Application of simulation models to measured traffic traces

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

2005.

Contents

1 Introduction **1**

2 Properties of traffic processes **3**
2.1 Self-similar processes 3
2.2 Multifractal processes 4

3 Statistical analysis methods **6**
3.1 Tests for long range dependency 6
3.1.1 Variance-time plot 6
3.1.2 R/S plot 7
3.1.3 Whittle estimator 8
3.2 Multifractal framework 8
3.2.1 Legendre spectrum 8
3.2.2 Haar wavelet 10
3.3 Stationary queue length behaviour 12
3.4 Transient queue length behaviour 13
3.5 Marginal distribution of the trace 14
3.6 Lag-k correlation 15

4 Simulation models **16**
4.1 Simulation models approximating long range dependency . . . 16
4.1.1 Multifractal wavelet model (MWM) [15] 16
4.1.2 Markov Arrival Process (MAP) based simulation . . . 19
4.1.3 Fractal signal models 24
4.2 PH renewal process 25
4.3 Particular simulation models 26
4.3.1 Poisson process 26
4.3.2 Randomly scrambled traces 27

5 Fitting the Bellcore trace with simulation models **30**
5.1 Application of the modified MWM 30

5.2 Application of the asymmetric MMPP 37
5.3 Application of the structurally modified MAP model 44
5.4 Application of the fractional Gaussian noise 50
5.5 PHFit 55

6 Conclusion **59**

A Application of simulation models to the DEC traffic trace **61**
A.1 Application of the MWM 61
A.2 Application of the MMPP model 65
A.3 Application of the fractional Gaussian noise model 69
A.4 PHFit 73

B Working environment **78**

Kivonat

A dolgozat célja, hogy megvizsgálja forgalmi folyamatok viselkedését, különös tekintettel a hosszútávú összefüggőségre és a multifraktális viselkedésre.

A megfigyelések alapját valós adatforgalmak, illetve a rájuk illesztett modellek által generált adatsorok képzik. Ennek megfelelően a szimulációs eljárásokat és a felhasználhatóságukat is vizsgáljuk.

A korábban megjelent szimulációs modellek [15, 5] jól illesztik a hosszútávú összefüggőséget és a multifraktális viselkedést, azonban a határeloszlást nem. A dolgozat célja úgy módosítani ezeket az eljárásokat, hogy ne csak a hosszútávú összefüggőséget, hanem a határeloszlást is jól illesszék. Másik célunk, hogy az egyik korábban felvázolt modell [5] segítségével illeszteni tudjuk a sorhossz-eloszlást kis kihasználtság esetén is.

Összehasonlításképp két másik szimulációs modellel is közelítjük a valós adatokat. Ezek a fraktális Gauss zaj generátor [15, 13] és a PH megújuló folyamat [5, 7].

Az eredeti és a módosított illesztő-eljárásokat számos statisztikai elemző eljárással, különböző szempontok szerint vizsgáljuk. Ezek a statisztikai vizsgálatok a Variancia-idő diagram, az R/S diagram, a Legendre spektrum, a Haar wavelet transzformáció, a sorhossz-eloszlás, a határeloszlás és a lag-k korreláció.

A dolgozat két valós adatsornak, a közelítésük lépéseinek illetve eredményének elemzését ismerteti mind az eredeti, mind a módosított illesztő eljárásokat vizsgálva. Az analízis során értékeljük az illesztési eljárás hatásait a szimulációs modellekre és ezt felhasználva optimalizáljuk az illesztést.

Abstract

The goal of this thesis is to examine the behaviour of traffic traces respecting the long range dependency and the multifractal behaviour.

The observations are based on the analysis of real traffic traces and simulation model generated traces that are fitted on the measurements. According to this the simulation methods and its usability is observed.

The simulation models presented in [15] and [5] are effective in fitting the long range dependent behaviour of measured traces, but they do not approximate the marginal distribution. One goal of this thesis is to modify those methods such that they capture not only the long range dependent behaviour, but also the marginal distribution. There is an aim to fit the queue length distribution at low utilization level with the traffic model introduced in [5].

For comparison two further simulation models are used to approximate real traffic behaviour. These are the fractional Gaussian noise generator referred in [15, 13] and the PH renewal process referred in [5, 7].

The original and the modified fitting methods are analysed from various point of view based on numerous statistical test the Variance-time plot, the R/S plot, the Legendre spectrum, the Haar wavelet transform, the queue length distribution, the marginal distribution and the lag-k correlation analysis.

The thesis reviews the statistical analysis of two measured data sets, the steps of approximating them with the 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 optimize the fitting procedure.

Chapter 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 [9], 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 thesis two long range dependent (LRD) traffic models were extended. Both of them is developed to capture not only the LRD, but also the marginal distribution and one of them to approximate the transient queue length behaviour of measured traffic traces. These two models are the Multifractal Wavelet Model (MWM) proposed in [15] and the Markov Arrival Process (MAP) model 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.

There are two other simulation models introduced to compare the results

of them and to analyse monofractals and non LRD traces. These are the fractional Gaussian noise (fGn) generator [13] and the Phase-Type renewal process based on PH fitting [7].

To approximate better the marginal distribution of real traffic traces 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. 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.

To approximate better the queue length distribution at low utilization several stochastic processes were analysed and finally the MAP model was changed. However the aim was to leave the statistical properties of the original simulation methods unchanged and according to the introduced modification they worsened.

The rest of the thesis is organised as follows. The considered traffic properties are shown in Chapter 2 and the applied statistical analysis methods are demonstrated in Chapter 3. Chapter 4 summarizes the main features of the considered traffic models and their application to the data sets. The following chapter, Chapter 5, provides the details of the fitting methods and the comparison of the main properties of the original "Bellcore data set" and its approximate traces. The thesis is concluded in Chapter 6. In Appendix A we describe the approximation of the TCP part of the "DEC trace". Finally in Appendix B we describe the working environment.

