Bypassing Modelling: an Investigation of Entropy as a Traffic Descriptor in the Fairisle ATM network

Simon Crosby[1], Ian Leslie[1], John T. Lewis[2],
Neil O'Connell[2,3], Raymond Russell[2], and Fergal Toomey[2,4].

February 1995

Abstract

The thermodynamic entropy of an ATM traffic stream is a useful tool for predicting cell-loss ratios and cell-delay. In [1] it was proposed that the entropy of a cell stream be estimated directly at a multiplexer; from this data, predictions can be made rapidly using algorithms which are simple enough to be executed on-the-fly. Preliminary investigations using simulated data have suggested that the method is practicable. In this paper we describe the results of an investigation of the method using real ATM traffic on the Fairisle ATM network [2]. Fairisle is an experimental ATM LAN whose hardware design permits a high degree of experimental flexibility, including high resolution clocks for timestamping and measurement of traffic, a general purpose CPU and memory on every switch port.

Fairisle is equipped with a wide range of ATM traffic sources, including compressed video, audio and LAN data. In addition traffic can be generated from a pre-recorded trace of traffic activity. One such source is the Star Wars movie, reconstructed from a trace of the output of a DCT-based VBR video codec at Bellcore [3]. We describe experiments involving the estimation of the entropy of both single-source and multiplexed cell streams based on the measured cell inter-arrival times. These estimates are used to calculate the decay-rate of the queue-length distribution. For multiplexed cell-streams, the decay-rate can be measured rapidly.

1 Introduction

It is central to the ATM concept that traffic streams with widely different characteristics and service requirements can be carried on the same network if use is made of statistical multiplexing to maximise utilisation of the available resources. But if statistical multiplexing is to work in practice then traffic will have to be characterised in some way that will allow rational decisions to be made with regard to resource allocation. What properties are required of a good traffic characterisation scheme? Firstly, traffic descriptors should contain just that information needed to describe the likely effect of the traffic on the network and no more. Secondly, the scheme should work generally, both with forseeable traffic types such as voice, video and data, and also with unforeseeable future applications. Thirdly, it is desirable that the scheme should not be overly complex or make use of descriptors that are difficult to determine in practice.

It is apparent that parametric schemes such as those based on traffic modelling do not satisfy any of these criteria. Other schemes, based on specifying parameters such as peak cell rate and cell delay variation tolerance, have more general applicability than those based on a particular model but do not provide sufficient information for the full statistical multiplexing gain to be acheived. An alternative scheme, based on measuring the large deviation rate-function of a traffic stream, has the potential to meet all three of the above requirements. This scheme, proposed by some of the present authors at the Eleventh UK Teletraffic Symposium [4], is founded on the following observation, made now by several authors [5]: under very general conditions, a single-server queue has a queue-length distribution with asymptotics of the form

$$\frac{1}{q} \ln \mathrm{P}\{Q > q\} \;\sim\; -\delta \;, \tag{1}$$

[1]University of Cambridge Computer Laboratory.
[2]Dublin Institute for Advanced Studies.
[3]Department of Statistics, Trinity College, Dublin.
[3]Author presenting paper.

and the asymptotic decay-rate δ can be calculated from the rate-function $I(\cdot)$ of the arrival process:

$$\delta \;=\; \min_{c>0} I(c+s)/c\;, \tag{2}$$

where s is the service rate. It follows that, if $I(\cdot)$ can be measured, then δ can be calculated for any value of s. From δ, such important quantities as cell-loss and cell-delay can be estimated.

We call $I(\cdot)$ the 'entropy' of the arriving traffic, to draw attention to an analogy with thermodynamics. The entropy function of a gas, which mathematically speaking is nothing other than a large deviation rate-function, contains just that information needed to calculate the bulk properties of the gas. It is not the practice of engineers to construct complex models of a gas in order to calculate its entropy — they measure it. A similar approach may prove fruitful in the study of statistical multiplexing.

