Application of simulation models to measured traffic traces

Levente Bodrog
Advisor: Miklós Telek associate professor
Department of Telecommunications, Technical University of Budapest

Abstract

There are several simulation models, which are out to fit measured datasets, i.e. a model capturing all the properties examined by the known statistical analysis methods.

The simulation models presented in [12] and [5] are effective in fitting the long range dependent behaviour of measured traces, but they do not approximate the marginal distribution. The main goal of this paper is to modify those methods such that they capture not only the long range dependent behaviour, but also the marginal distribution.

We develop fitting methods based on a selected set of statistical properties including long range dependency, Haar wavelet parameters and the marginal distribution. These properties are checked by the R/S plot, the Variance time plot, the Legendre spectrum, the Haar wavelet transform, the queue length distribution and the marginal distribution analysis.

The paper reviews the statistical analysis of a measured data set, the steps of approximating it with the modified simulation models and the analysis of the synthetized data sets generated by the modified and the original simulation models. During the analysis we evaluate the effect of the fitting method on the simulation model and using these observations we optimise the fitting procedure.

1 Introduction

In the late 80's, traffic measurement of high speed communication networks indicated unexpectedly high variability and burstiness over several time scales, which indicated the need of new modeling approaches capable to capture the observed traffic features. The first promising approach, the fractal modeling of high speed data traffic [7], resulted in a fertile research in traffic theory. Since that time a lot of effort was put also in the analysis and modeling of traffic traces of real communication networks.

Various statistical tests were applied and based on their results a series of traffic models were proposed to describe the experienced traffic behaviour of real data traffic. A group of these traffic models was developed for simulation based analysis of telecommunication systems. These models can generate new traffic traces whose statistical behaviour (at least with respect to some essential features) are similar to the one of the measured data set. There are various applications of simulation models. Occasionally, their use is easier in simulation experiences than to handle the original data set. Further more, they can be used for generating independent simulation runs to improve the confidence of simulation results. They also can be used in traffic measurement of real network, where an artificial "traffic generator" generates e.g. the background traffic of an analyzed traffic stream.

In this paper we extend two long range dependent traffic models to modify their marginal distributions while maintaining their long range behaviour. The first one is the Multifractal Wavelet Model (MWM) proposed in [12], and the second one is a special Markov Arrival Process (MAP) introduced in [5]. A common feature of these simulation models is that both of them approximate the multifractal scaling behaviour by fitting the Haar wavelet coefficients of the original data trace.

The main idea behind the modification of the MWM was to change the distribution of some aggregation level dependent multipliers (A_j), which affects the Marginal distribution, but maintains the Haar wavelet coefficients of the model (see Section 4.1.2). The modification of the MAP was somehow similar. In this case we used the same "multi dimensional" structure to ensure the same long range dependent behaviour, but we assumed an asymmetric behaviour of each "dimension" instead of the original symmetric structure. In both cases the main goal was to leave the above described statistical properties of the models unchanged and to capture the marginal distribution of the measured data set in the mean time.

The rest of the paper is organised as follows. The considered traffic properties are shown in Section 2 and the applied statistical analysis methods are demonstrated in Section 3. Section 4 summarizes the main features of the two considered traffic models and their application to the data sets. The following section, Section 5, provides the details of the fitting methods and the comparison of the main properties of the original trace and its approximate traces. Finally, the paper is concluded in Section 6.

2 Properties of traffic processes

The definition of long range dependence of traffic arrival processes is as follows. Let us divide the time axis into equidistant intervals of length Δ. The number of arrivals in the ith interval is denoted by X_i. $\mathcal{X} = \{X_i, i = 0, 1, \ldots\}$ is a stochastic process whose aggregated process is defined as follows:

$$\mathcal{X}^{(m)} = \{X_i^{(m)}\} = \left\{ \frac{X_1 + \ldots + X_m}{m}, \ldots, \frac{X_{mk+1} + \ldots + X_{(m+1)k}}{m}, \ldots \right\}$$

The autocorrelation function of $\mathcal{X}^{(m)}$ is:

$$r^{(m)}(k) = \frac{E\{(X_n^{(m)} - E(X^{(m)})) \cdot (X_{n+k}^{(m)} - E(X^{(m)}))\}}{E\{(X_n^{(m)} - E(X^{(m)}))^2\}}$$

The process $\mathcal{X}$ exhibits long-range dependence (LRD) of index β if its autocorrelation function can be realized as

$$r(k) \sim A(k)k^{-\beta}, \ k \to \infty$$

where A(k) is a slowly varying function.

2.1 Self-similar processes

Using the above definition of the aggregated process, $\mathcal{X}$ is

a) exactly self-similar if $\mathcal{X} \stackrel{d}{=} m^{1-H}\mathcal{X}^{(m)}$, i.e., if $\mathcal{X}$ and $\mathcal{X}^{(m)}$ are identical within a scale factor in finite dimensional distribution sense.

b) exactly second-order self-similar if $r^{(m)}(k) = r(k), \forall m,\ k \geq 0$

c) asymptotically second-order self-similar if $r^{(m)}(k) \to r(k),\ (k, m \to \infty)$

where H is the Hurst parameter, also referred to as the self-similarity parameter.

For exactly self-similar processes the scaling behavior, which is characterized by the Hurst parameter (H), can be checked based on any of the absolute moments of the aggregated process:

$$\log(E(|X^{(m)}|^q)) = \log(E(|m^{H-1}X|^q)) = q(H-1)\log(m) + \log(E(|X|^q)). \qquad (1)$$

According to (1), in case of a self-similar process, plotting $\log(E(|X^{(m)}|^q))$ against $\log(m)$ for fixed q results in a straight line. The slope of the line is $q(H-1)$. Based on the above observations the test is performed as follows. Having a series of length N, the moments may be estimated as

$$E(|X^{(m)}|^q) = \frac{1}{\lfloor N/m \rfloor} \sum_{i=1}^{\lfloor N/m \rfloor} |X_i^{(m)}|^q,$$

where $\lfloor x \rfloor$ denotes the largest integer number smaller or equal to x. To test for self-similarity $\log(E(|X^{(m)}|^q))$ is plotted against $\log(m)$ and a straight line is fitted to the curve. If the straight line shows good correspondence with the curve, then the process is self-similar and its Hurst-parameter may be calculated by the slope of the straight line. This approach assumes that the scaling behavior of all absolute moments, q, are the same and it is captured by the Hurst-parameter. If it is the case we talk about mono-fractal behavior. The variance-time plot, which is used widespread to gain evidence of self-similarity, is the special case with $q = 2$. It depicts the behavior of the 2nd moments for the centered data.

It is worth to point out that self-similarity and stationarity imply that either $E(X) = 0$, or $E(X) = \pm\infty$, or $H = 1$. But $H = 1$ implies as well that $X_i = X_j,\ \forall i, j$ almost surely. As a consequence, to test for statistical self-similarity makes sense only having zero-mean data, i.e., the data has to be centered before the analysis.

2.2 Multi-fractal processes

Statistical tests of self-similarity try to gain evidence through examining the behavior of the absolute moments $E(|X^{(m)}|^q)$. In case of monofractal processes the scaling behavior of all absolute moments is characterized by a single number, the Hurst parameter. Multifractal processes might exhibit different scaling for different absolute moments. Multifractal analysis looks at the behavior of $E(|X^{(m)}|^q)$ for different values q and results in a *spectrum* that illustrates the behavior of the absolute moments. This analysis procedure is detailed in Section 3.2.

3 Statistical analysis methods

3.1 Tests for long range dependency

Recently, it has been agreed [7, 9, 10] that when one studies the long-range dependence of a traffic trace the most significant parameter to be estimated is the degree of self-similarity, usually given by the so-called Hurst-parameter. The aim of the statistical approach, based on the theory of self-similarity, is to find the Hurst-parameter.

In this section methods for estimating the long-range dependence of data sets are recalled. Beside the procedures described here, several other can be found in the literature. See [1] for an exhaustive discussion on this subject.

It is important to note that the introduced statistical tests of self-similarity, based on a finite number of samples, provide an approximate value of H only for the considered range of scales. Nothing can be said about the higher scales and the asymptotic behavior based on these tests.

Throughout the section, we illustrate the application of the estimators on the first trace of the well-known Bellcore data set set that contains local-area network (LAN) traffic collected in 1989 on an Ethernet at the Bellcore Morristown Research and Engineering facility. It may be downloaded from the WEB site collecting traffic traces [13]. The trace was first analyzed in [3].

3.1.1 Variance-time plot

One of the tests for *pseudo* self-similarity is the variance-time plot. It is based on the fact that for self-similar time series $\{X_1, X_2, \ldots\}$

$$\mathrm{Var}(X^{(m)}) \sim m^{-\beta}, \quad \text{as} \quad m \to \infty, \ 0 < \beta < 1.$$

The variance-time plot depicts $\log(\mathrm{Var}(X^{(m)}))$ versus $\log(m)$. For *pseudo* self-similar time series, the slope of the variance-time plot $-\beta$ is greater than -1. The Hurst parameter can be calculated as $H = 1 - (\beta/2)$. A traffic process is said to be *pseudo* self-similar when the empirical Hurst parameter is between 0.5 and 1.

The variance-time plot for the analyzed Bellcore trace is depicted in Figure 1. The Hurst-parameter given by the variance-time plot is 0.80.