Chapter 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 or as in our observations X_i denotes the ith interarrival time. $\mathcal{X} = \{X_i, i = 0, 1, \ldots\}$ 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 and slowly varying means

$$\frac{A(t \cdot k)}{A(k)} \to 1, t > 0, k \to \infty.$$

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 behaviour, which is characterized by the Hurst parameter (H), can be checked based on any of the absolute moments of the aggregated process:

$$\log(E(|X^{(m)}|^q)) = \log(E(|m^{H-1}X|^q)) = q(H-1)\log(m) + \log(E(|X|^q)). \tag{2.1}$$

According to (2.1), 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 behaviour of all absolute moments, q, are the same and it is captured by the Hurst-parameter. If it is the case we talk about monofractal behaviour. The Variance-time plot, which is used widespread to gain evidence of self-similarity, is the special case with $q = 2$. It depicts the behaviour 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 Multifractal processes

Statistical tests of self-similarity try to gain evidence through examining the behaviour of the absolute moments $E(|X^{(m)}|^q)$. In case of monofractal processes the scaling behaviour 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 behaviour of $E(|X^{(m)}|^q)$ for different values q and results in a *spectrum* that illustrates the behaviour of the absolute moments. This analysis procedure is detailed in Section 3.2.

Chapter 3

Statistical analysis methods

3.1 Tests for long range dependency

Recently, it has been agreed [9, 11, 12] 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 behaviour 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 [16]. 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 3.1. The Hurst-parameter given by the Variance-time plot is 0.80.

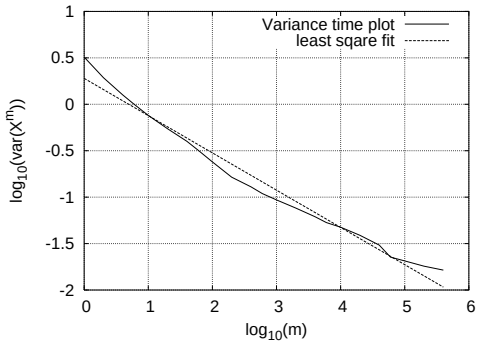


Figure 3.1: Variance-time plot and its least square fit for the Bellcore trace

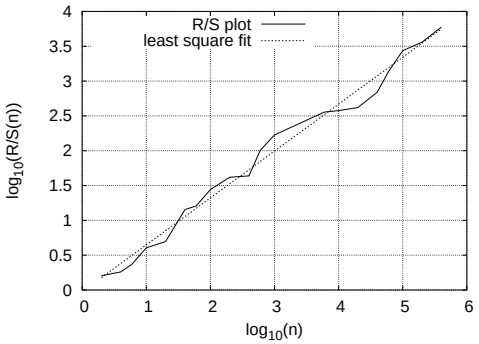


Figure 3.2: R/S plot and its least square fit for the Bellcore trace

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 [10]. 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 3.2. 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*

$$T(q) = \lim_{n\to\infty} \frac{1}{-n} \log_2 E\left[\sum_{k=0}^{2^n-1} |Y((k+1)2^{-n}) - Y(k2^{-n})|^q\right]. \tag{3.1}$$

Then, a multifractal spectrum, the so-called *Legendre spectrum* is given as the *Legendre transform* of (3.1)

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

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

$$T^*(\alpha) = q\alpha - T(q), \text{ and } (T^*)'(\alpha) = q \text{ at } \alpha = T'(q). \tag{3.2}$$

Let us mention here that there are also other kinds of fractal spectrum defined in the fractal world (see for example [14]). 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 (3.1) becomes

$$S_n(q) = \sum_{k=0}^{N/2^n-1} |X_k^{(2^n)}|^q, \ 0 \leq n \leq L, \tag{3.3}$$

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.

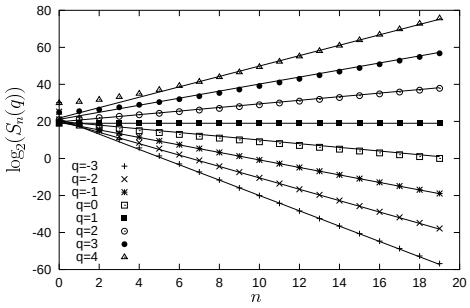


Figure 3.3: Scaling of log-moments with linear fits for the interarrival times of the Bellcore *pAug* trace

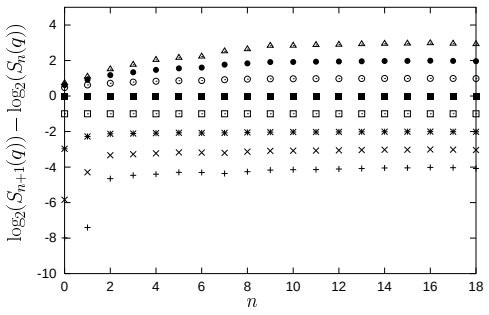


Figure 3.4: Increments of log-moments for the interarrival times of the Bellcore *pAug* trace

Figure 3.3, 3.4, 3.5 and 3.6 illustrate the above described procedure to obtain the Legendre spectrum of the Bellcore *pAug* traffic trace. Figure 3.3 depicts the scaling behaviour of the log- moments calculated by (3.3). 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 [15] 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 3.4 depicts the increments of the log-moment curves of Figure 3.3. Completely horizontal lines correspond linear log-moment curves.

The partition function $T(q)$ with various initial value of n is depicted in Figure 3.5. 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 is in the range $[0.8, 1.15]$), it is

not as informative as its Legendre transform is (Figure 3.6). 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 (3.2) 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 [15] for basic principles of interpreting the spectrum. We mention here only that a curve like the one depicted in Figure 3.6 reveals a *rich multifractal spectrum*.