A traffic characterisation scheme based on measuring the rate-function $I(\cdot)$ possesses the first and second of the desirable properties listed above: it provides just that information needed to describe the effect of the traffic on the network, and it can be applied with great generality to traffic of any type. An additional advantage is that, if the rate-functions of several individual traffic streams are known, then the rate-functions of the traffic which results when the streams are multiplexed together in various ways can be calculated using appropriate mathematical prescriptions. Hence, by measuring rate functions at the inputs to a network it may be possible to predict the network's performance at any of its nodes (this question is the subject of on-going theoretical research; see [6] for an introduction).

The purpose of the present paper, then, is to investigate whether the rate function scheme meets the third of our stated criteria: can $I(\cdot)$ be measured without great difficulty in the case of real traffic on a real ATM network? To answer this question we have performed several experiments with video traffic on the Fairisle ATM network, an experimental ATM platform built at the University of Cambridge Computer Laboratory to permit research into performance, architectural and management issues in ATM networks [2].

Fairisle is based on a network architecture composed of ATM switches which are interconnected to form a general topology ATM LAN. The Fairisle switch is based on a simple, 4×4 input buffered crossbar. Switches of degree greater than 4 are constructed by using multiple 4×4 crossbars in a 2-stage delta interconnection network.

The hardware design of the Fairisle network components permits a high degree of experimental flexibility owing to the use of field programmable gate arrays which implement a number of features to support experimental work, including high resolution clocks and hardware time stamping of cells. Fairisle can be configured as either an input or output buffered switch. Each port is equipped with 2K cell buffers, 4 MB of RAM and an ARM RISC processor, which implements all policy-specific traffic control functions, using a flexible control plane which exports RPC-style interfaces [7, 8] for signalling and management. This has enabled us to implement a wide range of measurement facilities to accurately monitor network performance. In the experiments reported in this paper, Fairisle was configured as an output-buffered switch with FIFO queueing.

2 Measuring the Rate Function

In this section we will describe briefly one method of estimating the rate-function. The method described here has already been applied, with considerable success, to simulated 'bursty' traffic (see [1]). Rather than estimating the rate function I itself, it is more convenient to estimate a transform of I called the *scaled cumulant generating function* (CGF), defined by

$$\lambda(\theta) = \lim_{n\to\infty} \frac{1}{n} \ln \mathrm{E}\; \exp(\theta A_n), \tag{3}$$

where A_n denotes the number of cells arriving in an interval of length n. The CGF λ is related to I by the Legendre transform:

$$I(x) = \max_\theta \{x\theta - \lambda(\theta)\}, \tag{4}$$

where the maximisation is performed over all values of θ. The decay-rate δ can be calculated directly from λ from the equation

$$\delta = \max\{\theta > 0 : \lambda(\theta) < s\theta\}. \tag{5}$$

Thus δ is the ordinate of the intercept of the graph of λ and a line of slope s through the origin (see figure 1).

How might we estimate λ? It can be shown that for a very large class of arrivals processes it is possible to find a block-length T such that the aggregated arrivals A_T are approximately independent and identically distributed; in this case the CGF is given approximately by

$$\lambda(\theta) \approx \lambda_T(\theta) = \frac{1}{T} \ln \mathrm{E}\ \exp(\theta A_T). \tag{6}$$

The most straightforward approach to estimating λ therefore is to use the CGF of the empirical distribution of the block arrivals:

$$\hat{\lambda}_T(\theta) = \frac{1}{T} \ln \frac{1}{N} \sum_{i=1}^{N} \exp(\theta A_T^{(i)}). \tag{7}$$

The advantages of this estimator are its simplicity and the fact that it makes minimal assumptions about the traffic. To use it in practice, we must make a choice concerning the block length T. This should be sufficiently large so that successive observations of A_T are approximately independent, but not so large that short bursts of activity, which make an important contribution to the queue-length, are effectively smoothed out; we will return again to this question. The graph of a typical CGF estimated using (7) is shown in figure 1.

