The Entropy of Cell Streams as a Traffic Descriptor in ATM Networks

*N. T. Plotkin**
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025, USA
ninatp@erg.sri.com
and
C. Roche
Laboratoire MASI
Université Paris VI
75252 Paris Cedex 05, France
roche@masi.ibp.fr

Abstract

We examine the properties of a promising new traffic descriptor for ATM networks, namely the entropy of cell streams. The entropy is a measure of the disorganization among cells within a traffic stream; alternatively we can say that entropy captures the amount of randomness in cell scattering. We study the entropy of ON-OFF sources with respect to the typical queue parameters of interest: average queue size, queue variance and equivalent buffers. The equivalent buffer is defined as the minimum buffer size needed to achieve a specific loss probability. We demonstrate that the average queue size and the variance of queue size are monotonically decreasing with increasing entropy in streams with the same fixed load. We find that the inverse of the entropy is closely linear, to within a good approximation, to the equivalent buffer. This simple relation demonstrates the appeal of the entropy estimator. In addition to a measure of cell scattering, our results suggest another interpretation of entropy as a measure of smoothness. Traffic streams with higher entropy (i.e. smoother) have less buffering needs in terms of average, variance and equivalent buffers. Based on this observation we introduce a traffic shaping mechanism whose goal is to boost the entropy of a stream.

Keywords

Traffic Characterization, Entropy, ATM Networks

---

*This work was carried out while this author was a visiting researcher at the Laboratoire MASI at Université Paris VI.

Appeared in: IFIP Performance of Communication Networks. October 1995

1 INTRODUCTION

Control of congestion in ATM networks is expected to be carried out through the use of preventive mechanisms. The overall approach to network control is based on call admission control and real-time source policing. When the user submits a connection request to a network agent, the agent uses a call admission control mechanism to decide whether to accept or reject the call. This decision is made based on the expected consumption of network resources needed by the traffic stream, and the available resources in the network. When a call is accepted, a traffic contract is established between the user and the network. This traffic contract should specify the anticipated characteristics of the traffic flow. During the lifetime of the connection, policing mechanisms at the entry point of the network ensure that the actual characteristics of the traffic stream match those specified in the contract. The goal is to avoid the occurrence of a congested state inside of the network.

Obviously the proper operation of such network control depends heavily on a correct characterization of the traffic flow upon establishment of the traffic contract. One way to characterize a flow is through a well chosen set of traffic descriptors [CCITT, 1992]. A good traffic descriptor should measure important properties of the traffic stream which influence the performance of the network. Traffic descriptors which exhibit simple relations to network performance parameters are desirable. Moreover, for traffic descriptors to be useful for source policing, they must be measurable "on-the-fly", or in real-time [Eckbern, 1992].

Common traffic descriptors include the peak rate, mean rate, utilization factor, and autocorrelation factors. A variety of descriptors based on the concept of a *burst* have also been proposed. The study of superposed Ethernet traffic reported in [Leland, 1993] hints that, although some of these quantities may be useful to characterize simple traffic streams (e.g., single source traffic), they prove inadequate for more complex traffic streams that appear in actual networks, in particular superposed streams. For example, the notion of burst length is ill-defined for highly-superposed LAN traffic. This motivates us to pursue the study of novel traffic descriptors that do not suffer from such limitations.

In this paper we examine the properties of a promising traffic parameter known as the *entropy* of a traffic stream. The idea of using entropy as a traffic metric was first introduced in [Plotkin, 1994], where entropy is used as a tool for studying properties of departure processes in a tandem queueing system. Here we are interested in exploring the relation of entropy to standard ATM QoS parameters, in order to assess its applicability to these networks. We consider traffic streams of discrete-time slotted systems which carry fixed sized cells, such as in ATM networks. Each slot either carries a cell or is empty, therefore a particular sample stream can be viewed as a binary sequence. Intuitively entropy measures the degree of randomness with which cells are dispersed over slots, within a given stream. If there is little randomness in cell placement among slots, then the cells will exhibit repeating patterns. In our binary sequence representation this corresponds to shorter bit patterns that appear repeatedly within a longer binary sequence. If there are few detectable repeating patterns, then the cells are considered more disorganized. The entropy measure grows with increasing disorganization. Since a deterministic stream has an exact organization, its entropy is zero. The terms cell *dispersion*, *disorganization*, and *scattering* can be used interchangeably. Entropy can thus be used as a measure of cell scattering in ATM networks.

For a parameter to be useful as a traffic descriptor it must be both measurable in real-time and meaningful to the performance of the network. In ATM network complex traffic streams arise because as single streams traverse the network they are merged, superposed, interleaved, queued, split apart and so on. Some of these effects will cause cell scattering or dispersion. Others will cause clustering or clumping. Entropy may be useful to capture this naturally occurring phenomenon of cell scattering. The second interesting point is the possibility to estimate the entropy of actual traffic streams. In [Taft-Plotkin, 1994] they illustrated an efficient and implementable method for entropy measurement in real-time systems. The algorithm developed can be used inside a traffic monitor to estimate the entropy of traffic streams on the fly, i.e., while watching cells bypass the monitor. This algorithm uses a variation on Lempel-Ziv data compression techniques.