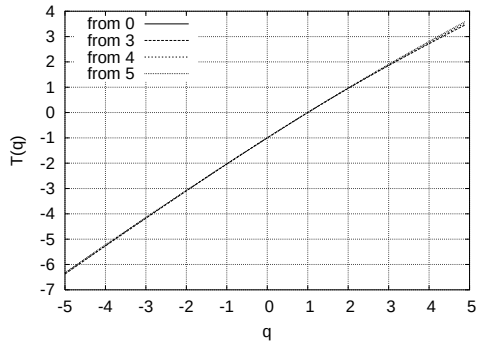


Figure 3.5: Partition function estimated through the linear fits shown in Figure 3.3, for different initial values of n

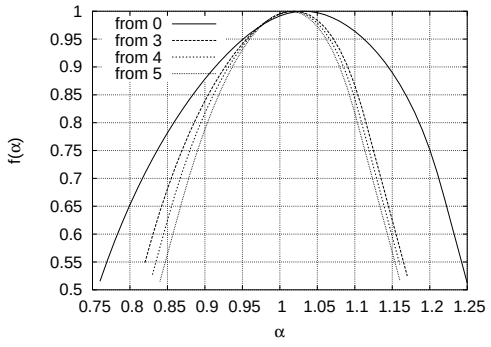


Figure 3.6: The Legendre transform of the partition functions shown in Figure 3.5 results in the Legendre spectra

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

The multiscale behaviour 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

$$c_{j,k} = c_{j-1,2k-1} + c_{j-1,2k}, \tag{3.4}$$

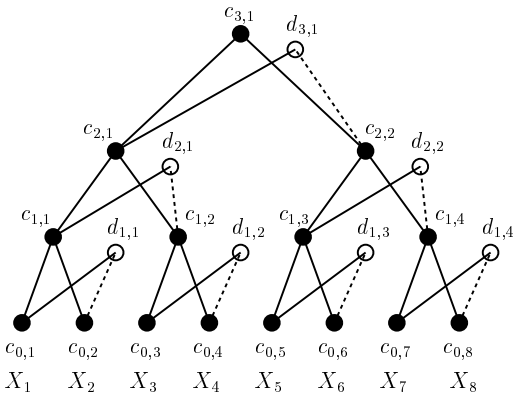


Figure 3.7: Haar wavelet transform

$$d_{j,k} = c_{j-1,2k-1} - c_{j-1,2k}, \tag{3.5}$$

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 is j the lower is 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 (3.4) and (3.5)

$$E[c_{j,k}^2] = E[c_{j-1,2k-1}^2] + 2E[c_{j-1,2k-1}c_{j-1,2k}] + E[c_{j-1,2k}^2], \tag{3.6}$$

$$E[d_{j,k}^2] = E[c_{j-1,2k-1}^2] - 2E[c_{j-1,2k-1}c_{j-1,2k}] + E[c_{j-1,2k}^2], \tag{3.7}$$

Let us assume that the series we analyze are stationary; then, by summing (3.6) and (3.7) and rearranging the equation, we have

$$E[c_{j-1}^2] = \frac{1}{4}\left(E[d_j^2] + E[c_j^2]\right). \tag{3.8}$$

Similarly, by consecutive application of (3.8) 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 4.1 we will discuss two methods that attempts to capture the multifractal scaling behaviour 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 3.8. We connect the $E[d_j^2]$ points only for better visualization.

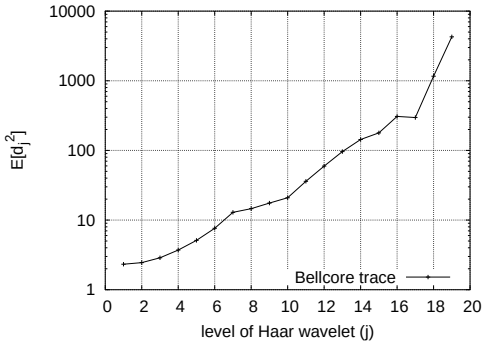


Figure 3.8: The Haar parameters $E[d_j^2]$ vs. j

3.3 Stationary queue length behaviour

One of our primary goals is to find a traffic model, by which the queue length behaviour 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 $\rho = 0.2$ and at $\rho = 0.8$ in Figure 3.9 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.

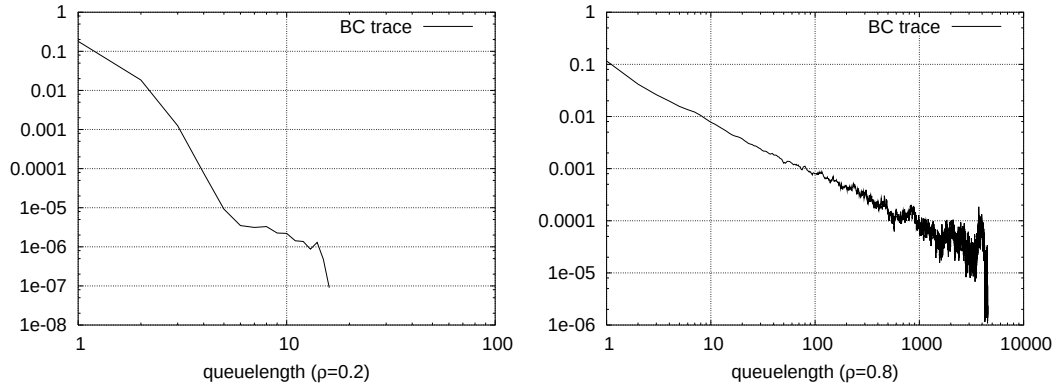


Figure 3.9: The probabilities vs. the queue length for BC trace with different utilizations

3.4 Transient queue length behaviour

As told above, in Section 3.3, one of the most important behaviour of a simulation model is to approximate well the queue length distribution of the original trace. Nevertheless another important parameter is the burstiness of the traffic on the network. This can be examined through the transient behaviour of the queue length.

