Characterization of ATM On-Off Traffic from Cell Traffic Measurements

Thomas M. Chen Dept. of Electrical Engineering SMU PO Box 750338 Dallas, TX 75275 Tel: 214-768-8541 Email: tchen@seas.smu.edu

ABSTRACT— The optimal design of broadband switching systems depends on an understanding of the statistical nature of the broadband traffic. An important practical problem is characterization of the network traffic by fitting stochastic models to actual traffic measurements. This paper studies the problem of characterization of ATM cell traffic by maximum likelihood estimation of the commonly used Markov on-off model. The maximum likelihood estimators (MLEs) and their probability distribution function are derived. The MLEs are shown to be unbiased, consistent, and asymptotically normal. Maximum likelihood estimation is compared numerically with the method of moment matching for a simulated on-off packet trace, which indicate that the MLEs result in smaller variance and better accuracy than moment matching.

I. INTRODUCTION

The optimal design of broadband switching systems depends on an understanding of the statistical nature of the broadband traffic. Typically, stochastic models are assumed for traffic sources, and queueing analysis or simulations are conducted to evaluate network performance. Many stochastic models have been proposed for broadband traffic [1]. A practical issue in traffic modeling and performance evaluation is the validity of any assumed stochastic traffic models. There has been relatively less work on statistical methods to fit stochastic models to actual traffic measurements taken at the sources or points within a broadband network.

The simple Markov on-off traffic model has been widely used for characterization of packet speech and data sources [1-3], and has served as the basis for stochastic fluid approximations [4,5]. In the model, active periods alternate with idle periods according to a continuous-time Markov chain. Active periods are exponentially distributed with mean $1/\alpha$, and idle periods are exponentially distributed with mean $1/\beta$. During an active period, packets are generated at regular periods of *T*.

This paper investigates the method of maximum likelihood estimation to infer the parameters of the Markov on-off model from ATM cell measurements (a trace of cell arrival times). Maximum likelihood estimation is often used because maximum likelihood estimators (MLEs) are well known to be consistent, asymptotically normal, and asymptotically efficient under certain "regularity" conditions [6,7]. In section 2, the Markov on-off model is described as a renewal process, and a simple method is presented for estimating the parameter T. In section 3, the MLEs for α and β are derived, and their mean and variance are found. The MLEs are shown to be unbiased, consistent, and asymptotically normal. Section 4 presents numerical results that compare the performance of the MLEs with the method of moment matching.

II. MARKOV ON-OFF MODEL

A. Model Parameters

In the model, active periods alternate with idle periods according to a continuous-time Markov chain. Active periods are exponentially distributed with mean $1/\alpha$, and idle periods are exponentially distributed with mean $1/\beta$. During an active period, packets are generated at regular periods of *T*. We assume that an active period typically consists of multiple consecutive packets ($\alpha T < 1$) which is the case of practical interest.

For estimation, the on-off source is more conveniently viewed as a renewal process with interarrival time distribution given by

$$F(t) = [1 - \alpha T e^{-\beta(t-T)}] u(t-T)$$
(1)

where u(t) is the unit step function [3].

A traffic trace is assumed to consist of a set of measured packet arrival times $\{a_0, a_1, \dots, a_N\}$, or equivalently, a set of packet interarrival times $\{t_1, \dots, t_N\}$ where $t_n \equiv a_n - a_{n-1}$. The given traffic trace is assumed to be free of jitter; in practice, physical transmission layer factors may cause small random variations in the packet arrival times that would complicate the problem of model fitting. The measured interarrival times $\{t_1, \dots, t_N\}$ are used as the basis for estimating the on-off model parameters: α , β , and *T*.

B. Estimation of T if Unknown

If *T* is not known *a priori*, a natural estimator for *T* is $\hat{T} = \min(t_1, ..., t_N)$. It can be easily shown that \hat{T} is a consistent but biased estimator. Fortunately, the bias rapidly diminishes when *N* becomes large. The probability distribution function for \hat{T} is

$$\Pr(\hat{T} \le t) = 1 - \Pr(t_1 > t) \cdots \Pr(t_N > t)$$
(2)
= $[1 - (\alpha T)^N e^{-N\beta(t-T)}] u(t - T).$

The mean is $E(\hat{T}) = T + \frac{(\alpha T)^N}{N\beta}$, therefore the estimator has a small positive bias $(\alpha T)^N / N\beta$ that diminishes quickly with *N*. The estimator will converge to *T* in probability since $\Pr(|T - \hat{T}| \le \varepsilon) \rightarrow 1$ as $N \rightarrow \infty$ for any $\varepsilon > 0$.