3 Comparing the Estimator with Empirical Observations

As we have already seen, the estimated CGF can be used to predict the slope of the log queue length distribution obtained when the traffic is fed through a buffer. This suggests a simple method of comparing the estimator with empirical observations. We pass the traffic through a virtual buffer simulated in software and then compare by eye the graph of $\ln \mathrm{P}\{Q > b\}$ with the straight line of slope

$$\hat{\delta} = \max\{\theta > 0 : \hat{\lambda}(\theta) < \theta s\}. \tag{8}$$

From the graph of a typical CGF (figure 1) we see that that as the service rate s increases the intercept of λ and the line of slope s through the origin moves towards larger and larger θ. Hence by varying s we can examine the accuracy of $\hat{\lambda}(\theta)$ over a range of values of θ.

It turns out that a single buffer with constant service rate is a fairly good representation of the actual situation in the Fairisle switch when configured as an output-buffered switch with FIFO queueing. To demonstrate this, we ran an experiment in which bursty traffic simulated by a two-state Markov model was passed through a switch, and then compared the observed end-to-end delay distribution with that predicted by the rate-function. (The rate- function can be calculated explicitly for this model). For a single-server queue where the asymptotic queue-length distribution is given by (2), the delay distribution is of the form,

$$\mathrm{P}\{D > d\} \approx B e^{-\delta s d}, \tag{9}$$

where s is the service rate. Figure 2 is a plot of the measured delay distribution and the bound provided by the rate function. Clearly the slopes are in close agreement. The measured load in the switch during this experiment was about 97 %, and a total of 10^7 cells were transmitted.

In the experiment, a single Fairisle switch was configured to ensure that all queueing occurred at the outputs. The simulated traffic was transmitted from a trace-driven traffic generator at the input of the switch, and directed to a single output port where the measurements were taken. The experiment required a high utilization of the output link, which was easily arranged by slightly reducing the transmission link rates from the switch (using a feature of the experimental equipment), enabling us to attain a load of 97% with the simulated sources.

This result shows that a simple single server queue can be used not only to test the estimator, but also to model the behaviour of a real switch under given traffic conditions.

4 The Starwars Video Traffic

We turn our attention in this section to results obtained in experiments with 'real' ATM traffic. We will discuss two experiments performed using as traffic the 'Starwars' video data set constructed at Bellcore. In the first of these the traffic was transmitted from a single source in the absence of any other network traffic. In the second, ten different sections of the Starwars data were transmitted simultaneously from different sources to a single sink, and a trace of the arriving multiplexed traffic was recorded.

4.1 The Starwars Data

By now well known in the ATM traffic research community, the Starwars data set was produced by Garrett and Vetterli at Bellcore. The data set has been studied in some detail, for example in [3], and comprises the information content in bytes per frame for 171,000 frames (approximately 2 hours) of the film, as transmitted by a Discrete Cosine Transform (DCT) based codec, with a compression algorithm similar to that of JPEG. The byte count per frame is broken down further to the "slice" level – each slice is 16 video lines. With 30 slices per frame and a frame rate of 24 per second (i.e. the original film rate), one slice represents the information transmitted in about 1.4 milliseconds.

Our data was constructed from the Starwars traces by encoding each slice of video as a single AAL5 PDU; each PDU was transmitted in such a fashion that its cells were evenly spaced throughout the slice time. This smoothing effect over the slice time serves to reduce the peak rate of the source from the line rate (100 Mb/s, if all cells in a slice are transmitted back-to-back) to about 24.2 Mb/s for the worst slice, and is typical of ATM video sources (see, for example [8]). The mean rate of the source is about 5.3 Mb/s.