This examination is based also on the assumption that a system with given service time is fed with a traffic trace, which can be either a measured data set or the result of a simulation.

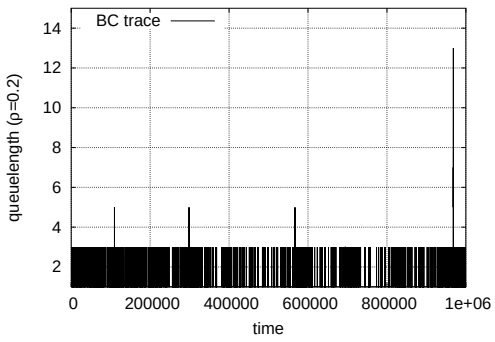


Figure 3.10: Transient queue of the BC trace at utilization $\rho = 0.2$

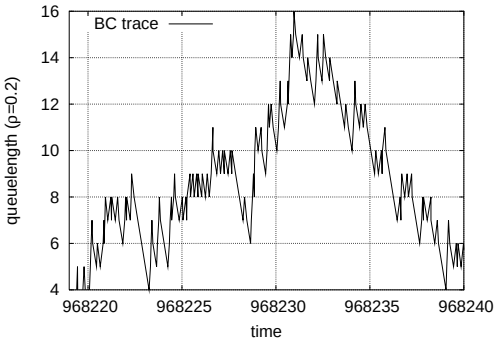


Figure 3.11: The burst in the BC trace at utilization $\rho = 0.2$

Figure 3.10 depicts the above described graph for the whole Bellcore data with $\rho = 0.2$. As one can see there is only one burst period, which makes the queue lengths larger than 5 (as shown in the left graph in Figure 3.9). This

interval is enlarged and depicted in Figure 3.11. The number of arrivals in the burst period was ~ 100.

3.5 Marginal distribution 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 divided 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 (M_i),
- normalize the value ($f_M(i)$) and plot it versus i.

The complementary cumulative distribution function (ccdf) is $1 - f_M(i)$. Figure 3.12 depicts the ccdf of the Bellcore trace.

Another way to compare two distributions is to calculate theirs moments. Two equivalent distributions have the same moments. We calculate the first twelve moments of the traces (Figure 3.13) and of its approximates. The empirical n^{th} moment of a data set can be calculated using this rule:

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

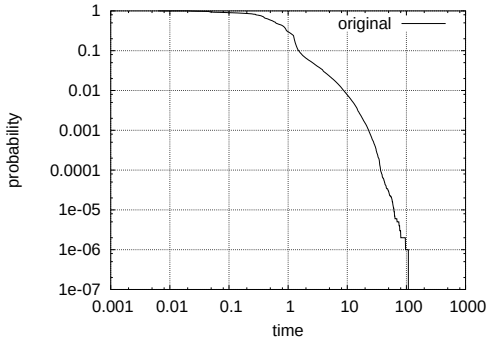


Figure 3.12: ccdf of the original Bellcore trace

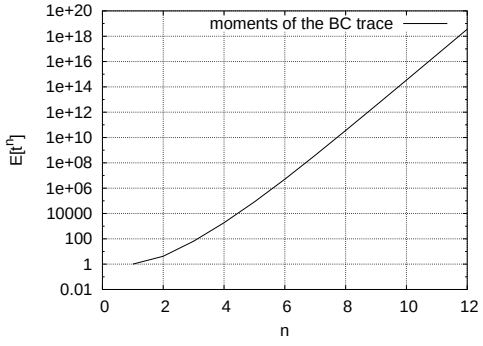


Figure 3.13: The first twelve moments of the Bellcore trace

3.6 Lag-k correlation

The lag-k correlation can be derived from the autocorrelation function of a random variable. Considering the interarrival times in the trace as a finite realization of a random variable and calculating the autocorrelation of the samples at a distance of k result the lag-k correlation. Doing this until a K value we get the first K lag correlation coefficients. The lag-0 correlation coefficient equals to 1 by its definition.

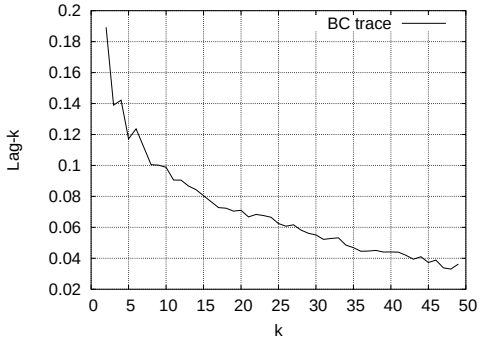


Figure 3.14: The lag-k correlation of the BC trace until $k = 50$

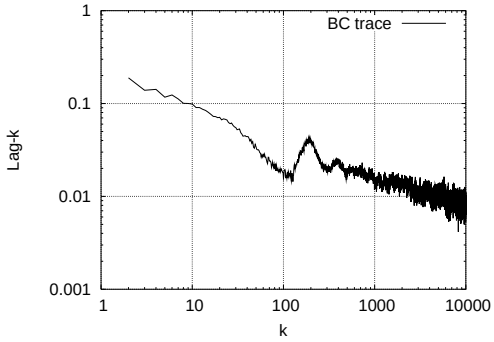


Figure 3.15: The lag-k correlation of the BC trace until $k = 10000$

Figure 3.14 depicts the lag-k coefficients of the Bellcore trace for the first 50 ks and Figure 3.15 for the first 10000. Although there is a break point in the graph at $k \simeq 20000$ as shown in [8] we have not calculated the lag-k coefficients for the higher k values because in our examinations we did not take care for the perfect fitting of this property and furthermore our observations have shown that it has less importance in the multifractal point of view.