In this paper we explore the meaning of entropy to discrete-time ATM-network queues. We consider a basic system in which an ON-OFF traffic stream is fed into a queue with a single deterministic server. In particular, we study the relationship between the entropy and each of the three typical queue parameters of interest: average queue size, variance of queue size, and equivalent capacity. We show that both the average queue size and the variance of the queue size are monotonically decreasing with increasing entropy, for fixed loads. We then examine the entropy of the input in relation to the *equivalent buffer*, i.e. the minimum buffer size needed to achieve a specific loss probability. We find that the equivalent buffer is closely linear, to within a good approximation, to the inverse of the entropy.

These results inadvertently suggest that entropy can be used as a measure of traffic “smoothness”. Traffic streams which - for a fixed load - have smaller equivalent buffers, find smaller average queue sizes and see smaller queue variance, can be considered *smoother* than those which have larger equivalent buffers, find larger average queue sizes and see larger queue variance. This interpretation is intuitive since traffic streams which require less buffers can be considered easier to handle, and hence smoother. This understanding of smoothness is based on the same concept of smoothness given in [Low, 1993]. With this interpretation, our results for ON-OFF streams imply that traffic streams with higher entropy are smoother than those with lower entropy. We therefore design a traffic smoother whose goal is to boost the entropy of a stream, and call this smoother an *entropy booster*. Simulation tests demonstrate that our entropy booster can be quite efficient in traffic smoothing and leads to performance improvement.

In section 2 we specify the class of input sources that we consider. Section 3 covers basic entropy definitions and presents the entropy of our sources in a useful form. Section 4 illustrates the relation between the first two moments of our queue system and the entropy of the source. Section 5 introduces our entropy booster and provides preliminary results from simulation testing of this smoothing mechanism. In section 6 we examine the influence of entropy on the equivalent buffer. Our conclusions and ideas for future research directions are discussed in section 7.

2 CHARACTERIZATION OF AN ON-OFF TRAFFIC PROCESS

An ON-OFF traffic process can be used to describe a single variable bit rate source with a bounded peak rate. It is given by a two-state discrete time Markov chain as depicted in Figure 1. In the OFF state the process generates 0 cells/slot, and in the ON state the process generates m cells/slot. The amount of time spent in the ON state, at each visit to the ON state, is given by a geometric random variable B with parameter α. Similarly the lengths of visits to the OFF state are given by a geometric random variable I with parameter β. The probability of transitioning to the OFF (ON) state given that we are in the ON (OFF) state is $1-\alpha$, $(1-\beta$, respectively). Therefore, the fraction of time the source is ON, also called the utilization factor, is given by

$$\sigma = \frac{E[B]}{E[B]+E[I]} = \frac{\frac{1}{1-\alpha}}{\frac{1}{1-\alpha}+\frac{1}{1-\beta}} = \frac{1-\beta}{2-\alpha-\beta} \tag{1}$$

In this model the peak rate is m and the average input rate is σm. The k-th order autocorrelation function for a single on-off input is (the derivation is provided in the Appendix)

$$\gamma(k) = (\alpha+\beta-1)^k \tag{2}$$

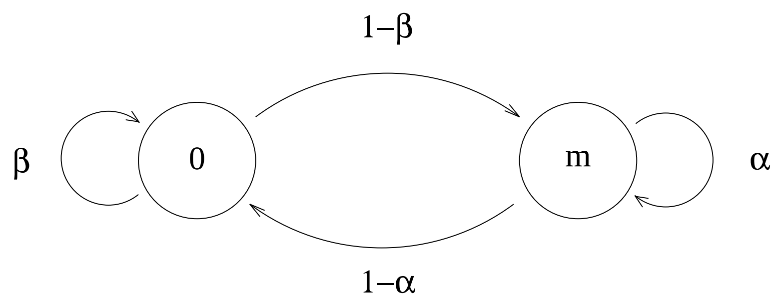

Figure 1 ON-OFF Source Model

A specific sample stream generated by this model can be represented not by a binary sequence of 0's and 1's, but rather by a binary sequence of 0's and m's. This on-off source is entirely described by the triple (α, β, m). This source is equivalently described by the triple (σ, γ, m), where $\gamma = \gamma(1)$, by using the transformation

$$\begin{aligned} 1-\alpha &= (1-\sigma)(1-\gamma) && (3) \\ 1-\beta &= \sigma(1-\gamma) && (4) \end{aligned}$$

The advantages of this transformation are that the parameters σ and γ have a natural physical meaning to the traffic stream generated by the source, and that it leads to simpler forms for the equations we develop below.

3 ENTROPY

The entropy of a random variable X with a probability mass function $p(x)$ defined in an information theoretic sense is given by

$$\mathcal{H}(X) = -\sum p(x) \log_2 p(x) \tag{5}$$

We use logarithms to base 2, and thus measure the entropy in bits. The entropy is a measure of the uncertainty of a random variable. For a Bernoulli process, with parameter $0 \leq p \leq 1$, where $Pr[X_i = 1] = p$, the entropy is

$$\mathcal{H} = -p \log p - (1-p) \log(1-p) \stackrel{def}{=} H(p) \tag{6}$$

and is plotted in Figure 2. We see here that the entropy is symmetric with respect to $p = 1/2$, i.e. $H(p) = H(1-p)$. Consider a binary sequence generated by a Bernoulli process. The symmetry means that the entropy doesn't distinguish between a given sample path and the same path in which the 0's and 1's are interchanged. Entropy is only concerned with the organization of the sample path. This example is simply that of coin tossing and we see that the entropy achieves its maximum for the case of an unbiased coin, i.e. $p = 1/2$. This is the case for which the random variable X_i is the most uncertain. It is known that $0 \leq \mathcal{H} \leq \log |\aleph|$ where $|\aleph|$ denotes the number of elements in the range of X_i. Since in this example $X_i \in \{0, 1\}$ we have $0 \leq \mathcal{H} \leq 1$. (For a complete coverage of entropy definitions and properties see [Cover, 1991].)