4.2 Experimental Configuration

We report on two sets of experiments using the Starwars data: the first was based on the transmission of a single Starwars video source across an output buffered Fairisle switch, as described previously. The absolute arrival time of each cell in the film was recorded, and the recorded data used offline to test the estimator, as described below. In the second set of experiments, ten 2 minute segments of the film, widely separated from one another in time, were transmitted across the network, and multiplexed onto a single output link from a Fairisle switch, where a trace of the activity (measured as the number of cells within a certain time period, called the "block size"), and the queueing delay were recorded.

The activity information was used as input to the CGF estimator described above. The results are presented below.

4.3 Discussion

Figures 3 and 4 show the measured activity on the network caused by a single Starwars source, on two timescales. The graphs have several notable features which are unpleasant from the point of view of statistical estimation and prediction. Firstly, they suggest that the traffic is *non-stationary*: its statistical characteristics are not independent of time. This is to be expected because the number of cells per slice is dependent on the degree of compression which can be achieved by the codec. Some frames of video have a large number of high-frequency components and do not compress well. These are typically in frames which are very "busy". The non-stationary nature of the source suggests that the traffic descriptors will have to be re-fitted to the data from time to time as the traffic pattern changes, making it imperative that descriptors should be as simple and easy to estimate as possible. Clearly, traffic descriptors which do not involve modelling have an advantage in this respect.

Secondly, the graphs show that the data contains many short bursts of activity which show up as 'spikes'. These spikes make the dominant contribution to the queue at the switch; this can be seen from the activity boxplot (figure 5). (The central horizontal bar is the median activity, the height of the solid box is the inter-quartile spread, and the vertical dashed bar spans the total range of the observations, except for the outliers which appear as horizontal bars outside the upper and lower whiskers. Observations which are separated from the median by a distance greater than 2.5 times the inter-quartile spread are considered to be outliers.) From the plot, we can see that the outliers span almost as great a range again as all the rest of the data. Thus although they are infrequent they must be taken into account when attempting to predict resource utilisation.

Finally from figure 4 it appears that the traffic is strongly correlated over short time scales. To investigate this, we have made a plot of the sample auto-correlation function for a relatively short, and therefore hopefully stationary, subset of the data. One must be careful in interpreting this type of plot in the presence of non-stationarities as these can give rise to the appearance of strong correlations. It is probably safe, however, to infer from the graph the existence of correlations over time-scales of up to several seconds. (The horizontal bars on the graph give a very approximate 95% confidence interval for the hypothesis of no correlation). The significance of correlations is that, unless they are treated appropriately, they lead to bias in many standard statistical estimators.

In summary then the Starwars traffic has many properties which make it difficult to work with. Applying our simple estimator (7) to the single source traffic we have found that good results are obtained only at small values of θ, and only for very large traffic samples. Figure 7 is a plot of the empirical queue length distributions obtained by passing thirty different partially overlapping subsets of the traffic through a buffer implemented in software at a load of 99%. Also plotted are the bounds predicted from the measured rate function. Clearly there is good agreement at this load (the slopes of the bounds are actually slightly conservative). The sample size here was about 700 seconds of data (representing a substantial proportion of the total), and the block length T used to estimate λ was about 30 seconds. At lower loads results are not as good; $\hat{\delta}$ consistently over-estimates the measured value of δ. This indicates that the estimate $\hat{\lambda}(\theta)$ performs poorly for large θ.

Figure 8 is a scatter plot of δ values obtained from the estimated CGF at 90% load against values estimated by fitting a straight line to the log queue length distribution by hand. It shows that, although the estimated values of δ are typically larger than the measured values, they nevertheless exhibit an approximately linear relationship. The respective ranges of the measured values and the estimates are also similar, as shown by the boxplot, figure 9.

Results obtained using the multiplexed traffic were similar to those for the single source experiment. We found however that a smaller sample size, of the order of 10 seconds, could be used in this case without deterioration of the estimates at high loads.