Chapter 4

Simulation models

4.1 Simulation models approximating long range dependency

The simulation models introduced in this section are out to fit the long range dependent behaviour of measured traffic traces.

4.1.1 Multifractal wavelet model (MWM) [15]

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

$$d_{j,k} = A_{j,k} * c_{j,k}, \tag{4.1}$$

where the $A_{j,k}$ multipliers can be considered as 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 3.7. 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.

Application of MWM to measured data sets

As mentioned in [15] 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

$$\begin{aligned} \eta_j &= \frac{E[d_{j+1,k}^2]}{E[d_{j,k}^2]} \\ &= \frac{2E[A_{j+1}^2]E[c_{j+1,k}^2]}{E[A_j^2]E[1 + A_{j+1}^2]E[c_{j+1,k}^2]} \\ &= 2\frac{E[A_{j+1}^2]}{E[A_j^2](1 + E[A_{j+1}^2])}. \end{aligned} \tag{4.2}$$

To match the variance decay of the measured data set we can recursively solve (4.2) for $E[A_j^2]$ in terms of η_j and $E[A_{j+1}^2]$ for $j = L-1, \ldots, 0$. First we initialize the calculation at the coarsest scale by setting

$$E[A_L^2] = \frac{E[d_{L,1}^2]}{E[c_{L,1}^2]}. \tag{4.3}$$

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 (4.3).

2. At scale j, generate the random multipliers $A_{j,k}$ according to the chosen distribution function and calculate each $d_{j,k}$ via (4.1) 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 $j = 0$.

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 4.1). Both of them are easy to calculate.

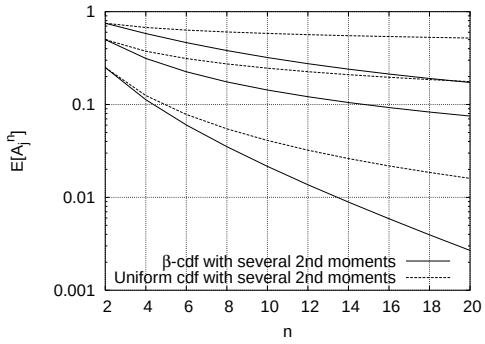


Figure 4.1: The even moments of the used distributions for the multipliers $A_{j,k}$ with 2nd moments $\{0.25; 0.5; 0.75\}$

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

$$g_A(x) = \frac{(1+x)^{p-1}(1-x)^{p-1}}{\mathcal{B}(p,p)2^{2p-1}}. \tag{4.4}$$

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:

$$var(A) = E[A^2] = \frac{1}{2p+1}. \tag{4.5}$$

The parameter p can be calculated from (4.5) 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 4.2 with several parameters p.

A special case of uniform distributions As written above in Section 4.1.1 the distribution function for multipliers have to be symmetric over $[-1, 1]$, with mean 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:

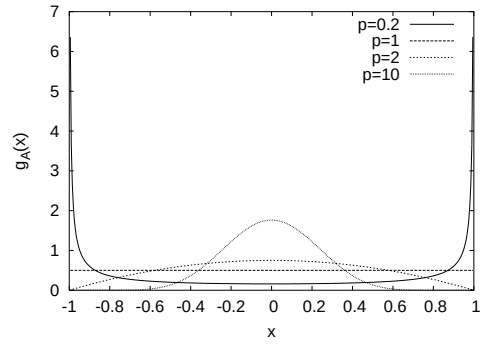


Figure 4.2: Beta distribution with several parameters

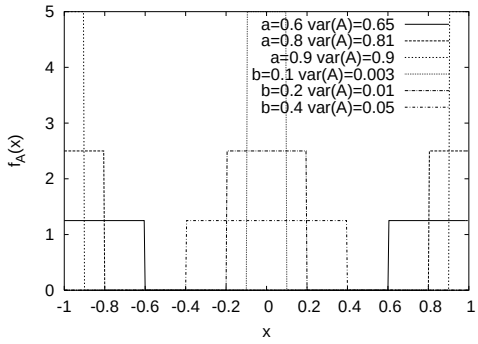


Figure 4.3: Uniform distribution with several parameters and variances

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

One of the results evaluating the quadratic equations for the parameters a and b matches the criteria $a, b \leq 1$.

The uniform distribution, with different values of the parameters a, b and the corresponding variances are depicted in Figure 4.3.

4.1.2 Markov Arrival Process (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 memoryless splitting.

Application of MMPP to measured data sets

In [6] a special MMPP structure is proposed to exhibit multifractal behaviour. 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. Without loss of generality, we assume that the time unit is such that the long term arrival intensity is one. An 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

$$\begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}$$

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 4.4) and generate the n-dimensional cube such that

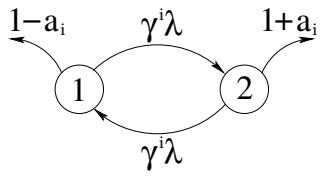


Figure 4.4: One dimension of the symmetric Markov chain

the behaviour 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

•	λ	$\gamma\lambda$	
λ	•		$\gamma\lambda$
$\gamma\lambda$		•	λ
	$\gamma\lambda$	λ	•

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 behaviour 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 [15], 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

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

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].

Asymmetric MMPP model

Based on the fitting method described in Section 4.1.2 with symmetric 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 are depicted in Figure 4.5.

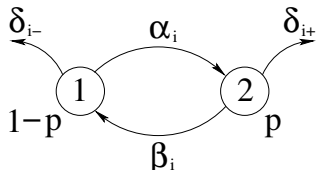


Figure 4.5: One dimension of the asymmetric Markov chain

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 new adjustable parameters.

Structurally modified MAP model

With the model proposed in Section 4.1.2 we get a sophisticated tool to fit the multifractal behaviour of a non-Markovian trace with a Markovian model, but we have to mind several practical considerations. The most important ones are those that react the queue length behaviour of the simulated trace. These are the approximation of the burstiness of a trace, i.e. the tail of the

queue length distribution and the better fitting of the body of the queue length distribution.