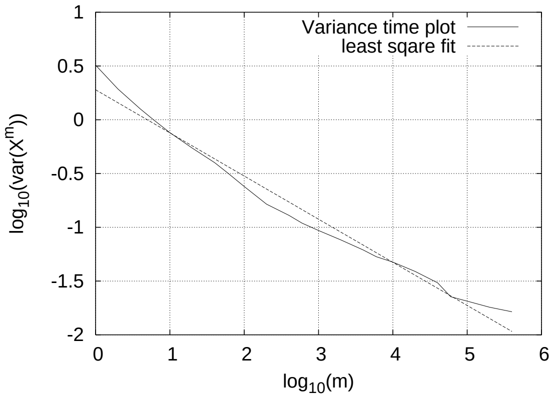


Figure 1: Variance-time plot and its least square fit for the Bellcore trace

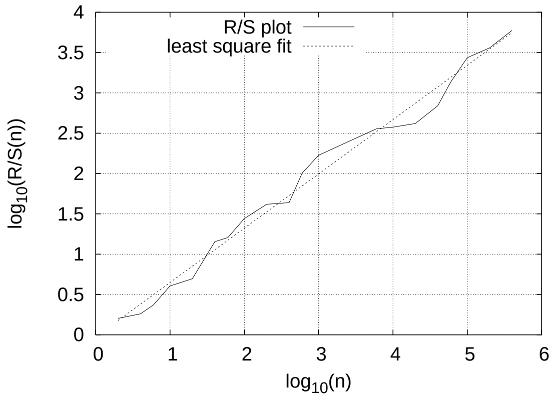


Figure 2: R/S plot and its least square fit for the Bellcore trace

3.1.2 R/S plot

The R/S method is one of the oldest tests for self-similarity, it is discussed in detail in [8]. For interarrival time series, $\mathcal{Z} = \{Z_i, i \geq 1\}$, with partial sum $Y_n = \sum_{i=1}^{n} Z_i$, and sample variance

$$S^2(n) = \frac{1}{n}\sum_{i=1}^{n} {Z_i}^2 - \frac{1}{n^2} \cdot Y_n^2,$$

the R/S statistic, or the rescaled adjusted range, is given by:

$$R/S(n) = \frac{1}{S(n)}\left[\max_{0 \leq k \leq n}\left(Y(k) - \frac{k}{n}Y(n)\right) - \min_{0 \leq k \leq n}\left(Y(k) - \frac{k}{n}Y(n)\right)\right].$$

$R/S(n)$ is the scaled difference between the fastest and the slowest arrival period considering n arrivals. For stationary LRD processes $R/S(n) \approx (n/2)^H$. To determine the Hurst parameter based on the R/S statistic the data set is divided into blocks, $\log[R/S(n)]$ is plotted versus $\log n$ and a straight line is fitted on the points. The slope of the fitted line is the estimated Hurst parameter.

The R/S plot for the analyzed Bellcore trace is depicted in Figure 2. The Hurst-parameter determined based on the R/S plot is 0.67.

3.1.3 Whittle estimator

The Whittle estimator is based on the maximum likelihood principle assuming that the process under analysis is Gaussian. The estimator, unlike the previous ones, provides the estimate through a non-graphical method. This estimation takes more time to perform but it has the advantage of providing confidence intervals as well. For details see [4, 1].

3.2 Multifractal framework

In this section we introduce two techniques to analyze multifractal processes.

3.2.1 Legendre spectrum

Considering a continuous-time process $\mathcal{Y} = \{Y(t), t > 0\}$ the scaling of the absolute moments of the increments is observed through the *partition function*

$$T(q) = \lim_{n\to\infty} \frac{1}{-n} \log_2 E\left[\sum_{k=0}^{2^n-1} |Y((k+1)2^{-n}) - Y(k2^{-n})|^q\right]. \quad (2)$$

Then, a multifractal spectrum, the so-called *Legendre spectrum* is given as the *Legendre transform* of (2)

$$f_L(\alpha) = T^*(\alpha) = \inf_q (q\alpha - T(q))$$

Since $T(q)$ is always concave, the Legendre spectrum $f_L(\alpha)$ may be found by simple calculations using that

$$T^*(\alpha) = q\alpha - T(q), \text{ and } (T^*)'(\alpha) = q \text{ at } \alpha = T'(q). \quad (3)$$

Let us mention here that there are also other kinds of fractal spectrum defined in the fractal world (see for example [11]). The Legendre spectrum is the most attractive one from numerical point of view, and even though in some cases it is less informative than, for example, the large deviation spectrum, it provides enough information in the cases considered herein.

In case of a discrete-time process $\mathcal{X}$ we assume that we are given the increments of a continuous-time process. This way, assuming that the sequence we examine consists of $N = 2^L$ numbers, the sum in (2) becomes

$$S_n(q) = \sum_{k=0}^{N/2^n-1} |X_k^{(2^n)}|^q, \ 0 \le n \le L, \quad (4)$$

where the expectation is ignored. Ignoring the expectation is accurate for small n, i.e., for the finer resolution levels. In order to estimate $T(q)$, we plot $\log_2(S_n(q))$ against $(L-n)$, $n = 0, 1, ..., L$, then $T(q)$ is found by the slope of the linear line fitted to the curve. If the linear line shows good correspondence with the curve, i.e., if $\log_2(S_n(q))$ scales linearly with $\log(n)$, then the sequence $\mathcal{X}$ can be considered a multifractal process.

Figure 3, 4, 5 and 6 illustrate the above described procedure to obtain the Legendre spectrum of the Bellcore *pAug* traffic trace. Figure 3 depicts the scaling behavior of the log- moments calculated by (4). With q in the range $[-3, 4]$, excluding the four finest resolution levels $n = 0, 1, 2, 3$ the moments show good linear scaling. For values q outside the range $[-5, 5]$ the curves are less and less linear. According to [12] one may look at non-integer values of q as well, but in general it does not provide notably more information on the process. To visualize the deviation from linearity better, Figure 4 depicts the increments of the log-moment curves of Figure 3. Completely horizontal lines correspond linear log-moment curves.

The partition function $T(q)$ with various initial value of n is depicted in Figure 5. Our review is about to find the widest range (in n), where it shows good linearity. Since the partition function calculated from the log-moments varies only a little (its derivative

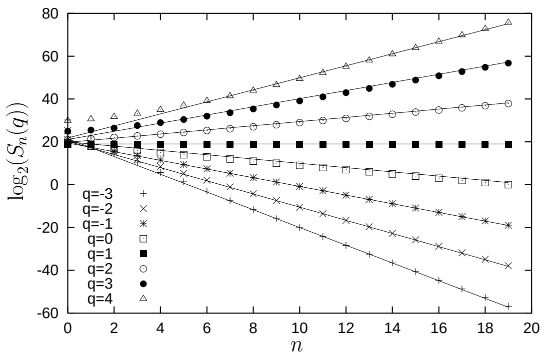


Figure 3: Scaling of log-moments with linear fits for the interarrival times of the Bellcore *pAug* trace

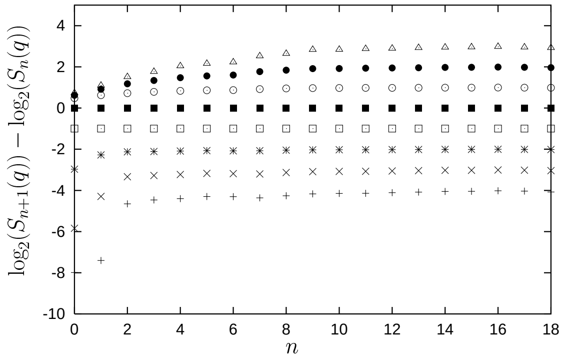


Figure 4: Increments of log-moments for the interarrival times of the Bellcore *pAug* trace

is in the range $[0.8, 1.15]$), it is not as informative as its Legendre transform is (Figure 6). As one can see using different intervals the result is not the same. In the rest of this paper we choose $n \in [4, 19]$ for the original Bellcore trace. For all the traces in the sequel we applied the same procedure to find the right range of n.According to (3) the Legendre spectrum is as wide as wide the range of derivatives of the partition function is, i.e., the more the partition function deviates from linearity the wider the Legendre spectrum is. The Legendre transform significantly amplifies the scaling information, but it is also sensitive to the considered range of the log-moments curves.

See [12] for basic principles of interpreting the spectrum. We mention here only that a curve like the one depicted in Figure 6 reveals a *rich multifractal spectrum*.

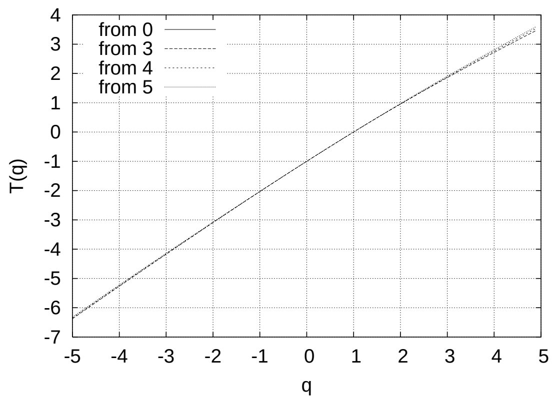


Figure 5: Partition function estimated through the linear fits shown in Figure 3, for different initial values of n

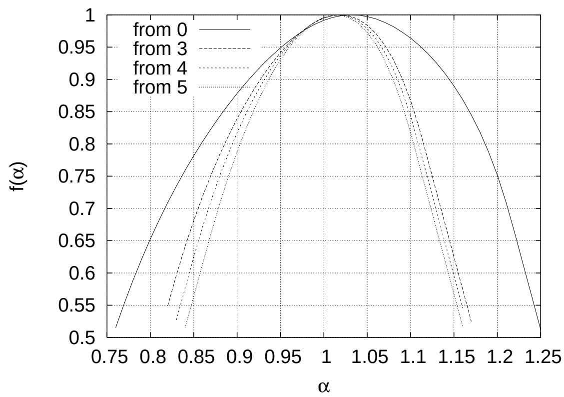


Figure 6: The Legendre transform of the partition functions shown in Figure 5 results in the Legendre spectrums

3.2.2 Haar wavelet

Another way to carry out multiscale analysis is the Haar wavelet transform. The choice of using the *unnormalized* version of the Haar wavelet transform is motivated by the fact

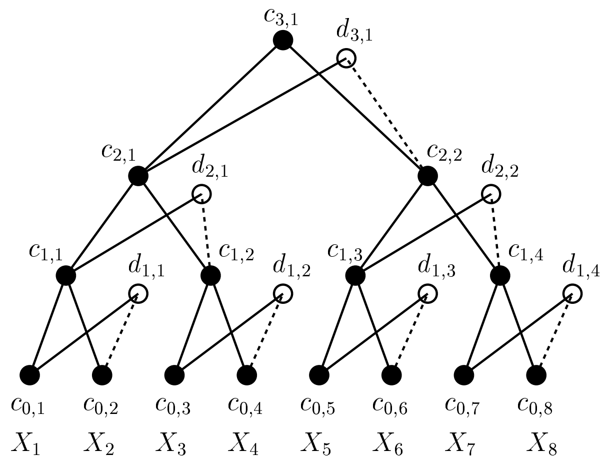


Figure 7: Haar wavelet transform

that it suits more the analysis of the Markovian point process introduced further on.