5 Towards an Improved Estimator

Our experience with the simple estimator (7) shows that the method proposed here works in principle, but that an improved estimator of λ will be required for the method to be useful in practice. At low loads the predominant contribution to the queue at a switch comes from the short bursts of high activity which are so evident in the activity plots. As we mentioned in section 3, these bursts will be smoothed out of the data if the block size chosen to estimate λ is too large. On the other hand, it may be neccessary to choose a large block size if the traffic is strongly correlated over time. The combination of strong correlations and large, short bursts in the Starwars traffic thus appears to be responsible for the poor performance of (7) as an estimator of λ. The results with the multiplexed traffic would seem to confirm this: multiplexing has the effect of smoothing the traffic somewhat, so that good estimates were obtainable using a smaller sample size.

The high activity observations, taken by themselves, are not strongly correlated over time. The correlations which neccessitate the use of a large block length are present mainly when the rate of arrivals is low or near its mean. This suggests that the estimator can be improved by allowing the block length to vary inversely with the activity: at high activity a small block length can be used so that bursts are not smoothed out, while at lower activity the block length can be expanded to deal with strong correlations. One way of acheiving this is to let the block length vary linearly with the activity by setting T to be the time taken for C cells, say, to arrive at the switch. If the arrivals rate is high, this time will be short.

In letting T depend on the arrivals rate we are effectively changing variables, from A_n the number of arrivals in a block of fixed length n, to T_C, the time taken for a fixed number C of cells to arrive. Recent work by one of us (R. R.) gives the relation between the rate-functions, and hence the CGFs, of the two processes. The CGF of the T_C process is given by,

$$\mu(\phi) = \lim_{C \to \infty} \frac{1}{C} \ln \mathrm{E}\ \exp(\phi C - \phi s T_C), \tag{10}$$

where s is the service rate, and δ may be obtained from $\mu(\cdot)$ as before:

$$\delta = \max\{\phi > 0 : \mu(\phi) > 0\}. \tag{11}$$

Details concerning this type of transformation and how it is applied may be found in [9].

As before, we construct an estimator of μ by fixing C and setting

$$\hat{\mu}_C(\phi) = \frac{1}{C} \ln \frac{1}{N} \sum_{i=1}^{N} \exp(\phi C - \phi s T_C^{(i)}). \tag{12}$$

Using this estimator on the single-source Starwars data, we obtained results significantly better than those obtained with (7). At high loads, (12) gives accurate estimates of the decay rate at sample sizes of just 10 seconds (corresponding to about 10^5 cells). Appropriate values for the constant C were of the order of 1000 cells. At loads below about 90% however, the modified estimator exhibits similar problems to those of (7); the estimates of δ were typically larger than the measured values.

6 Discussion

In this paper we have described three experiments carried out on the Fairisle network: bursty traffic simulated by a two-state Markov model was passed through a switch and the observed delay distribution was compared with that calculated on the assumption that the switch can be represented by a single buffer model; data from a single Starwars video source was transmitted across an output-buffered Fairisle switch, and the rate-function of the source was estimated from a record of the cell inter-arrival times; ten short segments of the film, widely separated in time, were multiplexed onto a single output link, and the rate-function of the multiplexed traffic was estimated from a trace of the activity. From the first experiment we conclude that the single buffer model is a useful representation of the output-buffered Fairisle switch. We conclude from the second and third experiments that collecting the data required to estimate the rate-function is straight forward, and can be done on the fly while the switch is operating. Using this data, we compared the decay rate δ estimated from the CGF with the empirical decay rate measured using a virtual buffer. In both experiments good agreement was obtained at high loads, but at lower loads the estimated decay rate was typically larger than the measured value. For single-source traffic, a large sample size of the order of 10 minutes worth of data was needed when the simple estimator (7) was used; for the multiplexed traffic good estimates were obtained with considerably shorter sample sizes, of the order of 10 seconds worth of data. When the improved estimator (12) was employed, the amount of single-source data needed was also reduced to 10 seconds worth.