We have minded to approximate simultaneously the tail and the body of the queue length distribution especially at utilization level $\rho = 0.2$. Doing this we have made some heuristic modification in the original MMPP model.

First we have examined the effects of the random scrambling of the trace. $2^n, n \in [0, 19]$ length blocks of the interarrival times have been scrambled. The detailed description of this step and the experiences can be found below in Section 4.3.2 and 4.3.1.

After we have recognized the correspondence between the considered time scales in the MMPP model and the queue length distribution depending on the utilization level we tried to fit the lower time scales that affect the queue length distribution at lower utilization. In the term used in Section 4.1.2 we set n smaller than 5, γ smaller than 8 and T_M was set smaller too. With these settings the body of the queue length distribution was captured better than before, but the tail was failed.

To solve this problem we tried to reset the asymmetry parameter and the arrival intensity of the fastest dimension in the asymmetric structure described in Section 4.1.2, but it has a rough effect on the tail of the queue length distribution. Resetting this parameter changed the properties of the half of the states in the Markov chain thus the interarrival times in the trace.

We found that there is no appropriate setting of the asymmetry parameter and the arrival intensity to capture the tail of the queue length distribution. According to this a new state in the CTMC was introduced. This state is responsible for the burst period in the trace thus for the peak value in the transient queue length, i.e. the tail of the queue length distribution. The settings of the special state are the followings:

- The arrival intensity is set according to the highest queue length,
- the stationary probability of the new state is set to 10^{-6}, according to that the process moves probably only once in that state during the simulation,
- the transition rates to the new state are set to $\frac{1}{10^6-1}$ with respect to the setting of the cycle time between the old and new states to 10^6
- the transition rates from the new state are set according to the stationary probabilities of the states in the other part of the Markov chain.

Figure 4.6 demonstrates the structure of the new MAP model with three dimensions, the additive state and the arrival intensity of the new state.

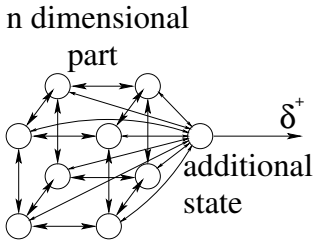


Figure 4.6: The structure of the new MMPP model

4.1.3 Fractal signal models

Fractals are geometric objects exhibiting an intricate, highly irregular appearance on all resolutions. Here, we are mainly interested in fractal signals, i.e. signals having fractal graph. Most known fractals are self-similar; if we "zoom" in or out of the fractal, we obtain a picture similar to the original [15].

Fractional Brownian motion

The fractional Brownian motion (fBm) $B(t)$ has been the most broadly applied fractal signal model [9]. Its power lies in its simplicity: fBm is statistically self-similar.

$$B(at) \stackrel{d}{=} a^H B(t).$$

Although it has rich statistical properties it remains amenable to a tractable analysis. The fBm is not stationary.

Fractional Gaussian noise

Although fBm is useful for theoretical analysis, its increments process (for finite increment (Δt))

$$G[n] := B(n\Delta t) - B((n-1)\Delta t),$$

known as fractional Gaussian noise, is often more useful in practice. While fBm is non-stationary, fGn is stationary.

The fGn with $1/2 < H < 1$ has proven useful for signal modeling. It is analytically LRD, theoretically its Variance-time plot is linear with a slope of $2 - 2H$.

Application of fGn to measured data sets

The fGn simulator has several adjustable parameters, which by the simulated trace can behave similar to the original one, these are:

- the Hurst parameter,
- the mean,
- the peakedness
- and the length of the trace.

4.2 PH renewal process

Fitting a phase type distribution to a measured traffic trace means that to find a phase type Markov chain and set its parameters e.g. the number of states, the time domain (continuous or discrete) etc, to approximate the marginal distribution function of the trace well.

Phase type distribution

$Z(t)$ is a continuous time Markov chain with n transient and one absorbing state. Its initial probability distribution is $\hat{\alpha}$ and generator matrix is $\hat{\mathbf{B}}$. The time to reach the absorbing state, T, is phase type distributed with representation $\alpha, \mathbf{B}$, where α is the sub-vector of $\hat{\alpha}$ and $\mathbf{B}$ is the sub-matrix of $\hat{\mathbf{B}}$ associated with the transient states. The cumulative distribution function (cdf), the probability density function (pdf) and the moments of this distribution are:

$$F_T(t) = 1 - \alpha e^{\mathbf{B}t} h, f_T(t) = \alpha \mathbf{B} e^{\mathbf{B}t} h, E[X^i] = i! \alpha (-\mathbf{B})^{-i} h,$$

where h is the column vector of ones. The number of unknown in the $\alpha, \mathbf{B}$ representation of a PH distribution is $\mathcal{O}(n^2)$ [5].

Application of PHFit to measured traffic traces

The detailed fitting method of phase type distribution to measured traffic traces is described in [7]. We applied this model to the measured dataset based on "engineering considerations" using the tool available at [17].

The parameters to fit the marginal distribution are:

- the time domain,
- the number of phases to fit either the body or the tail of the distribution function,
- measure to fit the body,

- the upper bound of fitting the body,
- the lower and the upper limit of fitting the tail
- and the distance of adjacent points.

4.3 Particular simulation models

In this section we describe the simulation trials and the scrambling of the trace, from which we drew the conclusions that the queue length distribution at different utilization levels relate to the fitted Haar wavelet levels and at very low utilizations the queue length distribution and the scaling behaviour are independent.

4.3.1 Poisson process

First we show the Haar parameters of a non-LRD trace. This was a Poisson process with parameter λ.