The multiscale behavior of the finite sequence $X_i, 1 \leq i \leq 2^L$ will be represented by the quantities $c_{j,k}, d_{j,k}, j = 0, \ldots, L$ and $k = 1, \ldots, 2^L/2^j$. The finest resolution is described by $c_{0,k}, 1 \leq k \leq 2^L$ which gives the finite sequence itself, i.e., $c_{0,k} = X_k$. Then the multiscale analysis based on the unnormalized Haar wavelet transform is carried out by iterating

$$c_{j,k} = c_{j-1,2k-1} + c_{j-1,2k}, \tag{5}$$

$$d_{j,k} = c_{j-1,2k-1} - c_{j-1,2k}, \tag{6}$$

for $j = 1, \ldots, L$ and $k = 1, \ldots, 2^L/2^j$. The quantities $c_{j,k}, d_{j,k}$ are the so-called scaling and wavelet coefficients of the sequence, respectively, at scale j and position k. At each scale the coefficients are represented by the vectors $\mathbf{c}_j = [c_{j,k}]$ and $\mathbf{d}_j = [d_{j,k}]$ with $k = 1, \ldots, 2^L/2^j$. For what concerns $\mathbf{c}_j$, the higher j the lower the resolution level at which we have information on the sequence. The information that we lost as a result of the step from $\mathbf{c}_{j-1}$ to $\mathbf{c}_j$, is conveyed by the sequence of wavelet coefficients $\mathbf{d}_j$. It is easy to see that $\mathbf{c}_{j-1}$ can be perfectly reconstructed from $\mathbf{c}_j$ and $\mathbf{d}_j$. As a consequence the whole $X_i, 1 \leq i \leq 2^L$ sequence can be constructed (in a top to bottom manner) based on a normalizing constant, $\mathbf{c}_L = c_{L,1} = \sum_{i=1}^{2^L} X_i$, and the $\mathbf{d}_j, j = 1, \ldots, L$ vectors.

By taking the expectation of the square of (5) and (6)

$$E[c_{j,k}^2] = E[c_{j-1,2k-1}^2] + 2E[c_{j-1,2k-1}c_{j-1,2k}] + E[c_{j-1,2k}^2], \tag{7}$$

$$E[d_{j,k}^2] = E[c_{j-1,2k-1}^2] - 2E[c_{j-1,2k-1}c_{j-1,2k}] + E[c_{j-1,2k}^2], \tag{8}$$

Let us assume that the series we analyze are stationary; then, by summing (7) and (8) and rearranging the equation, we have

$$E[c_{j-1}^2] = \frac{1}{4}\left(E[d_j^2] + E[c_j^2]\right). \tag{9}$$

Similarly, by consecutive application of (9) from one scale to another, the $E[d_j^2], j = 1, \ldots, L$ series completely characterize the variance decay of the $X_i, 1 \leq i \leq 2^L$ sequence apart of a normalizing constant ($\mathbf{c}_L = c_{L,1} = \sum_{i=1}^{2^L} X_i$). This fact allows us to realize a series with a given variance decay if it is possible to control the 2nd moment of the scaling coefficient with the chosen synthesis procedure. In Section 4 we will discuss two methods that attempts to capture the multifractal scaling behavior via the series $E[d_j^2], j = 1, \ldots, L$.

One of them fits accurately these parameters and the other one minimizes the distance of the $E[d_j^2]$ series of the trace and the simulation model. We visualize the $E[d_j^2], j = 1, \ldots, L$ series versus j in Figure 8. We connect the $E[d_j^2]$ points only for better visualization.

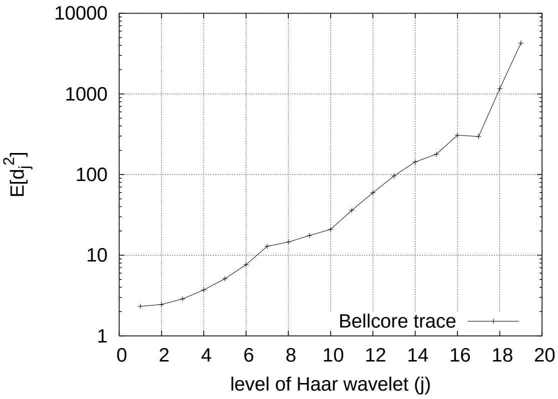


Figure 8: The Haar parameters $E[d_j^2]$ vs. j

3.3 Queue length behaviour

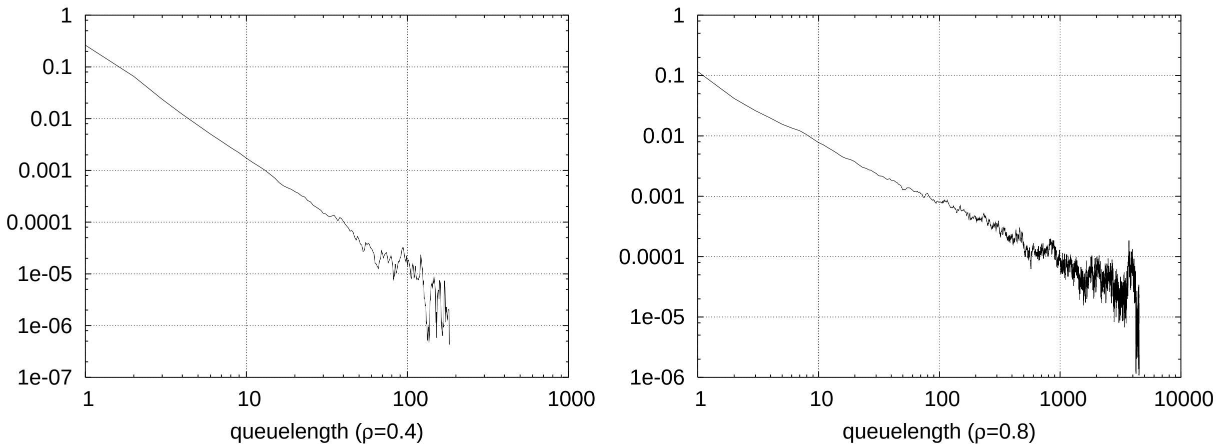
One of our primary goals is to find a traffic model, by which the queue length behavior is the same as by the original trace. For this reason, we examine the queue length distribution, using several utilizations (ρ). This is depicted at $\rho = 0.4$ and at $\rho = 0.8$ in Figure 9 for the Bellcore trace.

We used deterministic service time, to calculate the queue length using different traces as arrival process. The utilization of the system can be adjusted by the service time: $\rho = \frac{E[\text{service time}]}{E[\text{interarrival time}]}$, where $E[\text{service time}]$ is deterministic, and $E[\text{interarrival time}]$ is calculated from the trace.

3.4 Marginal distribution and the moments of the trace

Considering the interarrival times in the trace like a stochastic process, we can calculate its empirical marginal distribution with these rules:

- let the time axis divide into equidistant intervals between the smallest and the largest interarrival time,
- count the number of interarrival times which are in the $1^{st}, \ldots, i^{th}$ interval,
- normalize the value and then plot it versus the ordinal number of the interval.



$\rho = 0.4$ $\rho = 0.8$

Figure 9: The probabilities vs. the queue length for BC trace with different utilizations

The complementary cumulative distribution function (ccdf) equals to $1 -$ marginal distribution. We will observe the ccdf (see Figure 10 for the ccdf of the Bellcore trace) of the traces.

Another way to compare two distributions is to calculate theirs moments. Two equivalent distribution have the same moments. We will calculate the first twelve moments of the traces (Figure 11) and of its approximates. The empirical n^{th} moment of a data set can be calculated using this rule:

$$E[X^n] = \frac{\sum_{i=1}^{\text{number of arrivals}} \text{interarrival time}_i^n}{\text{number of arrivals}}$$

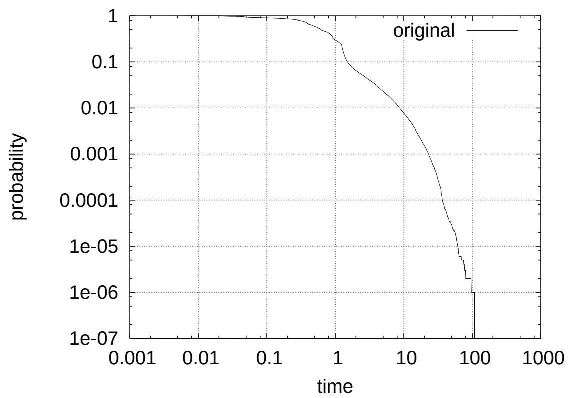


Figure 10: ccdf of the original Bellcore trace

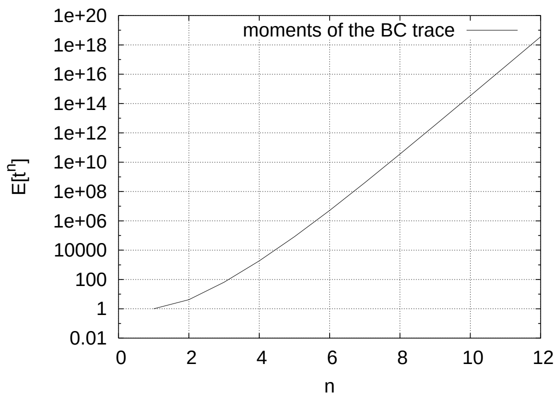


Figure 11: The first twelve moments of the Bellcore trace

4 Long range dependent simulation models

4.1 Multifractal wavelet model (MWM)

This wavelet-based simulation model is for positive, stationary, LRD data. While characterizing positive data in the wavelet domain is problematic for general wavelets, for the Haar wavelet, we have the simple condition: $X(t)$ is positive if and only if $|d_{j,k}| \leq c_{j,k}$ for all j, k.