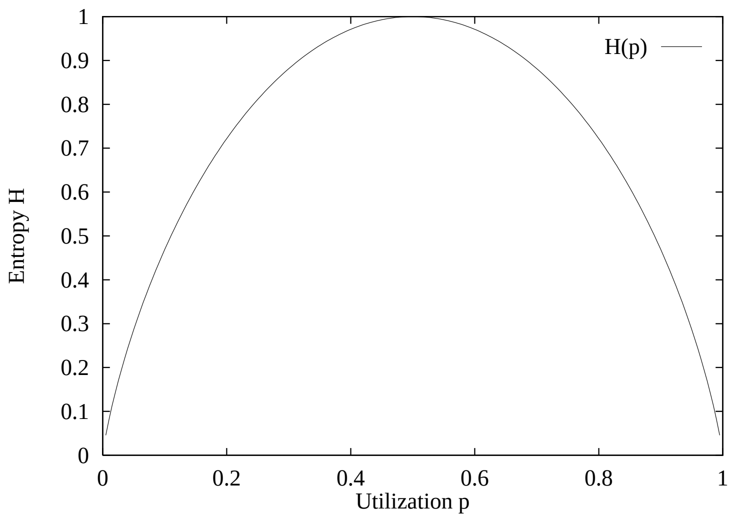

Figure 2 Entropy of a Bernoulli Process with parameter p.

The *entropy rate* (in bits per unit time) of a stochastic process $\{X_i\}$, $i = 0, 1, 2 \ldots$ is defined by $\mathcal{H} = \lim \frac{1}{n} \mathcal{H}(X_1, X_2, \ldots, X_n)$, when the limit exists, where $\mathcal{H}(X_1, X_2, \ldots, X_n)$

is defined as in (5) with $p(x)$ replaced by the joint probability distribution $p(x_1, x_2, \ldots, x_n)$. For a stationary Markov chain $\{X_i\}$ with invariant distribution μ and transition matrix P, the entropy rate is

$$\mathcal{H} = \mathcal{H}(X_2/X_1) = -\sum_{ij} \mu_i P_{ij} \log P_{ij} \tag{7}$$

According to this last equation, the entropy of our bursty on-off traffic source in terms of α and β is given by

$$\mathcal{H} = \frac{1-\beta}{2-\alpha-\beta}\, H(\alpha) + \frac{1-\alpha}{2-\alpha-\beta}\, H(\beta) \tag{8}$$

Using the transformation in equations (3) and (4), and the fact that $H(p) = H(1-p)$, the entropy can be rewritten in a more useful form in terms of σ and γ as follows

$$\mathcal{H} = \sigma\, H\left((1-\sigma)(1-\gamma)\right) + (1-\sigma)\, H\left(\sigma(1-\gamma)\right) \tag{9}$$

In this form, the equation clearly reveals the symmetry of the entropy of an ON-OFF source in the utilization factor σ. This symmetry indicates that the entropy should not be used alone as a traffic descriptor, but rather that it should be coupled with the utilization. Essentially we are studying the traffic descriptor (σ, H) since including the utilization factor allows us to distinguish between two streams with the same entropy but different queueing behavior. In the case of $\gamma = 0$, the right-hand term of equation (9) collapses to $H(\sigma)$ as it should since in this case the ON-OFF process reduces to a simple Bernoulli process with parameter σ. For our sources $X_i \in \{0, m\}$ hence $|\aleph| = 2$ and the range of the entropy values is again given by $0 \leq \mathcal{H} \leq 1$. Note that the value m does not appear in the formula for the entropy, since the entropy is only concerned with the organization of unoccupied slots and occupied slots; (this is true as long as the occupied slots always contain a fixed number of cells).

The interpretation of these definitions in terms of cell dispersion is as follows. Let the Markov Chain $\{X_i\}$ $i = 0, 1, 2\ldots$ be given by an ON-OFF source. In our binary sequence representation an example of a sample path generated by the source might is given by $X_0 = 0, X_1 = m, X_2 = m, X_3 = 0, X_4 = 0, X_5 = 0, X_6 = m$ or simply $0mm000m$. The sample stream $0m00mmm0mm00$ is considered more disorganized than the stream $mm0mm0mm00m0$ since in the latter stream the pattern $mm0$ occurs repeatedly.