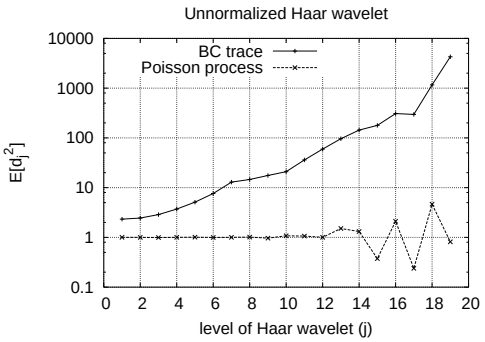


Figure 4.7: The Haar wavelet parameters of the Poisson process vs. j

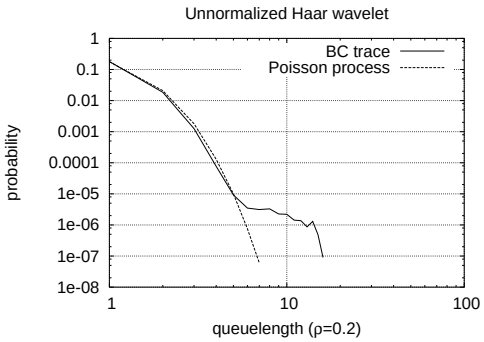


Figure 4.8: The queue length distribution of the Poisson process and the original trace

As one can see in Figure 4.7 the normalized Haar parameters of a non-LRD trace lie on a horizontal line. This graph shows the upper limit of reliable estimation of $E[d_j^2]$ based on 10^6 sample. This limit is the point, where the graph of the Poisson process starts to oscillate around the horizontal line.

Another astonishing conclusion was that the Poisson process approximates best the body of the queue length distribution (see Figure 4.8). This means that at low utilizations the long range dependent traces behaves like

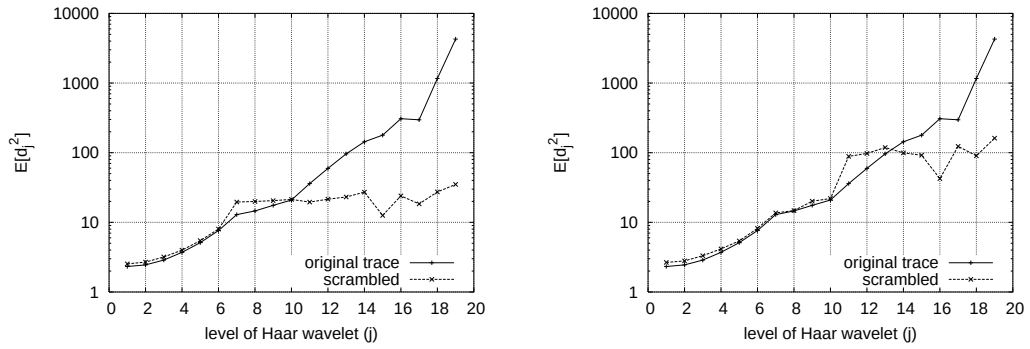
any non-LRD stochastic process this matches our imaginations that LRD dominates the systems with high utilization, when the queue of the system do not empties after almost every arrival.

4.3.2 Randomly scrambled traces

The random scrambling of the trace was organised as follows:

- the trace was divided into $2^n, n \in [0, 19]$ length blocks,
- the order of the blocks was permuted randomly.

Our expectation is that the scrambling of the 2^n length blocks results in a graph, in which the Haar parameters for j smaller or equal to n are the same as in the original, but for j greater than n the graph is a horizontal line. As one can see in Figure 4.9 the scrambled traces behaves like our expectations.



$n = 6$ $n = 10$

Figure 4.9: Scrambling of the blocks with different length of the BC trace

After this we observed the queue length distributions of the rearranged traces and experienced that for small ns, the queue length with low utilization level was unchanged and with high utilization the fitting was worsened, but for greater ns the distribution at all utilization levels was unchanged (see Figure 4.10). The conclusion was that the scrambling of the 2^n length blocks preserves the 2^n length dependency structure of the trace and eliminates the longer dependencies. This explains that the tail of the queue length distribution sometimes disappeared at utilization level 0.2.

The burst period of the BC trace at utilization level $\rho = 0.2$ contains ~ 100 arrival, this is depicted in Figure 3.11. When the blocks are much more smaller than 100 the tail disappeared during the scrambling. When the length of the blocks are approximately 100 the tail was shorter than

the original. When the blocks was much more longer than the burst period $2^n \gg 100$ the tail was unchanged, i.e. the 2^n length dependency remained. This is depicted in the upper left graph in Figure 4.10.

These observations showed us that the lower Haar parameters, i.e. the smaller time scales are responsible for the queue length distribution at lower utilization levels and the fitting of higher Haar coefficients are indifferent.