In the multifractal wavelet model, we ensure a positive signal output by modeling the wavelet coefficients as

$$d_{j,k} = A_{j,k} * c_{j,k}, \tag{10}$$

with the multipliers $A_{j,k}$, the realizations of the independent random variables A_j supported on $[-1, 1]$ with mean 0.

The MWM flows as a multiscale, coarse-to-fine synthesis down the tree in Figure 7. Given the approximation to $X(t)$ at resolution 2^{-j} $(c_{j,k})$, we compute the wavelet coefficients $d_{j,k} = A_{j,k} * c_{j,k}$ with random $A_{j,k}$. The approximation to $X(t)$ at resolution $2^{-(j-1)}$ $(c_{j-1,k})$ is then obtained from scaled sums and differences of the $c_{j,k}$ and $d_{j,k}$ as we told in Section 3.2.2. This process can be iterated until any desired resolution (signal length) reached.

4.1.1 Application of MWM to measured data sets

As mentioned in [12] we can control the energy decay of the simulated trace through the 2nd moment of the chosen distribution of A_j.

The way to control the energy scaling is to set the energy of the coarsest scale $(j = L)$ and then calculate ratios of energy for the other scales with $\eta_j = \frac{var(d_{j,k})}{var(d_{j-1,k})}$, $0 \leq j < L$. In our special case we can calculate η_j via

$$\begin{aligned}\eta_j &= \frac{E[d^2_{j+1,k}]}{E[d^2_{j,k}]} \\ &= \frac{2E[A^2_{j+1}]E[c^2_{j+1,k}]}{E[A^2_j]E[1 + A^2_{j+1}]E[c^2_{j+1,k}]} \\ &= 2\frac{E[A^2_{j+1}]}{E[A^2_j](1 + E[A^2_{j+1}])}.\end{aligned} \tag{11}$$

To match the variance decay of the measured data set we can recursively solve (11) for $E[A^2_j]$ in terms of η_j and $E[A^2_{j+1}]$ for $j = L-1, \ldots, 0$. First we initialize the calculation at the coarsest scale by setting

$$E[A^2_L] = \frac{E[d^2_{L,1}]}{E[c^2_{L,1}]}. \tag{12}$$

After that we have the 2nd moments of the distributions we are ready to start the following synthesis procedure:

1. Set $j = L$. Compute the 2nd moment of the multiplier's distribution with the rule (12).

2. At scale j, generate the random multipliers $A_{j,k}$ according to the chosen distribution function and calculate each $d_{j,k}$ via (10) for $k = 0, \ldots, 2^j - 1$

3. At scale j, use $c_{j,k}$ and $d_{j,k}$ to calculate $c_{j-1,2k}$ and $c_{j-1,2k+1}$, the scaling coefficients at scale $j-1$ for $k = 0, \ldots, 2^j - 1$.

4. Iterate steps 2 and 3, replacing j by $j-1$ until $j = 0$.

4.1.2 Distributions for the multipliers

To capture the marginal distribution of the measured data sets we have used two distributions (the beta and the uniform distributions) for the multipliers $A_{j,k}$ with significant differences in theirs moments (see Figure 12). Both of them are easy to calculate.

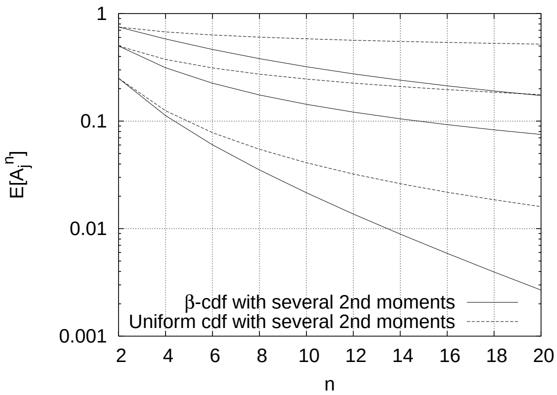


Figure 12: The even moments of the used distributions for the multipliers $A_{j,k}$ with 2nd moments $\{0.25; 0.5; 0.75\}$

Symmetric beta distribution A $\beta(p,p)$ random variable A, symmetrically distributed over $[-1,1]$, has the following pdf:

$$g_A(x) = \frac{(1+x)^{p-1}(1-x)^{p-1}}{\mathcal{B}(p,p)2^{2p-1}}. \tag{13}$$

Here $\mathcal{B}(.,.)$ is the beta function, and $p > 0$ is a shape factor. In this symmetric case the variance equals to the second moment, thus:

$$var(A) = E[A^2] = \frac{1}{2p+1}. \tag{14}$$

The parameter p can be calculated from (14) knowing $E[A_j^2]$ (which recursively depends on η_j and $E[A_{j+1}^2]$). One can set the required energy decay through this parameter. The symmetric beta distribution is depicted in Figure 13 with several parameters p.

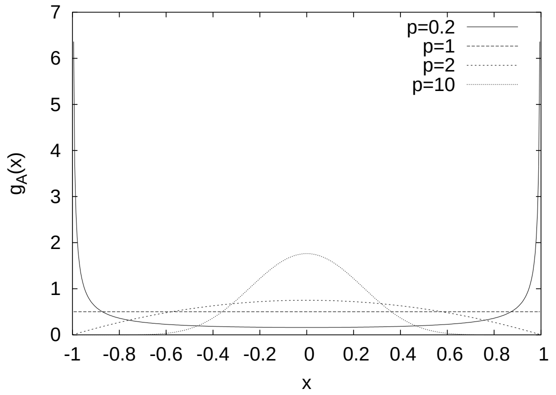


Figure 13: Beta distribution with several parameters

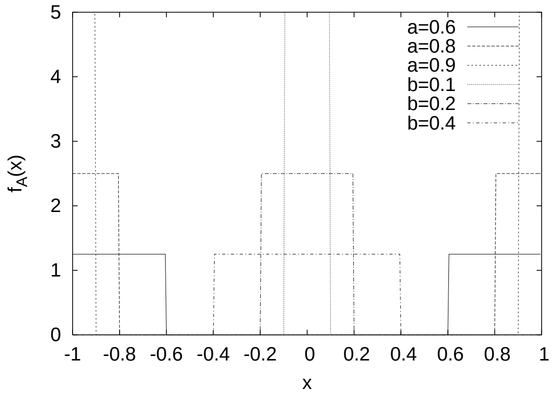


Figure 14: Uniform distribution with several parameters

A special case of uniform distributions As written above in Section 4.1 the distribution function for multipliers have to be symmetrical over $[-1, 1]$, with mean value 0.

The symmetric uniform distributions over $[-1, 1]$ match these criteria. It has parameter a or b depending on the variance (remember it is a symmetric distribution thus $var(A) = E[A_j^2]$). If the second moment is in the range $]1/3, 1]$ then we use the distribution described by the parameter a, if it is less or equal to $1/3$ we use the another one described by b. The following density function defines this random variable:

$$f_A(x) = \begin{cases} \dfrac{1}{2(1-a)}, \text{ if } 1/3 < E[A_j^2] = \dfrac{a^2+a+1}{3} \leq 1 \\ \dfrac{1}{2b}, \text{ if } 0 \leq E[A_j^2] = \dfrac{b^2}{3} \leq 1/3 \end{cases}$$

The uniform distribution, with different values of the parameters a and b, is depicted in Figure 14.

4.2 MAP based simulation

Let $Z(t)$ be an irreducible Markov chain with finite state space of size m and generator $\mathbf{Q}$. An arrival process (this is the so-called Markov Arrival Process – MAP) is associated with this Markov chain in the following way:

- while the Markov chain stays in state i arrival occurs at rate λ_i,
- when the Markov chain undergoes a state transition from i to j arrival occurs with probability p_{ij}.

The standard description of MAPs are given with matrices $\mathbf{D_0}$ and $\mathbf{D_1}$ of size $(m \times m)$, where $\mathbf{D_0}$ contains the transition rates of the Markov chain which are not accompanied with arrivals and $\mathbf{D_1}$ contains the transition rates which are accompanied with arrivals, i.e.:

- $\mathbf{D_0}_{ij} = (1 - p_{ij})\mathbf{Q}_{ij}$, for $i \neq j$ and $\mathbf{D_0}_{ii} = \mathbf{Q}_{ii} - \lambda_i$;
- $\mathbf{D_1}_{ij} = p_{ij}\mathbf{Q}_{ij}$ for, $i \neq j$ and $\mathbf{D_1}_{ii} = \lambda_i$.

Many familiar arrival processes represent special cases of MAPs:

- the Poisson process (MAP with a single state),
- interrupted Poisson process: a two-state MAP in which arrivals occur only in one of the states and state jumps do not cause arrival,
- Markov modulated Poisson process (MMPP): state jumps do not give rise to arrivals.

The class of MAPs is closed for superposition and Markovian splitting.

4.2.1 Application of MMPP to measured data sets

In [6] a special MMPP structure is proposed to exhibit multifractal behavior. The background continuous time Markov chain (CTMC) of the MMPP has a symmetric n-dimensional cube structure and the arrival intensities are set according to the variation of the arrival process at the different time scales. The special choice of the structure is motivated by the generation of the Haar wavelet transform. Basically the Haar wavelet transform evaluates the variation of the data set at different aggregation levels (time scales), and similarly, the proposed MMPP structure provide different variation of the arrival rate at different time scales.

The composition of the proposed MMPP structure is similar to the generation of the Haar wavelet transform (the another procedure for traffic trace generation based on this transform is introduced above in Section 4.1). Without loss of generality, we assume that the time unit is such that the long term arrival intensity is one. A MMPP of one state with arrival rate 1 represents the arrival process at the largest (considered) time scale.

At the next time scale, $1/\lambda$, an MMPP of two states with generator

$$\boxed{\begin{array}{cc} -\lambda & \lambda \\ \lambda & -\lambda \end{array}}$$

and with arrival rates $1 - a_1$ and $1 + a_1$ $(-1 \leq a_1 \leq 1)$ represents the variation of the arrival process. This composition leaves the long term average arrival rate unchanged.