Because the estimator is highly likely to be exactly *T* with large samples, we will assume that *T* has been accurately estimated before proceeding to estimate the other model parameters, α and β . In all the numerical simulations in this study, we found that \hat{T} quickly found the true value of *T* within a few samples.

III. MAXIMUM LIKELIHOOD ESTIMATION

A. Likelihood Function

Maximum likelihood estimators (MLEs) are often used because they are known to be consistent, asymptotically normal, and asymptotically efficient under certain "regularity" conditions [6,7]. Because interarrival times are independent samples of the probability distribution function (1), the likelihood function is the product

$$L_N(\alpha,\beta) = \prod_{n=1}^N f(t_n)$$
(3)

where f(t) is the probability density function for interarrival times:

$$f(t) = (1 - \alpha T)\delta(t - T) + \alpha\beta T e^{-\beta(t - T)}u(t - T)$$
(4)

and $\delta(t) = du(t) / dt$ is the Dirac delta function. The maximum likelihood estimators (MLEs) are the values that maximize the likelihood function (3) or the log-likelihood function

$$l_N(\alpha,\beta) = \sum_{n=1}^N \ln f(t_n).$$
(5)

The likelihood function is unfortunately difficult to maximize due to the singularities (delta functions) in the probability density function (4). For computation, we can approximate the log-likelihood function by

$$l_N(\alpha,\beta) = \sum_{n=1}^N \ln \frac{1}{2\Delta} [F(t_n + \Delta) - F(t_n - \Delta)].$$
(6)

That is, for very small Δ and all $t \ge T$, we replace the probability density function f(t) by the central difference approximation:

$$\frac{1}{2\Delta} [F(t+\Delta) - F(t-\Delta)]$$

$$= \begin{cases} 0 , \text{ if } t < T - \Delta \\ (1 - \alpha T) / 2\Delta , \text{ if } T - \Delta \le t < T + \Delta \\ \alpha \beta T e^{-\beta(t-T)} , \text{ if } t \ge T + \Delta \end{cases}$$
(7)

The particular value of Δ will not be important in determining the MLEs.

The MLEs for α and β can be found by taking partial derivatives of the log-likelihood function (6) with respect to α and β , setting the result equal to 0, and then solving for α and β , respectively. Let *K* be the number of samples that are

equal to *T* (or less than $T + \Delta$ for arbitrarily small Δ), then the MLEs for α and β are given by:

$$\hat{\alpha} = \frac{N-K}{NT}, \quad \hat{\beta} = \frac{N-K}{\sum_{n=1}^{N} t_n - NT}.$$
(8)

The MLEs are consistent with intuition. First, $\hat{\alpha}T$ is estimating the relative fraction of interarrival times that are greater than *T* which has a probability of αT . It will be more convenient at this point to define $b \equiv 1/\beta$ and use the MLE

$$\hat{b} = \frac{\sum_{n=1}^{N} t_n - NT}{N - K} \tag{9}$$

instead of the MLE $\hat{\beta}$ because the probability distribution of \hat{b} will be easier to find. The MLE \hat{b} is estimating the mean length of interarrival times that are greater than *T*. In the special case when K = N, the MLE \hat{b} is undefined but we will select $\hat{b} = 0$.

B. Properties of the MLE

We are interested in the mean and variance of the MLEs as measures of their accuracy for a given sample size. Also, asymptotic properties of the MLEs give an indication of their behavior for large samples.

Proposition 1: The MLEs $\hat{\alpha}$ and \hat{b} have the probability distribution functions:

$$\Pr(\hat{\alpha} \le x) = \sum_{k=0}^{\lfloor xNT \rfloor} \frac{N!}{(N-k)!k!} (\alpha T)^k (1-\alpha T)^{N-k} \quad (10)$$

$$\Pr(\hat{b} \le x) = \sum_{n=1}^{N} \sum_{i=n}^{\infty} \frac{N!}{i! (N-n)! n!} (\alpha T)^{n} (1 - \alpha T)^{N-n} (\beta n x)^{i} e^{-\beta n x} + (1 - \alpha T)^{N}$$

for $x \ge 0$, where $\lfloor x \rfloor$ is the floor function for the largest integer less than or equal to *x*.