4 ON-OFF/D/1 QUEUE

We begin by examining the basic case of a infinite FIFO queue with a deterministic server and an ON-OFF input source. ON-OFF streams are suitable for this study since the ON-OFF traffic model is a common model for ATM networks and because it allows us to vary the amount of cell scattering by tuning its characteristic parameters. These processes also include Bernoulli processes which are a degenerate case of ON-OFF processes when $\alpha + \beta = 1$, or $\gamma = 0$. In [Bruneel, 1994] they determine the z-transform of the steady state buffer occupancy process for an ON-OFF/D/1 queue, whose input is described by the triple (α, β, m). This generating function is given by

$$U(z) = \sum_{k=0}^{\infty} p(k) z^k = \frac{(1-\lambda)(z-1)[\beta + (1-\alpha-\beta)z^{m-1} + (1-\beta)z^m]}{z - \beta - (1-\alpha-\beta)z^{m-1} - \alpha z^m} \tag{10}$$

where $p(k)$ denotes the probability that there are k cells in the queue and where the average input rate λ is given by

$$\lambda = \frac{1-\beta}{2-\alpha-\beta} m \tag{11}$$

We now use this function to determine the mean and variance of the corresponding queue occupancy process. The average number in the queue is obtained from

$$E(N) = \frac{dU(z)}{dz}|_{z=1} \tag{12}$$

After the transformation into the parameter set (σ, γ, m) we obtain

$$E(N) = \frac{m\sigma[2(m-1) - (1-\gamma)(2m\sigma + m - 3)]}{2(1-\gamma)(1-m\sigma)} \tag{13}$$

We now have, with equations (9) and (13), formulas for both the average queue size and the entropy of the input in terms of the same triple (σ, γ, m). By fixing the load (i.e. m and σ) and varying the autocorrelation factor γ we can trace a parametric plot of the average queue length in terms of the entropy. This plot is given in Figure 3. We see that the average queue length increases when the entropy decreases to zero, which corresponds to $\gamma \to 1$. The backtracking behavior that we observe at the bottom of the curves corresponds to $\gamma \to -1$ and the change of direction occurs at the point $\gamma = 0$. For bursty traffic models, the interesting values of γ are those that correspond to positive correlations, i.e. $\gamma > 0$. In this region, $E(N)$ is a monotonically decreasing function of H.

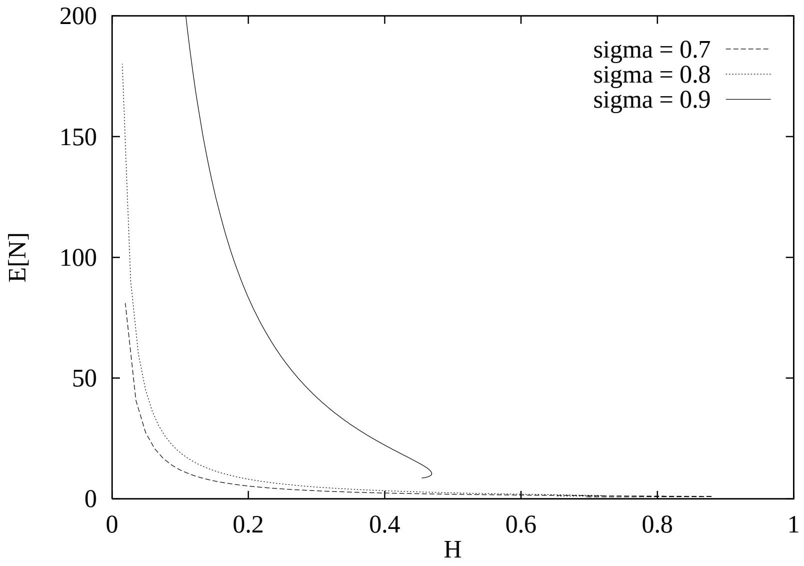

Figure 3 Average Queue Size vs. Entropy

The second moment of the queue size is determined by $E(N^2) = U''(1) + U'(1)$, and is given below.

$$\begin{aligned} E(N^2) &= \frac{\sigma m[6(\sigma m)^2 + (m^2 + 3m - 16)\sigma m + 2m^2 - 9m + 13]\gamma^2}{6(1-\gamma)^2(1-\sigma m)^2} \\ &+ \frac{\sigma m[-12(\sigma m)^2 + (-2m^2 + 6m + 20)\sigma m + 8m^2 - 18m - 2]\gamma}{6(1-\gamma)^2(1-\sigma m)^2} \\ &+ \frac{\sigma m[6(\sigma m)^2 + (m^2 - 9m - 4)\sigma m + 2m^2 + 3m + 1]}{6(1-\gamma)^2(1-\sigma m)^2} \end{aligned} \tag{14}$$

In Figure 4 we trace the variance $V(N) = E(N^2) - E(N)^2$ in terms of the entropy via a parametric plot. We observe similar behavior here as in the previous plot; namely that $V(N)$ is a monotonically decreasing function of H in the region of interest $\gamma > 0$.

We have found that for a fixed load, ON-OFF input streams with higher entropy find lower average queue sizes and see smaller queue size variation. Note that the ON-OFF model does **not** include the case of deterministic periodic streams since there is no setting for α and β that can generate such streams (other than $X_i = 0 \ \ \forall i$ or $X_i = m \ \ \forall i$). Since $H = 0$ in our model corresponds to a stream of either all 0's or either all m's, the above claims - which hold for ON-OFF traffic - cannot be trivially compared to other results for deterministic periodic streams. We believe that such relationships between entropy and queueing, as depicted in Figures 3 and 4, may be generalized to similar models that exhibit a variable amount of randomness such as Interrupted Poisson Process (IPP).