In the rest of the composition we perform the same step. We introduce a new dimension (Figure 15) and generate the n-dimensional cube such that the behavior at the already set time scales remains unchanged. E.g., considering also the $1/\gamma\lambda$ time scale an MMPP of four states with generator

$$\left|\begin{array}{cc|cc} \hline \bullet & \lambda & \gamma\lambda & \\ \lambda & \bullet & & \gamma\lambda \\ \hline \gamma\lambda & & \bullet & \lambda \\ & \gamma\lambda & \lambda & \bullet \\ \hline \end{array}\right|$$

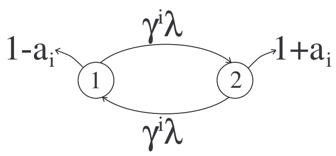


Figure 15: One dimension of the symmetric Markov chain

and with arrival rates $(1-a_1)(1-a_2)$, $(1+a_1)(1-a_2)$, $(1-a_1)(1+a_2)$ and $(1+a_1)(1+a_2)$ $(-1 \leq a_1, a_2 \leq 1)$ represents the variation of the arrival process. With this MMPP, parameter a_1 (a_2) determines the variance of the arrival process at the $1/\lambda$ $(1/\gamma\lambda)$ time scale.

A level n MMPP of the proposed structure is composed by 2^n states and it has $n+2$ parameters. Parameters γ and λ defines the considered time scales, and parameters $a_1, a_2, \ldots, a_n$ determines the variance of the arrival process at the n considered time scales. It can be seen that the ratio of the largest and the smallest considered time scales is γ^n. Having a fixed n (i.e., a fixed cardinality of the MMPP), any large ratio of the largest and the smallest considered time scales can be captured by using a sufficiently large γ.

A simple numerical procedure can be applied to fit a MMPP of the given structure to a measured data set. This heuristic approach is composed by "engineering considerations" based on the properties of the measured data set and a parameter fitting method.

First, we fix the value of n. According to [5] a "visible" multiscaling behavior can be obtained from $n = 3 \sim 4$. The computational complexity of the fitting procedure grows exponentially with the dimension of the MMPP. The response time with $n = 6$ (MMPP of 64 states) is still acceptable (in the order of minutes).

Similarly to [12], we set γ and the λ based on the inspection of the data set. Practically, we define the largest, T_M, and the smallest, T_m, considered time scales and calculate γ and λ from

$$T_M = \frac{1}{\lambda} \; ; \qquad T_m = \frac{1}{\gamma^n \lambda}.$$

Having γ and λ, we apply a downhill simplex method to find the optimal values of the variability parameters $a_1, a_2, \ldots, a_n$. The goal function that our parameter fitting method minimizes is the sum of the relative errors of the second moment of Haar wavelet coefficients up to a predefined time scale S:

$$\min_{a_1,\ldots,a_n} \sum_{j=1}^{S} \frac{|E(d_j^2) - E(\hat{d}_j^2)|}{E(d_j^2)} \ .$$

The goal function can be calculated analytically as it is described in [6].

4.2.2 Asymmetric MMPP model

Based on the fitting method described above in Section 4.2.1 with symmetrical transition rates in each dimension of the Markov chain we have set the stationary probabilities of the two states in each level to $1-p$ for state 1 (and to p for state 2). Thus the transition rates become: $\alpha_i = \frac{\gamma^i \lambda}{2(1-p)}$, $\beta_i = \frac{\gamma^i \lambda}{2p}$, according to the aim to leave the cycle time of the given level of the Markov chain untouched ($\frac{2}{\gamma^i \lambda} = \frac{1}{\alpha_i} + \frac{1}{\beta_i}$). The arrival rates in each dimension were changed to ensure that the average arrival rate ($\frac{(1-p)\delta_{i-} + p\delta_{i+}}{2}$) equals to 1. According to this the arrival intensities was set to $\delta_{i-} = (1 - \frac{a_i}{1-p})$ and $\delta_{i+} = (1 + \frac{a_i}{p})$. The changes depicted in Figure 16.

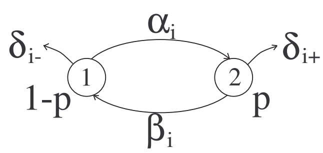


Figure 16: One dimension of the asymmetric Markov chain

With these changes we can control the arrival rates of the MMPP and the length of cycle time in each level of the Markov chain. The moments of the generated trace can be set through these parameters.

5 Application of the new simulation models to measured traffic traces

In this section we present the new fitting methods and the differences between the synthetic and the original traces.

5.1 Application of the modified MWM

Since we already computed the statistical measures of the Bellcore trace (in Section 3) the only thing we have to do is to compare it with the synthetic one.

As written in Section 4.1.1 we used two distributions for multipliers $A_{j,k}$. The MWM with beta distribution was applied and analysed in [12, 2] and found that approximates well the long range dependency and the multifractal behaviour. Throughout the fitting method with the new uniform distribution, our goal was to keep all of these good properties and additionally capture the marginal distribution of the original trace better.

Variance time plot Based on the assumption that the process is monofractal and characterized by a single number, this test indicates the long range dependency of the trace through the Hurst parameter, which is also referred to self-similarity parameter. The calculation of the Hurst parameters is the following.

- Plot $\log(\mathrm{Var}(X^{(m)}))$ versus $\log(m)$.
- Fit a straight line to the graph.

- When the slope of the line is $-\beta$ then the Hurst parameter is calculated as: $H = 1 - (\beta/2)$.

A trace with Hurst parameter between 0.5 and 1 based on this test is said to be *pseudo* self-similar. The variance time plot of the Multifractal Wavelet model with the beta and with the uniform distribution and the original trace are depicted in Figure 17.

The self-similarity parameters of the different traces are

- $H_{\text{original}} = 0.80$,
- $H_{\beta\text{MWM}} = 0.75$,
- $H_{\text{uniform}} = 0.76$.

The Hurst parameters of the trace and the MWMs are fairly close. Which means, that both models fit the LRD quite well. We note that a perfect fitting in the whole range of aggregation levels (m) means only that the second order statistical parameters are identical.

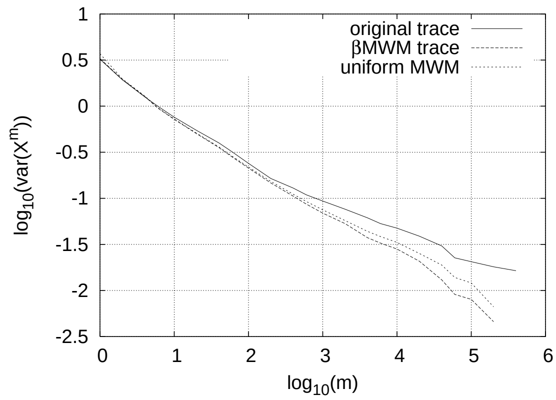


Figure 17: The Variance time plot for the original and the different MWM traces

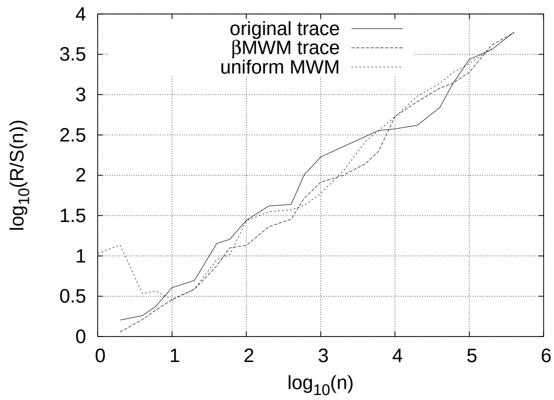


Figure 18: The R/S plot for the original and the different MWM traces

R/S plot Another test for *pseudo* self-similarity is the R/S plot. This plot is used as follows:

- the logarithm of the scaled sums between fastest and slowest arrival period considering n arrivals is depicted versus the logarithm of n,
- then a line is fitted on it and the slope of the line is the Hurst parameter.

As one can see in the Figure 18 the three plots are similar in almost the whole range of n except that the MWM with uniform distribution differs in low scales, which means that the model can not capture the long range dependency at low time scales, or the variance of the values considering few arrivals, is high. Apart from this all the three plots show good linearity, and fitting the line to the uniform-MWM only in the range $n \in [0.6, 5.3]$ we get the following Hurst parameters:

- $H_{\text{original}} = 0.67$,
- $H_{\beta\text{MWM}} = 0.71$,
- $H_{\text{uniform}} = 0.70$.

This test confirms that the simulation models approximate the LRD of the trace reasonably well.

Legendre spectrum Executing the algorithm of Section 3.2.1 both for the βMWM and for the uniform-MWM models we have the Legendre Spectrums with parameter q in the $[-5, 5]$ range and fitting straight lines to the $n = 4$ to $n = 19$ range. The results of these settings are depicted in Figure 19 and 20.

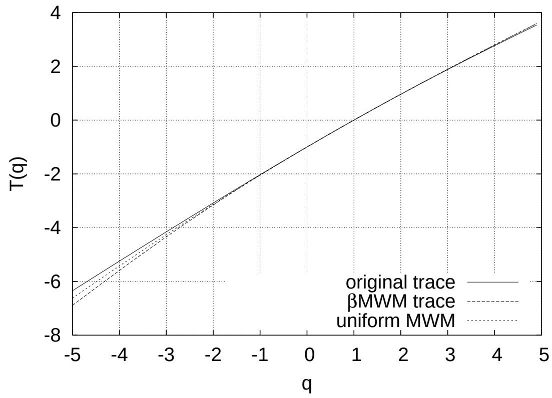


Figure 19: The partition functions of the original and the two kinds of MWM traces

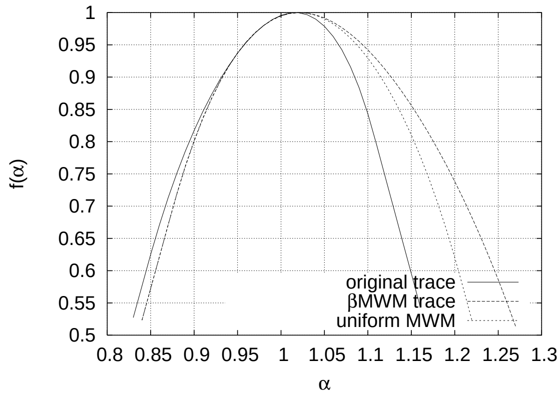


Figure 20: The Legendre transform of the partition functions in Figure 19

As one can see the Legendre spectrums of the Multifractal Wavelet Models differs from the original at high αs. This is because the high negative moments of the traces were not the same, but according to our statements above we examine only the *richness* of the *multifractal spectrum* and all the three curves depicted in Figure 20 have a rich multifractal spectrum. But we notice that the differences of the Legendre spectrum might be caused by the numerical method which we used to calculate the approximate traces. Considering narrower ranges of n the lines fitted to the log-moments curves resembles better and better to the original one, but we can not say anything about the lower time scales.