Proof: Appendix A. ■

We can immediately find the means of the MLEs as

$$E(\hat{\alpha}) = \alpha, \quad E(\hat{b}) = b$$
 (11)

and the variances of the MLEs as $var(\hat{\alpha}) = \frac{\alpha(1 - \alpha T)}{NT}$ (12)

$$\operatorname{var}(\hat{b}) = \sum_{n=1}^{N} \frac{b^2}{n} \frac{N!}{(N-n)!n!} (\alpha T)^n (1-\alpha T)^{N-n}.$$

The mean and variance are important measures of the accuracy of the estimators for a given sample size of N. The MLE usually displays desirable asymptotic properties such as consistency and normality.

Proposition 2: The MLEs $\hat{\alpha}$ and \hat{b} are unbiased, consistent, and asymptotically normal.

Proof: Appendix B.

IV. NUMERICAL RESULTS

We simulated on-off cell traffic traces to verify the performance of the MLEs and to compare the method of maximum likelihood estimation to another method. Because interarrival times in the on-off model are i.i.d. (independent and identically distributed), another natural choice for parameter estimation is the technique of moment matching.

The method of moments is a straightforward technique that matches the sample mean and sample variance of the measured interarrival times to their statistical mean and variance, respectively:

$$m = \frac{1}{N} \sum_{n=1}^{N} t_n = \frac{T(\alpha + \beta)}{\beta}$$
(13)
$${}^2 = \frac{1}{N-1} \sum_{n=1}^{N} (t_n - m)^2 = \frac{\alpha T}{\beta^2} (2 - \alpha T).$$

The solutions for α and β are the moment-matching estimators:

S

$$\hat{\alpha} = \frac{2}{T} \frac{(m-T)^2}{s^2 + (m-T)^2}$$
(14)
$$\hat{\beta} = \frac{2(m-T)}{s^2 + (m-T)^2}$$

For comparison with the MLEs, we use the momentmatching estimator for $b \equiv 1/\beta$ given by

$$\hat{b} = \frac{s^2 + (m - T)^2}{2(m - T)}.$$
(15)

It is well known that the sample mean and sample variance will converge to the true mean and variance when the samples are independent and identically distributed, and therefore the moment-matching estimators will converge to their true values if the interarrival time distribution is truly given by (1). Unfortunately, the mean and variance of the moment-matching estimators are difficult to obtain, hence we investigate their performance by means of simulations.

We simulated M = 8 independent cell traffic traces with interarrival time distribution given by (1) and the true values T = 1, $\alpha = 0.2$, and b = 1.0. Figure 1 shows that the experimentally measured means of the MLEs $\hat{\alpha}$ and \hat{b} are consistent with their statistical means calculated in (11), as expected. Likewise, Figure 2 shows that the experimentally measured variances of MLEs $\hat{\alpha}$ and \hat{b} are consistent with their statistical variances calculated in (12).





Fig.1. Experimentally measured mean of MLE $\hat{\alpha}$ compared to true mean. (b) Experimentally measured mean of MLE \hat{b} compared to true mean.



Figure 3 shows simulation results for the momentmatching estimators (14). Their measured means are close to

the true values of α and *b*, indicating that these estimators are unbiased (or their bias is small).



of moment-matching \hat{b} compared to true value of b.

Finally, Figure 4 compares the measured variances of the MLEs and the moment-matching estimators as a function of N. The results indicate that the MLEs have significantly less variance than the moment-matching estimators.

V. CONCLUSIONS

The MLEs and their probability distribution function are derived. They have been shown to be unbiased, consistent, and asymptotically normal estimators. In numerical results with a simulated cell trace, the MLEs compare favorably with moment-matching estimators in terms of lower variance and better accuracy.



Fig. 4. Experimentally measured variance of MLE $\hat{\alpha}$ compared to moment-matching estimator. (b) Experimentally measured variance of MLE \hat{b} compared to moment-matching estimator.

REFERENCES

- [1] A. Adas, "Traffic models in broadband networks," *IEEE Commun. Mag.*, vol. 35, pp. 82-89, July 1997.
- [2] J. Daigle, J. Langford, "Models for analysis of packet voice communication systems," *IEEE J. on Selected Areas in Commun.*, vol. SAC-4, pp. 847-855, Sept. 1986.
- [3] H. Heffes, D.Lucantoni, "A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance," *IEEE J. on Sel. Areas in Commun.*, vol. SAC-4, pp. 856-867, Sept. 1986.
- [4] D. Anick, et al., "Stochastic theory of a data-handling system with multiple sources," *Bell System Tech. J.*, vol. 61, pp. 1871-1894, Oct. 1982.
- [5] B. Maglaris, et al., "Performance models of statistical multiplexing in packet video communications," *IEEE Trans. on Commun.*, vol. 36, pp. 834-844, July 1988.
- [6] I. Basawa and B. Prakasa Rao, Statistical Inference for Stochastic Processes. NY: Academic Press, 1980.
- [7] S. Wilks, *Mathematical Statistics*. NY: Wiley & Sons, 1975.