Appeared in: IFIP Performance of Communication Networks. October 1995

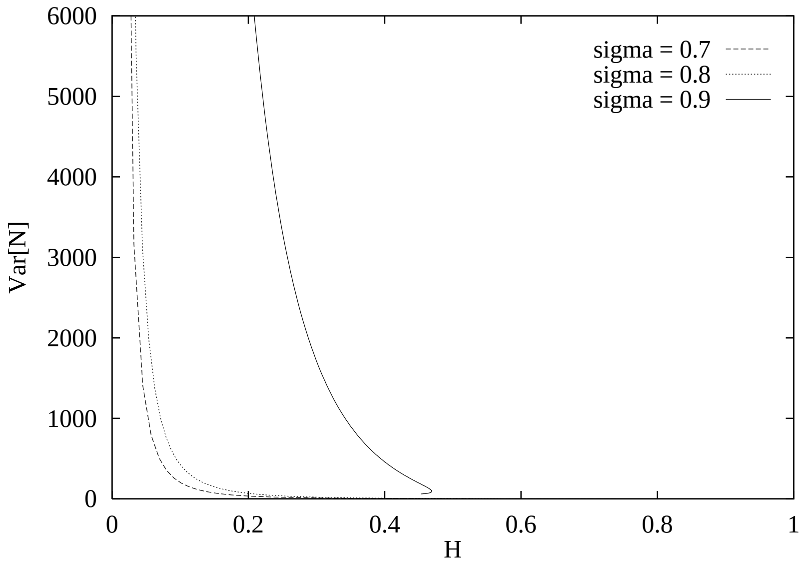

Figure 4 Variance of Queue Size vs. Entropy

As mentioned in the Introduction, we can interpret the *smoothness* of a traffic stream to be related to the amount of buffer space needed to handle a specific stream for a given load. One stream is then considered smoother than another if it requires less buffer space. This suggests a first application of the entropy metric, namely as a measure of smoothness of a traffic stream. Alternatively, we can say that the goal of a smoothing technique should be to increase the entropy of the traffic stream.

5 ENTROPY BOOSTER

Attractive smoothing techniques are those which reduce average queue size and queue variance, because reducing the average queue size leads to a reduction in average delays, and reducing the queue variance leads to a reduction in jitter. Lowering average delays and jitter are desirable from the point of view of network performance. We can see from Figures 3 and 4, that increasing the entropy, for a given stream with a given load, would lead to improved network performance. We introduce the following definition of smoothness.

DEFINITION: Consider two traffic streams X and Y whose entropies are given by $H(X)$ and $H(Y)$ respectively. If $H(Y) > H(X)$, then stream Y is considered *smoother* than stream X.

We therefore design a smoothing technique whose goal is to boost the entropy of a stream, and we call this an *Entropy Booster*. Our entropy booster is depicted in Figure 5. The idea is to achieve $H' > H$ which in turn leads to $E(Q2) < E(Q1)$. One entropy boosting method is is to insert "holes" into the stream at random. (There are a number

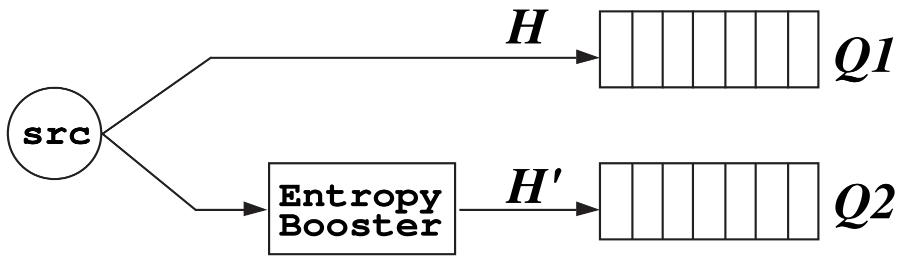

Figure 5 Randomized Smoother

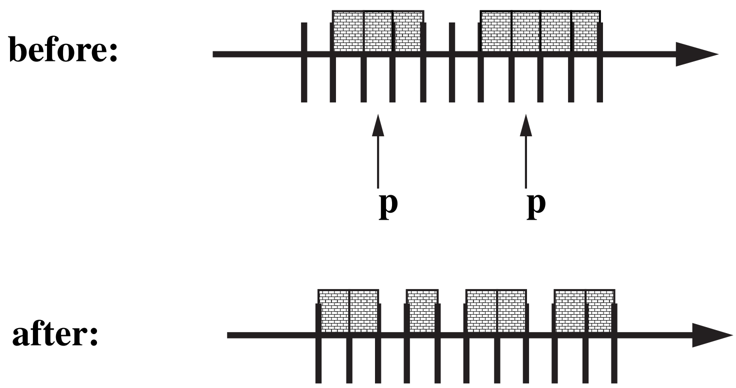

Figure 6 Entropy Boosting Scheme.

of ways to boost the entropy.)

> <u>Entropy Boosting Scheme</u>
> At each slot time, flip a coin with bias p.
> IF the coin yields a head AND there is a cell in the slot,
>   Insert a hole into the stream.
> ELSE do nothing.

By "inserting a hole" we mean that if there is a cell in the slot, then the cell should be held back (i.e., buffered) by one slot time, so that the slot in question is left empty. This method is depicted in Figure 6. This procedure of randomly introducing spaces between consecutive cells breaks up correlations and scatters the cells, thus increasing the stream's entropy. We call the coin bias p the boost parameter. (For simplification in implementation, one need not actually flip the coin at *every* time slot, but rather once every n slots where n is chosen to suit the nature of the traffic being carried.)

To check the validity of the entropy booster scheme, we carried out preliminary simulation experiments. The testing environment emulated the scenario shown in Figure 5. We computed the entropy at the entrance to the queues using the entropy measurement algorithm described in [Taft-Plotkin, 1994].

