routers treat IPv4 and IPv6 traffic identically and how this influences their performance deserves further experimental investigation. In addition, comparative performance analysis between flows routed over wireless technologies and their wired counterparts is to be further pursued.

References

[1] P. Abry and D. Veitch. Wavelet Analysis of Long-Range Dependence Traffic. In IEEE Transactions on Information Theory, 1998

[2] J. Beran, Statistics for Long-Memory Processes. Monographs on Statistics and Applied Probability, Chapman and Hall, New York, NY, 1994

[3] M. S. Borella and G. B. Brewster, Measurement and Analysis of Long-Range Dependent Behaviour of Internet Packet Delay, Proceedings, IEEE Infocom '98, pp. 497-504, Apr. 1998

[4] N. Brownlee, K. C. Claffy, Understanding Internet Traffic Streams: Dragonflies and Tortoises, IEEE Communications Magazine, Volume 40, Issue 10, pp. 110- 117, October 2002

[5] R. Chakravorty, J. Cartwright, I. Pratt, Practical experience with TCP over GPRS, IEEE GLOBECOM’02, Taiwan, 2002

[6] M. E. Crovella and A. Bestavros, Self-similarity in World Wide Web traffic: Evidence and possible causes, IEEE/ACM Trans. on Networking, 5(6): 835 - 846, December 1997

[7] D. Delignieres, S. Ramdani, L. Lemoine, K. Torre, M. Fortes, G. Ninot, Fractal analyses for ‘short’ time series : a re-assessment of classical methods, Journal of Mathematical Psychology, Vol. 50, 2006

[8] A. B. Downey, Lognormal and Pareto distributions in the Internet. Computer Communications 28(7): 790-801, 2005

[9] E. Erramilli, O. Narayan, W. Willinger, Experimental queuing analysis with long-range dependent packet traffic, IEEE/ACM Trans. on Networking, 4(2):209-223, 1996

[10] G. Fotiadis, V. Siris, Improving TCP throughput in 802.11 WLANs with high delay variability”. 2nd IEEE Int’l Symposium on Wireless Communication Systems (ISWCS’05), Italy, September, 2005

[11] J. Hannig, J.S. Marron, G. Samorodnitsky, F.D. Smith, Log-normal durations can give long range dependence, in Mathematical Statistics and Applications: Festschrift for Constance van Eeden, IMS Lecture Notes, Monograph Series, Institute of Mathematical Statistics, 2001, pp. 333-344

[12] T. Karagiannis, M. Molle, M. Faloutsos, Understanding the limitations of estimation methods for long-range dependence, Technical Report, University of California, Riverside, TR UCR-CS-2006-10245, 2006

[13] W. E. Leland, M. Taqqu, W. Willinger, D. V. Wilson, On the Self-Similar Nature of Ethernet Traffic (extended version), IEEE/ACM Trans. on Networking, 2(1): 1 – 15, 1994

[14] Q. Li and D. L. Mills. On the long-range dependence of packet round-trip delays in internet. In Proceedings of IEEE ICC'98, Atlanta, 1998

[15] K. Park, G. Kim, M. Crovella, On the Relationship Between File Sizes, Transport Protocols, and Self-Similar Network Traffic, IEEE International Conference on Network Protocols (ICNP’96), Ohio, USA, Oct. 29- Nov. 1, 1996

[16] K. Park and W. Willinger (eds.), Self-Similar Network Traffic and Performance Evaluation. John Wiley & Sons, New York, NY, USA, 2000

[17] V. Paxson and S. Floyd, Wide-area traffic: The failure of Poisson modelling, IEEE/ACM Trans. on Networking, 3(3):226-244, 1994.

[18] D. P. Pezaros, D. Hutchison, F. J. Garcia, R. D. Gardner, J. S. Sventek, In-line Service Measurements: An IPv6-based Framework for Traffic Evaluation and Network Operations, IEEE/IFIP Network Operations and Management Symposium (NOMS’04), Seoul, Korea, April 19-23, 2004

[19] D. P. Pezaros, M. Sifalakis, D. Hutchison, End-To-End Microflow Performance Measurement of IPv6 Traffic Over Diverse Wireless Topologies, Wireless Internet Conference (WICON'06), Boston, MA, August 2-5, 2006

[20] D. P. Pezaros, M. Sifalakis, L. Mathy, Fractal Analysis of Intraflow Unidirectional Delay over W-LAN and W-WAN Environments, in Proceedings of the third International Workshop on Wireless Network Measurement (WiNMee'07) and on Wireless Traffic Measurements and Modelling (WiTMeMo'07), Limassol, Cyprus, April 20, 2007

[21] M. S. Taqqu, V. Teverovsky, and W. Willinger, Estimators for long-range dependence: An empirical study. Fractals 3, 4 (1995), 785-798

[22] W. Willinger, V. Paxson, M. S. Taqqu, Self-similarity and Heavy Tails: Structural Modelling of Network Traffic, in A Practical Guide to Heavy Tails: Statistical Techniques and Applications, Adler, R., Feldman, R., and Taqqu, M.S. (eds.), Birkhauser, Boston, 1998

[23] W. Willinger, M. S. Taqqu, R Sherman, D. V. Wilson, Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level, IEEE/ACM Transactions on Networking, 5(1):71-86, January, 1997