The improvement in the estimator was based on the idea that block length should be adjusted as the activity changes. By doing this it is possible to avoid smoothing out short bursts of high activity, while retaining the long block length at low activity needed to avoid correlation-induced bias. We are confident that further developments in this direction will lead to an estimator that can be used over the range of loads that will be typical in ATM networks.

The traffic profiles in figures 3 and 4 are consistent with the expectation that traffic from video sources is in general non-stationary, making modelling impracticable. Nevertheless, the experiments described here suggest that there are periods of stationarity which are long compared with the time required to collect sufficient data to characterise the traffic using the rate-function as a descriptor. This suggests that the method described here could be useful in call acceptance control, in contrast with model-based methods. Operators will need to keep track of changing traffic patterns; for this purpose, measurements made on a current call will be useful in making decisions on the acceptance of new calls.

Acknowledgements: We thank Amir Dembo for suggestions which led us to use the simple estimator (7). The rapid computation of δ using the improved estimator (12) was made possible by some convex analysis supplied by Nick Duffield, to whom we are grateful.

References

[1] N. G. Duffield, J. T. Lewis, N. O'Connell, R. Russell, and F. Toomey. Entropy of ATM traffic streams. To appear in Special Issue of IEEE JSAC on Advances in the Fundamentals of Networking, 1995; available as ftp://stp01.stp.dias.ie/DAPG/dapg9430.ps.

[2] R. Black, I. Leslie, and D. McAuley. Experiences of building an ATM switch for the Local Area. In *SIGCOMM*, volume 24(4), September 1994.

[3] M. W. Garrett and M. Willinger. Analysis, modelling, and generation of self-similar vbr video traffic. In *Proceedings, 1994 SIGCOMM Conference*, pages 269–280, London, 1994.

[4] N. G. Duffield, J. T. Lewis, N. O. Connell, R. Russell, and F. Toomey. The entropy of an arrivals process: a tool for estimating QoS parameters of ATM traffic. In *Proceedings of the 11th UK Teletraffic Symposium*, Cambridge, March 1994.

[5] N. G. Duffield and N. O. Connell. Large deviations and overflow probabilities for the general single server queue, with applications. To appear in Proc. Cam. Phil. Soc., available as: ftp://stp01.stp.dias.ie/DAPG/dapg9330.ps.

[6] N. O. Connell. Large deviations in queueing networks. DIAS Technical Report; available as ftp://stp01.stp.dias.ie/DAPG/dapg9413.ps.

[7] R. Black and S. Crosby. Experience and results from the implementation of an ATM socket family. In *Proceedings USENIX Winter Technical Conference*, San Francisco, January 1994.

[8] U. of Cambridge Systems Research Group. ATM document collection. Technical note, University of Cambridge Computer Laboratory, Pembroke Street, Cambridge, CB2 3QG, UK, Mar. 1994. available as: http://www.cl.cam.ac.uk/Research/SRG/bluebook.html.

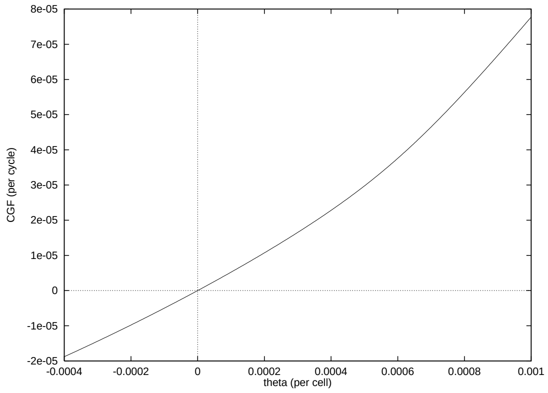

Figure 1: Graph of a typical estimated CGF.