Figure 7 shows an ON-OFF stream with a load of 0.5 both before and after it has been smoothed. We varied the average burst size (and corresponding idle period size) while

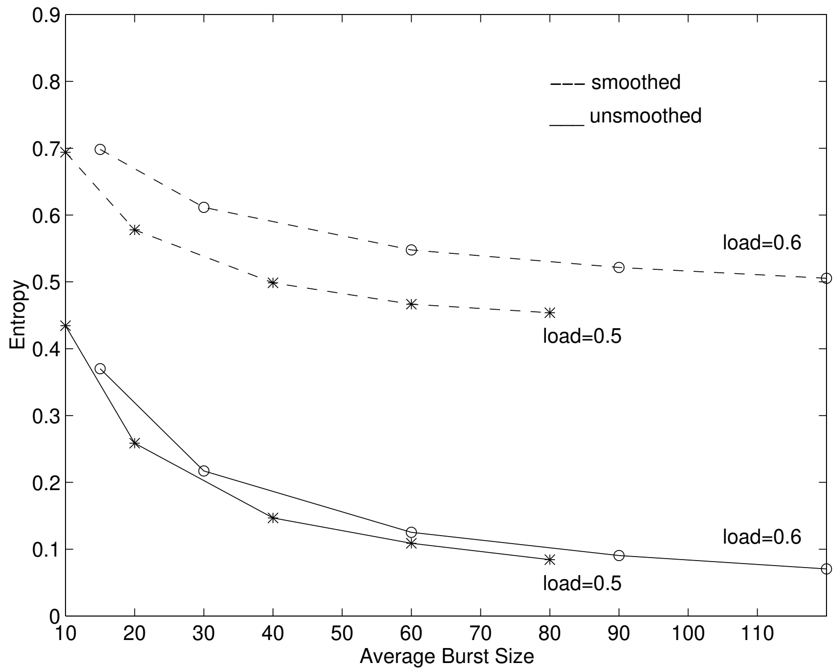

Figure 7 Change in entropy after smoothing.

keeping the load constant. The entropy boost parameter used was 0.2. The example of a load with 0.6 is also given. We see that, for a given burst size, the entropy booster is quite efficient in raising the entropy of a stream. This Figure should be directly compare to Figure 8 in which the exact same situations are considered, only here we plot the corresponding queue size. For both loading cases, the smoothed streams result in much smaller average queue sizes. These tests indicate the correct behavior of the proposed entropy boosting scheme.

It is interesting to note that this approach to traffic smoothing is seemingly in contrast to some of the implicit assumptions in the literature that a smoother stream is one which has been more precisely shaped by a leaky bucket. Future studies which directly compare the entropy-boosting smoother to a token bucket smoother could help to elucidate the relative advantages and disadvantages of each approach. It may be possible to combine the best of both philosophies in the same network.

6 EQUIVALENT BUFFER

Since we saw in Sections 4 and **??** that smooth streams have less buffering needs than non-smooth ones, we would therefore expect that smooth streams also have smaller equivalent buffers than nonsmooth ones. This motivates us to study the precise relationship between the equivalent capacity and the entropy of an ON-OFF traffic stream. In [Guerin, 1991] the authors determine the equivalent bandwidth for a single ON-OFF source using a fluid-flow model in which both the ON and OFF periods are exponentially distributed. Since

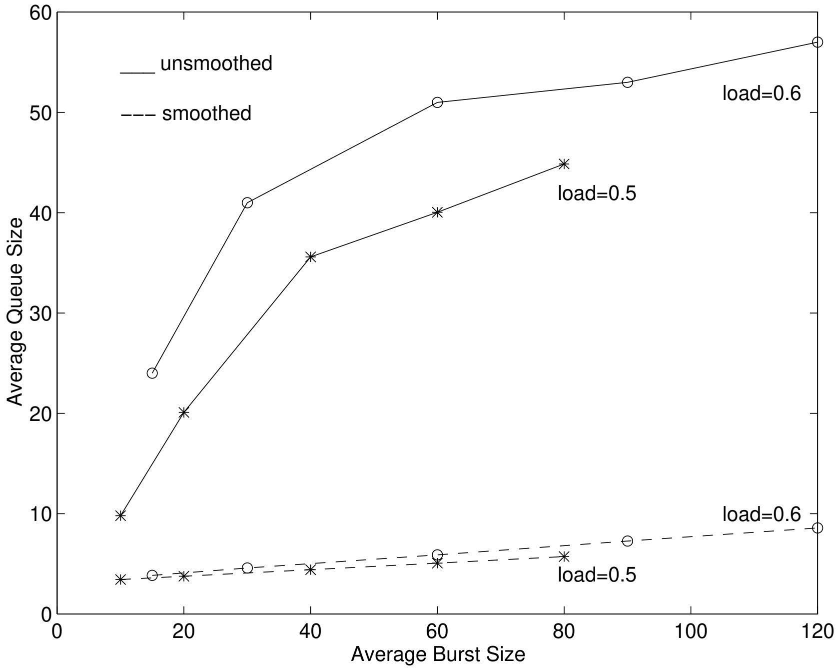

Figure 8 Expected queue size for smoothed and unsmoothed streams.

our traffic model is simply the discrete time version of theirs, we can apply their results. They determine the equation for the overflow probability ϵ to be

$$\epsilon = \nu \exp\left(-\frac{B(c-\sigma m)(1-\gamma)}{c(m-c)}\right) \tag{15}$$

where c is the queue service rate, B is the buffer needed to provide an overflow probability of ϵ, and the constant ν is given by

$$\nu = \frac{(c-\sigma m)+\epsilon\sigma(m-c)}{(1-\sigma)c} \tag{16}$$

We solve equation (15) for the equivalent buffer B which yields

$$B = \frac{ac(m-c)}{(c-\sigma m)(1-\gamma)} = \frac{B_0}{1-\gamma} \tag{17}$$

where the constant a stands for the quantity $\ln(\nu/\epsilon)$. The equivalent buffer B is expressed in number of cells. Let B_0 represent the minimum value of the equivalent buffer in the region of interest. B reaches this value when γ is null, i.e., in the case of a completely uncorrelated stream.