Haar wavelet The other way to observe the multifractal behaviour of a trace is through the Haar wavelet transform. In our case it has a special role because both models approximates the Haar wavelet parameters $E[d_j^2]$. These parameters are depicted in Figure 21 where we connect the points again for better visualization.

The significant differences between of the $E[d_j^2]$ parameters at higher j values are because the number of wavelet coefficients is not sufficient for a confident estimate of the experimental 2nd moment. (E.g. we have 2 values to consider when $j = 19$ and for values when $j = 18$.)

The $E[d_j^2]$ parameters of the original and the approximating trace can differ for all the scales because theoretically we exactly fit the second moments of the Haar wavelet parameters of the trace, but the parameters of the simulation models are calculated based on a finite sample series. If we had an infinite number of $A_{j,k}$s the fit would be perfect.

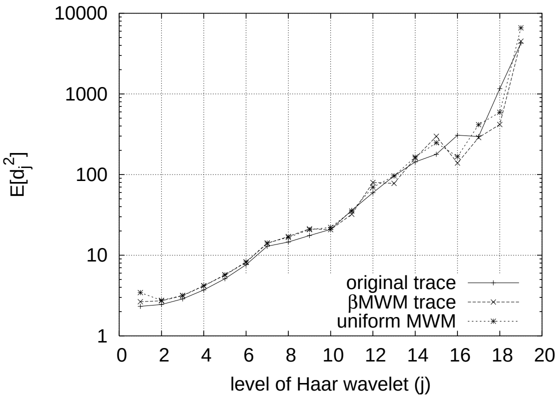


Figure 21: The Haar wavelet parameters of the original and the two kinds of MWM traces vs. j

Queue length behaviour The queue length distribution is one of the most important performance parameter in practice. We also compare behaviour of our simulation models with respect to the queue length distribution. Throughout the examination of the queue length we assume that the system has an infinite buffer and a deterministic service time. Based on this assumptions we can set the utilization (ρ) of the system by properly choosing the service time according to the following relation

$$\rho = \frac{E[\text{service time}]}{\text{E[interarrival time]}}. \tag{15}$$

We have depicted the queue length distribution of the system with four values of ρ in Figure 22 according to (15).

Taking a look at Figure 22 one can say that both of the two models approximates well the queue length distribution of the original trace with higher utilizations but with low system utilization the difference is large.

A possible reason of the poor approximation of the queue length distribution at low utilization is the presence of strange regular sequences of rather short length in the original trace. The investigation of this behaviour requires further research, especially because neither the dependency structure nor the marginal distribution (as it is presented below) of the approximate models cause this difference in an evident way.

Marginal distribution Regarding the traffic trace as a finite realization of a stochastic process we can estimate the marginal distribution and its moments. The experimental ccdf and moments are calculated as it is introduced in Section 3.4.

These properties of the Bellcore trace and of its Multifractal Wavelet Model approximations can be seen in Figure 23 and 24.

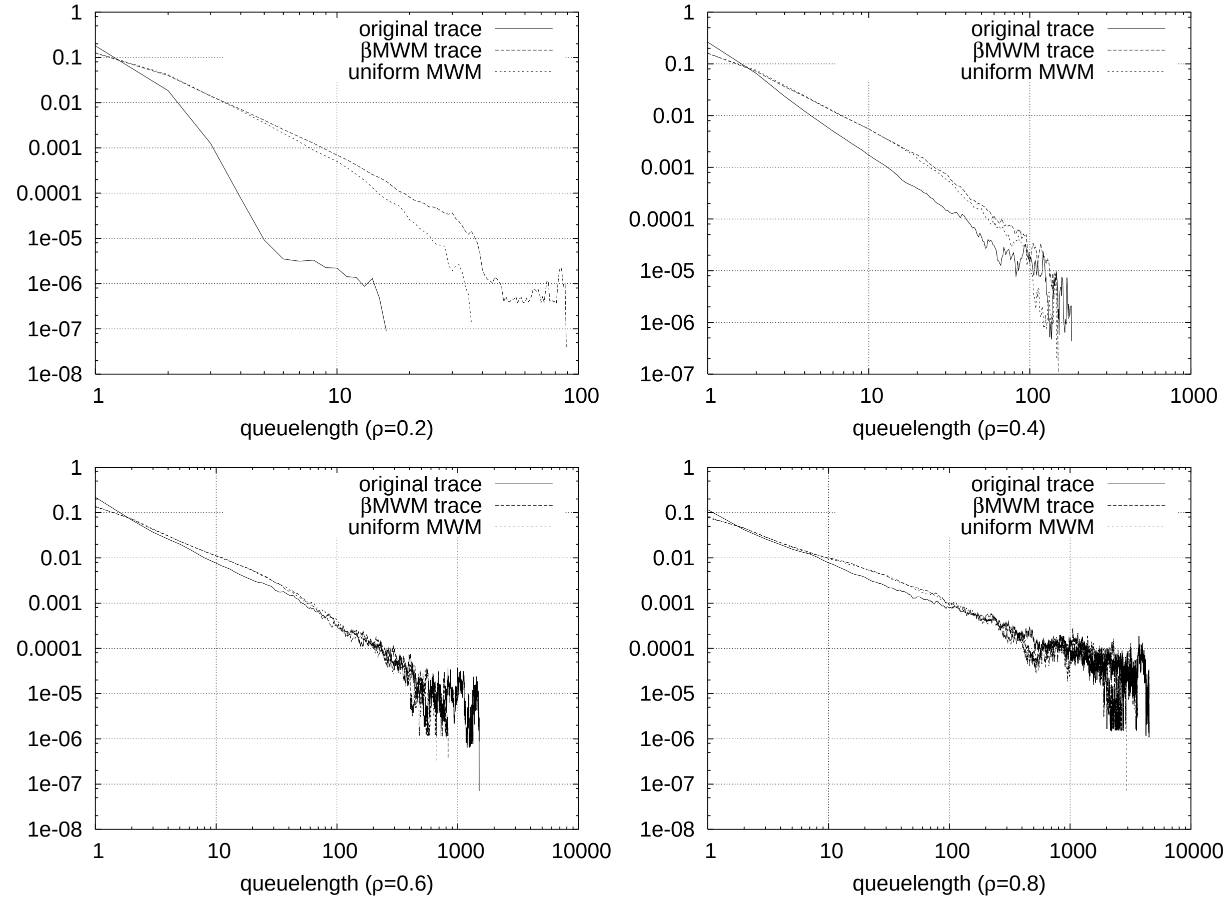


Figure 22: The probabilities vs. the queue length for the original and the MWM traces with several utilizations

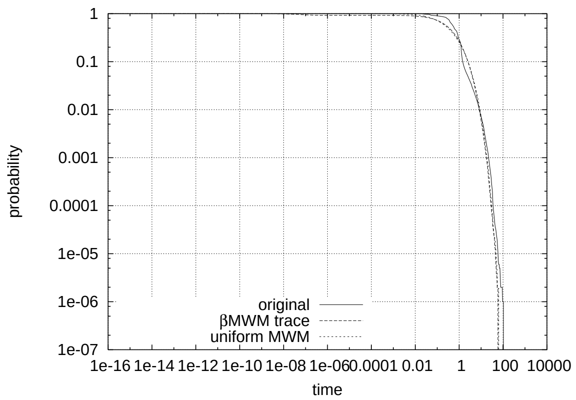


Figure 23: The complementary cumulative distribution functions of the three traces

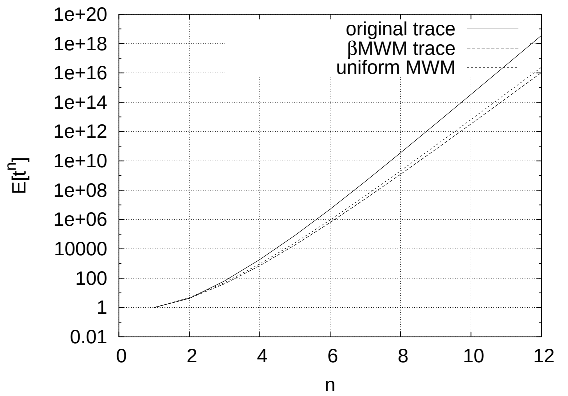


Figure 24: The first twelve moments of the three observed traces

Based on Figure 23 the marginal distributions looks very similar, but the differences are more visible from the series of moments on Figure 24 (where we also connected the discrete points with straight lines). In both cases the two kinds of MWM are very close to each other. Taking a look at the complementary cumulative distribution function we can see that the original trace is "ragged" while the distribution of the models are not, which it is rather natural for analytical models like our simulation models. This is the reason why the higher moments of the MWM simulated traces differs from the moments of the original trace. But the small differences between the distribution of the MWM traces suggests that the moments of the marginal distribution of the MWM process is rather independent of the moments of the distribution used for the $A_{j,k}$ multipliers.

Dependence on the initial seed value In [12] the authors report a strong dependency of the generated traces on the initial value of the random number generator. We also investigated this dependence, but we draw a different conclusion based on our experiences. In the majority of the cases we obtained very similar results with different initial seed values of the random number generator as it is in Figure 25 and 26.