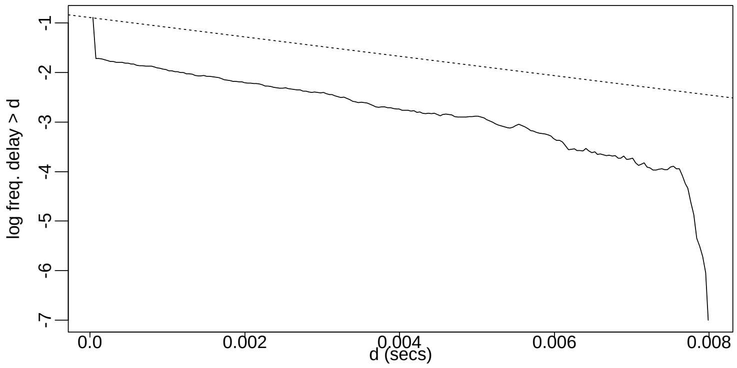

Figure 2: Measured delay distribution (solid curve) and estimated bound (dashed line) obtained while transmitting simulated traffic across the Fairisle switch.

[9] R. Russell. Rate-functions of traffic streams. DIAS Technical Report.

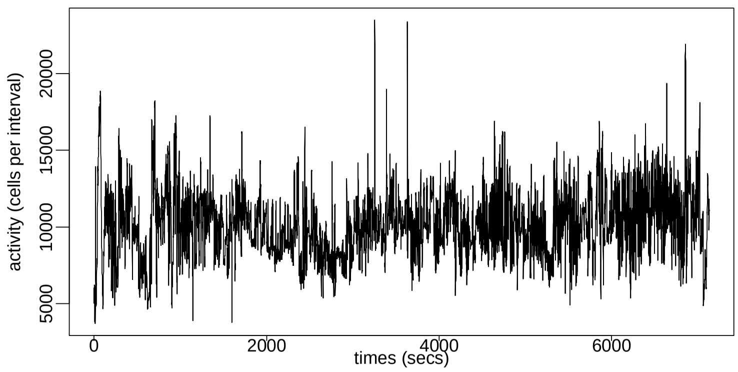

Figure 3: Starwars traffic profile

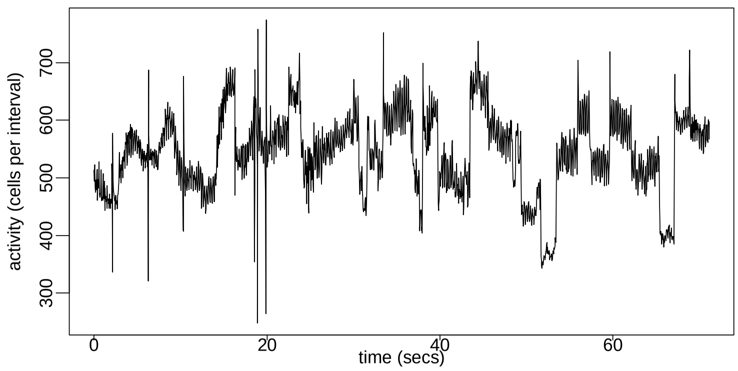

Figure 4: Starwars traffic: close-up

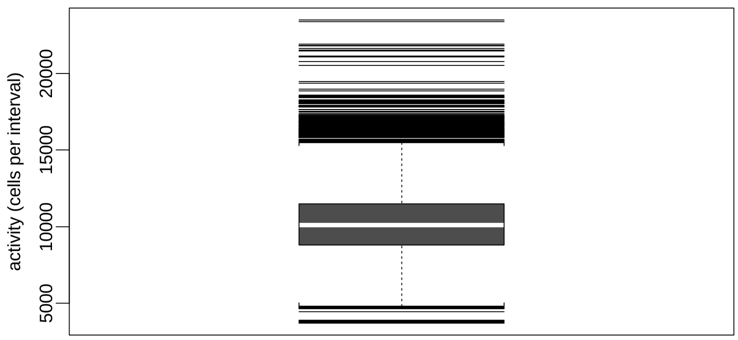

Figure 5: Starwars traffic boxplot

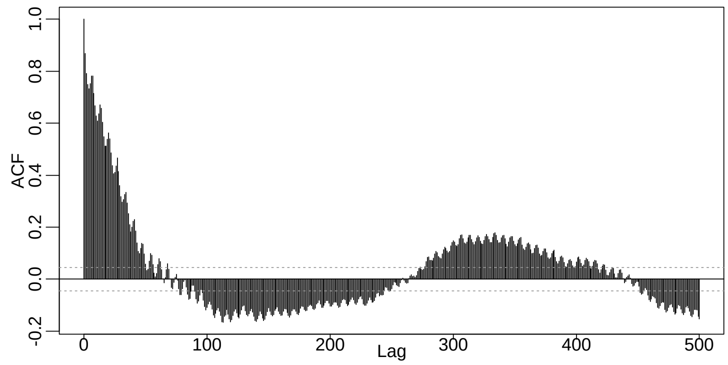