By eliminating the quantity $(1-\gamma)$ between equations (9) and (17), we can rewrite the entropy as a function of the equivalent buffer like this:

$$\begin{aligned}\mathcal{H} &= -\sigma\left\{(1-\sigma)\frac{B_0}{B}\log\left((1-\sigma)\frac{B_0}{B}\right)+\left(1-(1-\sigma)\frac{B_0}{B}\right)\log\left(1-(1-\sigma)\frac{B_0}{B}\right)\right\}\\ &\quad -(1-\sigma)\left\{\sigma\frac{B_0}{B}\log\left(\sigma\frac{B_0}{B}\right)+\left(1-\sigma\frac{B_0}{B}\right)\log\left(1-\sigma\frac{B_0}{B}\right)\right\} \qquad (18)\end{aligned}$$

This formula is not easy to handle as it is. So we use a development to the first order in $1/B$, which yields the following approximation

$$\mathcal{H}\approx 2\sigma(1-\sigma)\frac{B_0}{B}\log\left(\frac{e}{\sqrt{\sigma(1-\sigma)}}\frac{B}{B_0}\right) \qquad (19)$$

where e is the base of natural logarithms. The residual error that we make when approximating $\mathcal{H}$ by the preceding quantity is in $O(1/B^2)$, that is to say extremely small as soon as B is large.

The form of this approximation motivates us to plot $1/\mathcal{H}$ with respect to B, using the exact value given by equation (18). This plot is given in Figure 9 for different values of the utilization factor σ. We used the numerical values $\epsilon = 10^{-9}$ and $c = 1$. Since the inequalities $\sigma m \le c \le m$ must be satisfied for the system to be stable, we chose $m = 1.105$. This leads to $B_0 \approx 9.47$.

We see that these curves are very close to straight lines, which means that $1/\mathcal{H}$ is almost linear with respect to B. In fact we found that this quasi-linearity property holds on any interval $[B_0, B_m]$, with a slope that depends on the upper bound B_m. Due to the properties of the function $B/\log B$, the curves exhibit the same general shape for any value of B_m. This property can be interpreted as follows: once an upper bound (or order of magnitude) B_m is given for B, B is essentially proportional to $1/\mathcal{H}$, i.e. $B \approx \mu/\mathcal{H}$, where the proportionality coefficient μ is given by

$$\mu = 2\sigma(1-\sigma)B_0\log\left(\frac{e}{\sqrt{\sigma(1-\sigma)}}\frac{B_m}{B_0}\right) \qquad (20)$$

The simplicity of this relationship implies that, in the case of Markovian ON-OFF processes, the measure of the entropy of a cell stream can be used directly to estimate the equivalent buffer of the stream. Since the utilization factor σ is used in determining the proportionality coefficient, we see that it is the pair (σ, H) which is a useful traffic descriptor. This result is interesting since it gives us a first idea of the meaning of the entropy of a cell stream from the point of view of network performance. Moreover, if we are able to extend this result to more general processes, this may open interesting perspectives for entropy as a traffic descriptor in ATM networks.

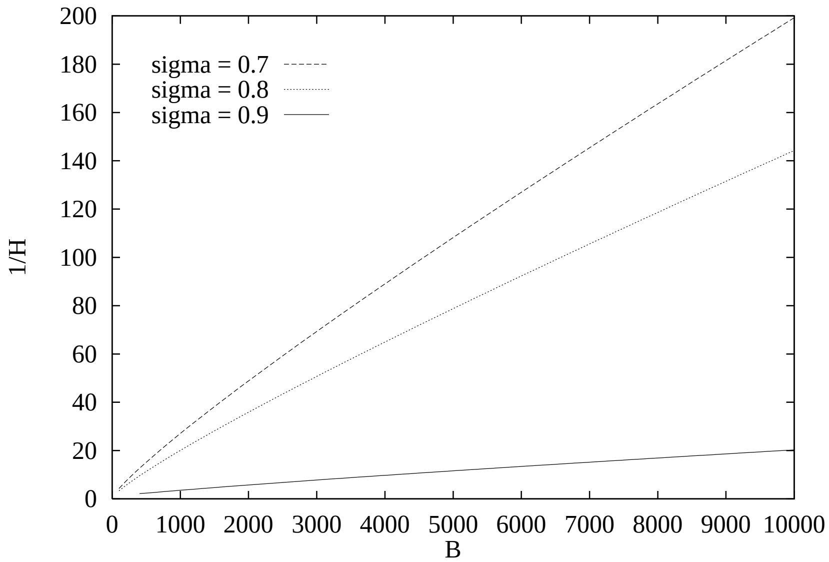

Figure 9 Inverse of Entropy vs. Equivalent Buffer

7 CONCLUSIONS

In this paper we have explored the properties of a promising new traffic descriptor for ATM networks. The entropy is a measure of the disorganization among cells and can be used to estimate the degree of scattering in cell streams. The traffic descriptor (σ, H) appears to be a promising traffic descriptor since the two quantities are both measurable and meaningful. First, we have demonstrated that the relationship of the entropy of an ON-OFF input source to both the average and variance of the corresponding queue can be found. In particular, we showed that the average and variance of the queue are monotonically decreasing with increasing entropy. Second, our study of equivalent buffer revealed a remarkable property that the inverse of the entropy behaves essentially according to the function $B/\log B$. This straightforward relationship between the entropy and the network performance parameter, equivalent buffer, renders the entropy an attractive traffic descriptor.