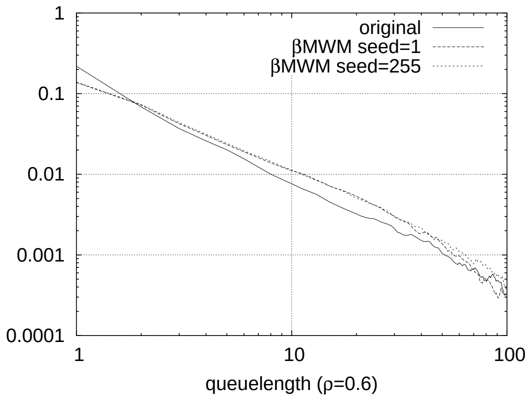


Figure 25: Queue length distribution with different seeds ($\rho = 0.6$)

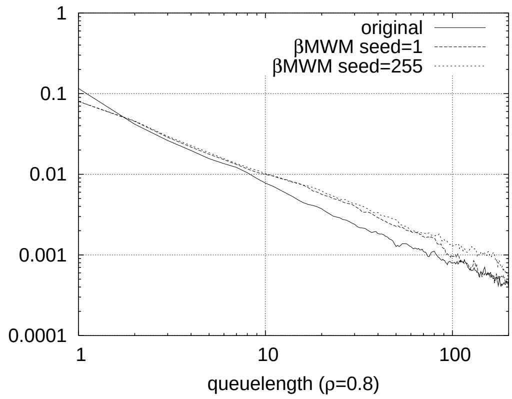


Figure 26: Queue length distribution with different seeds ($\rho = 0.8$)

5.2 Application of the asymmetric MMPP

In this section we complete a similar analysis for the MMPP models as before for the MWM. The goal of our investigation is to evaluate the effect of the modification of the original symmetric MMPP model presented in [5] to the modified asymmetric one presented in Section 4.2.2. With this modification we intend to maintain the dependency structure of the process, but modify the marginal distribution.

The used settings of the MMPP model was introduced and applied in [5]. According to the descriptions in Section 4.2.1 and in [5] the "user defined" parameters of the fitting method were:

- $n = 5$, according to that the complexity of the calculations grows exponentially with the number of the dimensions of the background Markov chain,
- $\frac{1}{\lambda}$ was set to the mean time of $N/4$ arrivals,
- γ was set to 8

- and various setting of the asymmetry parameters of the asymmetric model ($\{p_1, p_2, p_3, p_4, p_5\}$) were investigated.

Asymmetry parameters setting In the asymmetric MMPP model of Section 4.2.2 we introduced a level independent asymmetry parameters setting, i.e. $p_i = p, \forall i$, but for the investigation of the effect of the symmetry parameters requires us to evaluate the asymmetric MMPP model with level dependent asymmetry parameters. Note that, $p = 0.5$ represents the symmetric case.

First, we investigated the effect of using level independent asymmetry parameters different from 0.5. These kind of change have had a rude effect on the properties of the synthesized trace. It destroyed close fitting of the queue length distribution and had no effect on the marginal moments of the trace. Lower p values (more asymmetric setting) resulted in worse queue length fitting.

Next, we evaluated various level dependent settings. Changing the asymmetry only in the lower ($\{0.15, 0.5, 0.5, 0.5, 0.5\}$) or in the higher levels ($\{0.5, 0.5, 0.5, 0.5, 0.15\}$) of the Markov chain caused the tail of the queue length distribution to elongate. But changing the asymmetry parameter in the middle level of the Markov chain ($\{0.5, 0.5, 0.15, 0.5, 0.5\}$) improved the performance of the approximation a bit. This case asymmetry parameter setting is used in the analysis below.

Optimization with modified goal function To stress the role of the marginal distribution fitting we also tried to optimize the model parameters with a goal function which was sensitive to the marginal moments matching. We observed that the traces generated with the obtained MMPP model differed much from the original trace and from the trace generated by the symmetric MMPP.

Variance time plot The long range dependency of the original Bellcore trace and of the two kinds of MAP traces are compared through their self-similarity parameters, which we approximate first with variance time plot and than with R/S plot.

It can be seen in Figure 27 that all the three traces looks to be second order self-similar, i.e. long range dependent. The estimated Hurst parameters based on these test are:

- $H_{\text{original}} = 0.80$,
- $H_{\text{symmetric MAP}} = 0.82$,
- $H_{\text{asymmetric MAP}} = 0.80$,

but we note that in this case we depicted and calculated the Hurst parameters right after exploration of the plots and did not plotted the points (the time scales) which differed much from the linear trend. These points were the points of the symmetric and the asymmetric MAP simulated trace at high aggregation levels. It is because in the high aggregation levels there are only few values and they are not enough to calculate a reliable variance.

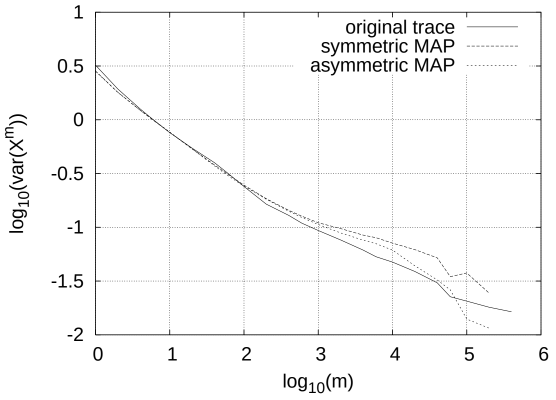


Figure 27: The Variance time plot for the original and for the two different MMPP traces

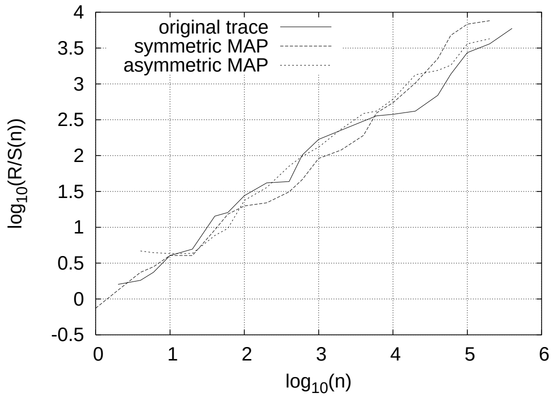


Figure 28: The R/S plot for the original and for the two different MMPP traces

R/S plot The second way to estimate the long range dependency is the use of R/S plot. The results of the fitting is depicted in Figure 28. The calculated Hurst-parameters are:

- $H_{\text{original}} = 0.67$,
- $H_{\text{symmetric MAP}} = 0.76$,
- $H_{\text{asymmetric MAP}} = 0.71$.

We have not plotted the points differing much from linear. According to the variance time plot and the R/S plot results we can say that the traces generated by the MMPP models approximate well the original trace.

Legendre spectrum When we use the Legendre spectrum to examine the multifractal behaviour of traffic traces the first thing we analyse is the widest range of n where the log-moments of the trace shows good linearity. Based on the algorithm presented in Section 3.2.1 we choose $n = 4$ for lower and $n = 19$ for upper bound for all traces.

According to our choice the Partition function (Fig. 29) and the Legendre spectrums (Fig. 30) are very close to each other and to the original trace too. All three traces shows *rich multifractal spectrum.*

Haar wavelet The scaling behaviour of the models were compared with their Haar wavelet parameters as well. We connect the discrete points of $E[d_j^2]$ versus j also in Figure 31.

As one can see the variance parameters of the higher scales differs significantly from the original. This is a result of the applied fitting method. The setting of $T_M = 1/\lambda$ determines the largest considered time scale. We set λ, such that T_M is the mean time of $N/4$ arrivals. It means that the Haar wavelet parameters are fitted up to level $j = 17$. All other differences are the consequence of the fact that in the finite trace generating method

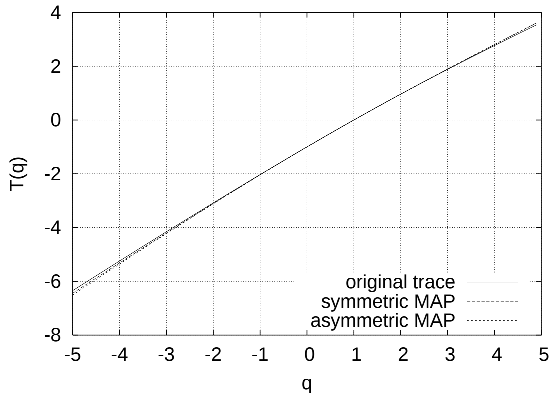


Figure 29: The partition functions of the original and the two kinds of MAP traces

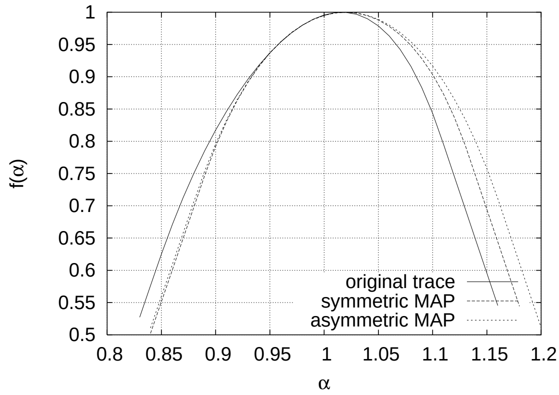


Figure 30: The Legendre transform of the partition functions in Figure 29

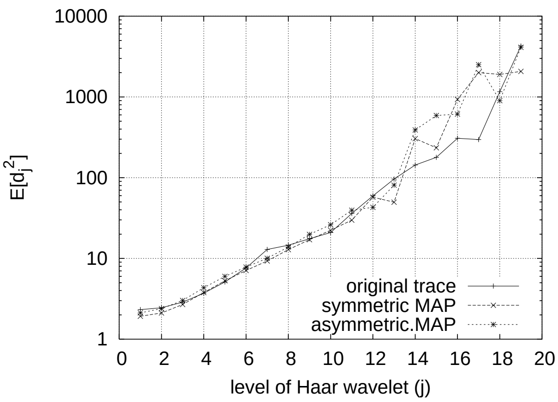


Figure 31: The Haar wavelet parameters of the original and the two kinds of MAP traces vs. j

there are finitely many samples of random variables. Based on Figure 31, we can say that the two MAP based simulation models approximate the Haar wavelet parameters reasonable well.