APPENDIX A

If *K* is defined as the number of samples equal to *T*, then *N* - *K* is the number of samples greater than *T*. According to (1), *N* - *K* is binomial with the probability mass function $Pr(K - N = x) = \frac{N!}{(N - x)! x!} (\alpha T) x (1 - \alpha T)^{N - x}$ for

x = 0, ..., N. The probability distribution of $\hat{\alpha}$ in (10) follows from the observation that

$$\Pr(\hat{\alpha} \le x) = \Pr\left(\frac{N-K}{NT} \le x\right) = \Pr(N-K \le xNT).$$
(A.1)

Given that N - K = n, the MLE \hat{b} can be written as $\hat{b} = \tau / n$ where τ is the sum of n exponential random variables with mean $1/\beta$. That is, τ is a gamma (n,β) random variable with probability density function $p(t) = \frac{\beta^n t^{n-1} e^{-\beta t}}{(n-1)!}$ for $x > 0, n \ge 1$. The conditional probability distribution function of \hat{b} is

$$\Pr(\hat{b} \le x | N - K = n) = \int_{0}^{nx} \frac{\beta^{n} t^{n-1} e^{-\beta t}}{(n-1)!} dt$$
(A.2)
$$= \sum_{i=n}^{\infty} \frac{(\beta n x)^{i}}{i!} e^{-\beta n x}.$$

The probability distribution function of \hat{b} given in (10) follows from the earlier observation that N - K is binomial (and recall that $\hat{b} = 0$ for the case N - K = 0).

APPENDIX B

The mean and variance of the MLE $\hat{\alpha}$ can be found from (8) as

$$E(\hat{\alpha}) = \frac{E(N-K)}{NT} = \frac{\alpha NT}{NT} = \alpha$$
(B.1)
$$var(\hat{\alpha}) = \frac{var(N-K)}{(NT)^2} = \frac{N\alpha T(1-\alpha T)}{(NT)^2} = \frac{\alpha(1-\alpha T)}{NT}$$

This shows that $\hat{\alpha}$ is unbiased and approaches α in probability as $N \to \infty$. By the well-known DeMoivre-Laplace limit theorem, the binomially distributed N - K will approach a normal distribution with mean $N\alpha T$ and variance $N\alpha T(1-\alpha T)$ in the limit as $N \to \infty$. Hence $\hat{\alpha}$ is asymptotically normal with mean α and variance $\frac{\alpha(1-\alpha T)}{NT}$.

As noted earlier, given that N - K = n, the MLE \hat{b} can be written as $\hat{b} = \tau / n$ where τ is the sum of n exponential random variables with mean $1/\beta$. The conditional mean of \hat{b} can be readily shown to be

$$E(\hat{b}|N-K=n) = \frac{E(\tau)}{n} = \frac{nb}{n} = b$$
 (B.2)

which is independent of *n*, hence \hat{b} is unbiased. The conditional variance of \hat{b} is

$$\operatorname{var}(\hat{b}|N-K=n) = \frac{\operatorname{var}(\tau)}{n^2} = \frac{nb^2}{n^2} = \frac{b^2}{n}$$
 (B.3)

hence the unconditional variance of \hat{b} is given by (12).

The MLE \hat{b} is a consistent estimator if the variance can be demonstrated to reduce to zero when $N \to \infty$. The binomially distributed N - K will approach a normal distribution with mean $N\alpha T$ and variance $N\alpha T(1 - \alpha T)$ in the limit as $N \to \infty$. Then the variance approaches

$$\operatorname{var}(\hat{b}) = \int_{0}^{\infty} \frac{b^{2}}{x} \frac{1}{\sqrt{2\pi N\alpha T (1 - \alpha T)}}$$
(B.4)
$$\cdot \exp\left\{-\frac{(x - N\alpha T)^{2}}{2N\alpha T (1 - \alpha T)}\right\} dx$$

The integral on the right hand side will diminish to zero as $N \rightarrow \infty$, therefore \hat{b} will converge to its mean b.

Finally, given N - K = n, the MLE \hat{b} can be written as $\hat{b} = \tau / n$ where τ is the sum of *n* exponential random variables with mean $1/\beta$. When *N* increases, N - Kwill approach $N\alpha T$ and \hat{b} will be the sample mean of an increasing number of exponential(β) random variables. By the central limit theorem, \hat{b} will be asymptotically normal with mean *b* and variance $\frac{b^2}{N\alpha T}$.