We introduced a definition of smoothness of a traffic stream which is measured in terms of its entropy. We then presented a sample smoother whose goal is to boost the entropy of the traffic stream. A few simple tests were conducted to demonstrate that this entropy boosting smoother can be quite efficient in reducing delays. Further testing of this traffic smoothing approach under more general traffic conditions is needed. Future research could explore the suitability of the entropy metric as an estimator of the effectiveness of traffic shaping mechanisms. Since the entropy measure is independent of the traffic shaping mechanism itself, it could be used to compare shaping schemes with one another. In particular, it would be interesting to compare the buffered leaky bucket mechanism with the entropy boosting mechanism since these mechanisms are based on opposing philosophies.

Another direction is to examine the usefulness of the entropy metric is measuring dependence across streams which can arise, for example, in multimedia traffic flows. As another example, it has been shown in [Floyd, 1993] that independent periodic processes (e.g. in the TCP protocol, in periodic routing protocols, and others) can become inadvertently synchronized and hence dependent. Such behavior could potentially be measured using entropy estimation techniques. They showed that this dependence can degrade network performance. They also adopted the approach of introducing randomness into the network timing in order to avoid performance degradation.

This work indicates that it would be useful to continue to examine the relationship between entropy and queueing. The extended application of this descriptor to a broader class of sources, such as superposed ON-OFF processes, should be studied. Future work could also examine other traffic models with more complex correlational structures than those reflected in traditional Markov chain models.

APPENDIX

We show here how the k-th order autocorrelation coefficient $\gamma(k)$ given in equation (2) was derived. Let A_n denote the number of cells generated in slot n by an ON-OFF process. Thus $A_n \in \{0, m\}$. The k-th autocorrelation coefficient is defined by

$$\gamma(k) = \frac{E(A_{n+k}A_n) - E(A_{n+k})E(A_n)}{\sqrt{Var(A_{n+k})Var(A_n)}} = \frac{E(A_{n+k}A_n) - E(A_n)^2}{Var(A_n)} \tag{21}$$

The second equality is true due to stationarity. For the term $E(A_{n+k}A_n)$ we have

$$E(A_{n+k}A_n) = \sum_{x\in\{0,m\}} \sum_{y\in\{0,m\}} xy\Pr(A_{n+k} = x, A_n = y) = m^2\Pr(A_{n+k} = m/A_n = m)\sigma \tag{22}$$

where the second equality is true since most terms in the summation are zero. To find $\Pr(A_{n+k} = m/A_n = m)$ let

$$\begin{aligned} s(k) &= \Pr(A_{n+k} = m/A_n = m) \\ q(k) &= \Pr(A_{n+k} = 0/A_n = m) \end{aligned} \tag{23}$$

We can define the following set of recursive equations.

$$\begin{aligned} s(k+1) &= \alpha s(k) + (1-\beta)q(k) \\ q(k+1) &= (1-\alpha)s(k) + \beta q(k) \end{aligned} \tag{24}$$

Solving these equations for $s(k)$, with the initial conditions $s(1) = \alpha$ and $q(1) = 1 - \alpha$, and substituting back into the definition for $\gamma(k)$ yields $\gamma(k) = (\alpha + \beta - 1)^k$.

REFERENCES

Bruneel, H. and Kim, B. (1994) *Communications Systems Including ATM.* Kluwer Academic Publishers.

CCITT Recommendation I.371 (1992) *Traffic control and congestion control in B-ISDN.* Geneva, Switzerland. June.

Cover, T. and Thomas, J. (1991) *Elements of Information Theory.* Wiley.

Eckberg, A. (1992) B-ISDN/ATM Traffic and Congestion Control. *IEEE Network Magazine*, September, 6(5):28–37.

Floyd, S. and Jacobson, V. (1993) The Synchronization of Periodic Routing Messages. *ACM SigComm Proceedings*, September.

Guerin, R. and Ahmadi, H. and Naghshineh, M. (1991) Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks. *IEEE Journal On Selected Areas in Communications*, September.

Leland, W. and Taqqu, M. and Willinger, W.and Wilson, D. (1993) On the Self-Similar Nature of Ethernet Traffic. *ACM SigComm Proceedings*, September.

Low, S. and Varaiya, P. (1993) Burstiness Bounds for Some Burst Reducing Servers. *IEEE InfoCom'93 Proceedings*, March.

Plotkin, N. and Varaiya, P. (1994) The Entropy of Traffic Streams in ATM Virtual Circuits. *IEEE InfoCom'94 Proceedings*, June.

Takagi, H. (1993) *Queueing Analysis Volume 3: Discrete-Time Systems.* Elsevier Science.

Taft-Plotkin, N. (1994) *High-Speed Network Traffic: Characterization and Control.* PhD Thesis, University of California, Berkeley.