Queue length distribution The queue length distributions are calculated also for the two MAP based models with infinite buffer and deterministic service time.

The queue length distribution at utilization ($\rho = \frac{E[\text{service time}]}{\text{E[interarrival time]}}$) 0.2, 0.4, 0.6 and 0.8 are depicted in Figure 32.

In the case of MAP based simulations the synthetic traces closely approximate the original one even with low utilization level.

Marginal distribution During the analysis of the asymmetric MAP models with different sets of asymmetry parameters ($\{0.5, 0.5, 0.5, 0.5, 0.15\}$, $\{0.15, 0.5, 0.5, 0.5, 0.5\}$, $\{0.15, 0.15, 0.15, 0.15, 0.15\}$, etc) the difference between the marginal distribution of the original and the approximating traces remained large in all the cases. In those cases where the marginal distribution was somewhat closer to the original one the queue length dis-

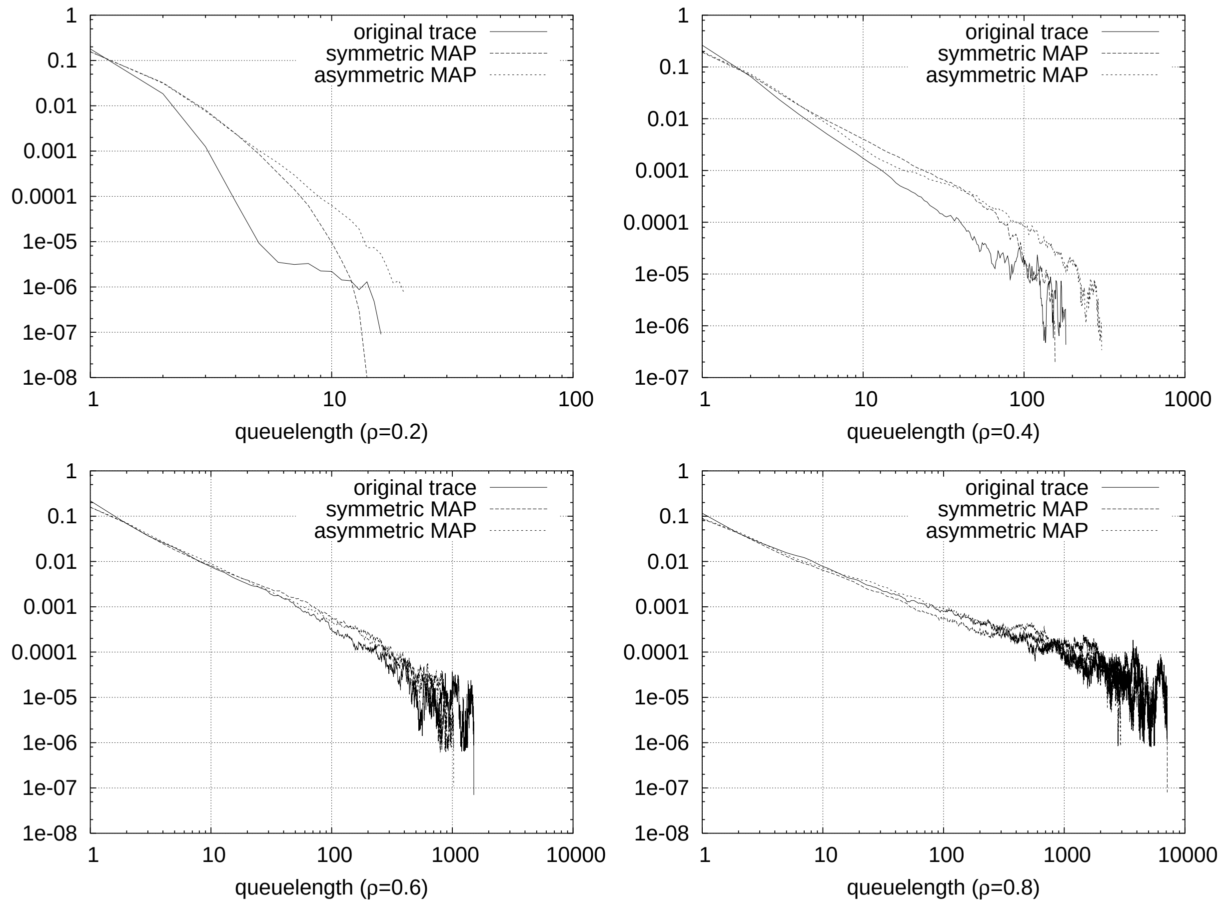


Figure 32: The probabilities vs. the queue length for the original and the MAP traces with several utilizations

tribution was rather far from the original (the approximate queue length distribution oscillates around the original one with larger amplitude). The results of one of the best asymmetric MMPP models are depicted in Figure 33 and 34.

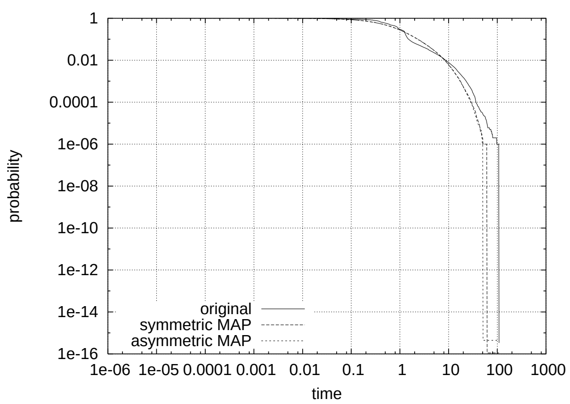


Figure 33: The complementary cumulative distribution functions of the three traces

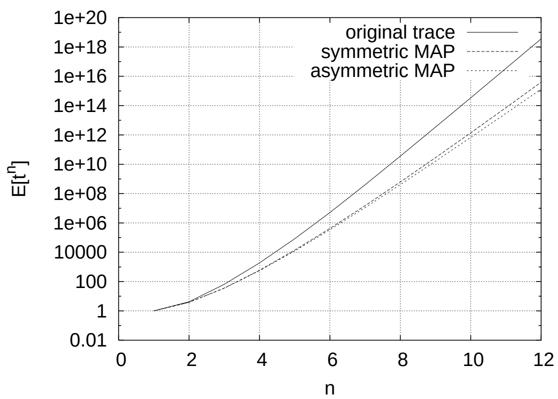


Figure 34: The first twelve moments of the three observed traces

Dependence on the initial seed value Similar to the Multifractal Wavelet Model the initial seed value of the random number generator has only a small effect on the properties of the trace generated by the MMPP simulation model.

6 Conclusion

We modified two long range dependent traffic models, the MWM and the MAP model, in order to approximate not only the long range dependent behaviour, but also the marginal distribution of measured traffic traces. We investigated their performance in approximating a commonly available real data trace of the internet archive [13] with respect to a set of statistical tests, which contains the most popular tests in time series analysis.

Our experiments indicate that both traffic models can exhibit long range dependent behaviour with mono- and multifractal scaling over several time scales.

The queueing behaviour is one of the most important properties in practice. Due to the long range dependent property of the original traffic trace, the experimental queue length distribution shows a polynomially decaying tail. Replacing the original trace with its modified MWM or MAP model the observed queue length distributions are reasonably close to the original one when the queue utilization was medium or high ($\rho \geq 0.5$). Both models preserve the polynomial queue length decay up to a given break point, which is very close to the limit statistical errors due to the finite length of the traces.

Based on the analysis of these two modified simulation models we can conclude that the accurate fitting of the queue length distribution and the marginal distribution are in conflict. The better is the fitting of the marginal distribution the worse is the associated approximate queue length distribution.

Acknowledgement

The author would like to acknowledge the help of András Horváth (Department of Computer Science, University of Torino) and Gábor Horváth (Department of Telecommunications, Technical University of Budapest). They provided their implementations of the MAP fitting and trace generator method and helped in theirs use. A special thanks to Miklós Telek for his instructions and help.

References

[1] J. Beran. *Statistics for long-memory processes*. Chapman and Hall, New York, 1994.

[2] L. Bodrog, G. Horváth, and M. Telek. Comparison of simulation models for long range dependent traffic traces. In *Proc. of The 3rd International Workshop on Rare Event Simulation and Combinatorial Optimization*, Budapest, Hungary, Sept 2004.

[3] H. J. Fowler and W. E. Leland. Local area network traffic characteristics, with implications for broadband network congestion management. *IEEE JSAC*, 9(7):1139–1149, 1991.

[4] R. Fox and M. S. Taqqu. 'large sample properties of parameter estimates for strongly dependent stationary time series. *The Annals of Statistics*, 14:517–532, 1986.

[5] A. Horváth and M. Telek. Markovian modeling of real data traffic: Heuristic phase type and map fitting of heavy tailed and fractal like samples. In *Tutorial of the IFIP WG 7.3 International Symposium on Computer Performance Modeling, Measurement and Evaluation (PERFORMANCE 2002)*, volume 2459 of *Lecture Notes in Computer Science*, Rome, Italy, Sept 2002.

[6] A Horváth and M. Telek. A markovian point process exhibiting multifractal behaviour and its application to traffic modeling. In *Proc. of MAM4*, Adelaide, Australia, 2002.

[7] W. E. Leland, M. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions in Networking*, 2:1–15, 1994.

[8] B. B. Mandelbrot and M. S. Taqqu. Robust R/S analysis of long-run serial correlation. In *Proc. of the 42nd Session of the International Statistical Institute*, volume 48, Book 2, pages 69–104, Manila, 1979. Bulletin of the I.S.I.

[9] I. Norros. A storage model with self-similar imput. *Queueing Systems*, 16:387–396, 1994.

[10] I. Norros. On the use of fractional brownian motion in the theorem of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13:953–962, 1995.

[11] R. H. Riedi. An introduction to multifractals. Technical report, Rice University, 1997. Available at `http://www.ece.rice.edu/~riedi`.

[12] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk. A multifractal wavelet model with application to network traffic. *IEEE Transactions on Information Theory*, 45:992–1018, April 1999.

[13] The internet traffic archive. http://ita.ee.lbl.gov/index.html.