Figure 6: Starwars traffic sample auto-correlation function (1 lag is about 30 milli-secs)

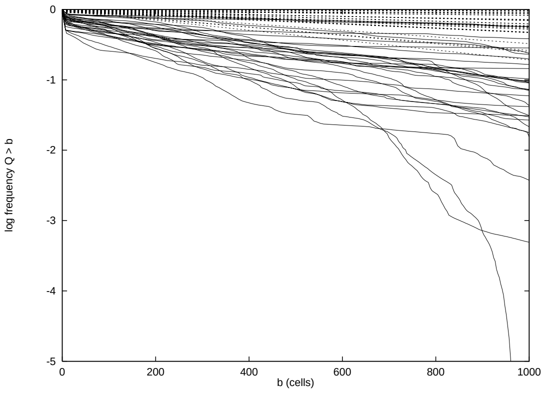

Figure 7: Starwars single source traffic: measured buffer occupancy (solid curves) and estimated bounds (dashed lines) at load 99%.

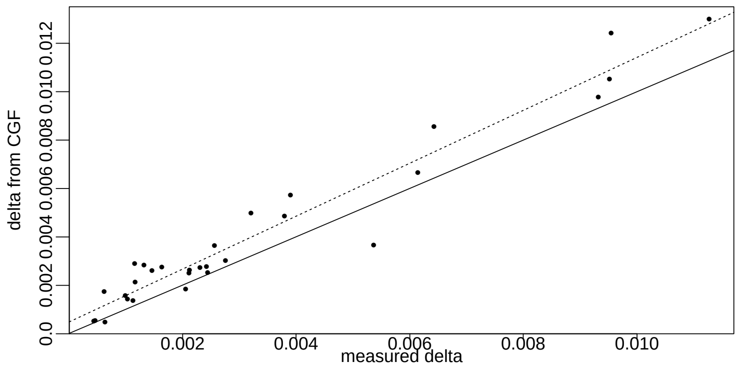

Figure 8: Scatter plot of estimated against measured values of δ at load 90%. The solid line is a line of slope 1 through the origin, and the dashed line is the fitted regression line.

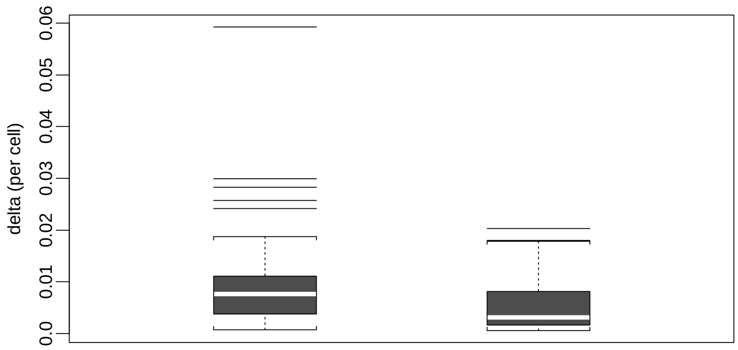

Figure 9: Boxplots of estimated (on the left) and measured (on the right) values of δ at 80% load.