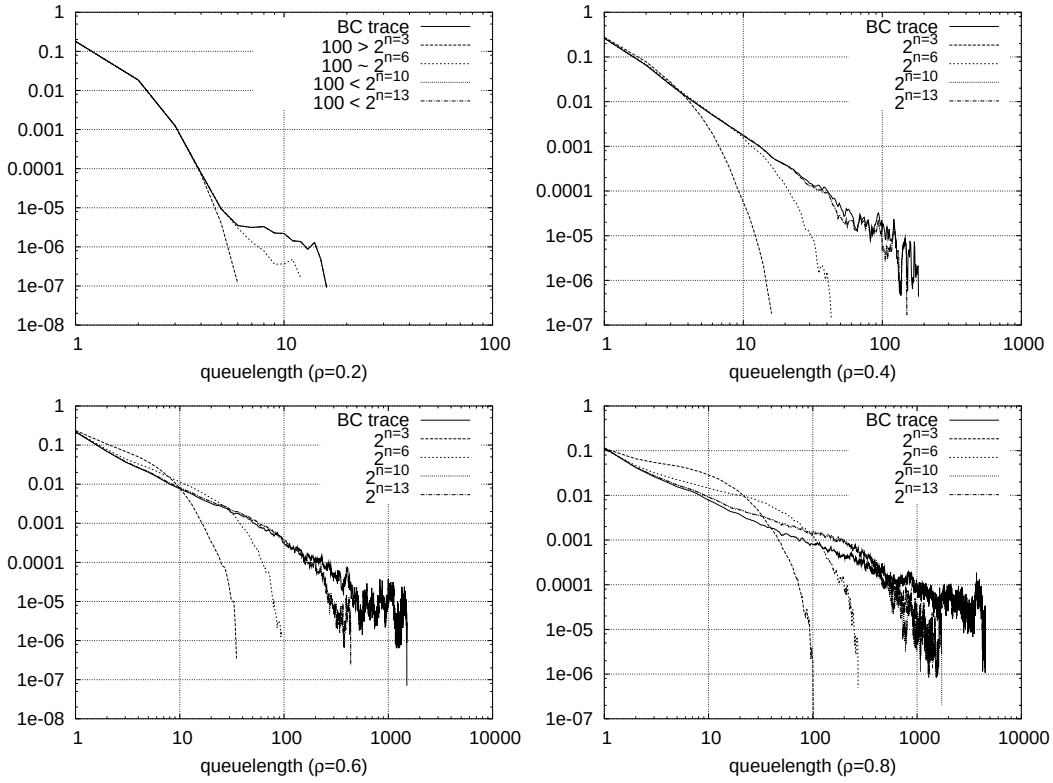


Figure 4.10: The probabilities vs. the queue length for the scrambled traces and the original with several utilizations

The conclusion that at very low utilizations the queue length distribution and the fitting of the Haar parameters, i.e. the scaling behaviour are independent is drawn from the examination of the Haar parameters. The comparison is showed in Figure 4.11 and 4.12.

Although the Poisson process and the randomly scrambled trace captured the body of the queue length distribution best, we have chosen the MAP model with three states $\gamma = 2$ and $T_M = \frac{1}{\lambda_0}$ (where λ_0 is the average arrival rate of the original trace) to be the basics for the new structure referred as "structurally modified MAP model" in Section 4.1.2. This is because the fitting of the queue length behaviour was closer to the original and the fitting

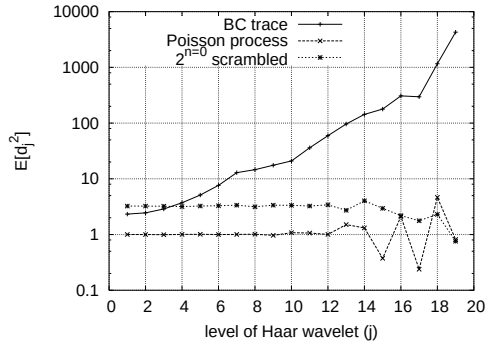


Figure 4.11: Comparison of the Haar parameters of the original trace, the Poisson process and the $n = 0$ scrambled trace

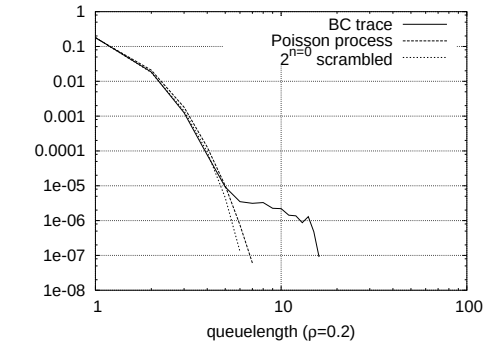


Figure 4.12: Comparison of the queue length distribution of the original trace, the Poisson porocess and the $n = 0$ scrambled trace

of the Haar parameters "was not worse" than the two traces described in this section, but it has an advantage, we can add a new state to the background Markov chain, responsible for the burstiness of the simulated trace.

Chapter 5

Fitting the Bellcore trace with simulation models

In this section we present the applied fitting methods and the differences between the synthetic and the original traces. For all models we test the goodness of the approximation of the results with the corresponding tests introduced in Chapter 3.

5.1 Application of the modified MWM

Since we already computed the statistical measures of the Bellcore trace (in Chapter 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 the $A_{j,k}$ multipliers. The MWM with beta distribution was applied and analysed in [15, 2] and found that approximates well the long range dependency and the multifractal behaviour. With the modification of the MWM our goal was to keep all these properties and in addition to capture the marginal distribution of the original trace.

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. 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 5.1.

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.

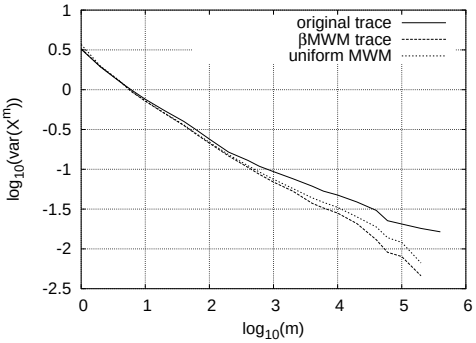


Figure 5.1: The Variance-time plot for the original and the different MWM traces

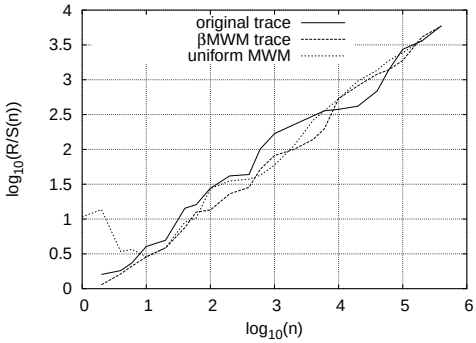


Figure 5.2: The R/S plot for the original and the different MWM traces

R/S plot

As one can see in the Figure 5.2 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 Spectra 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 5.3 and 5.4.

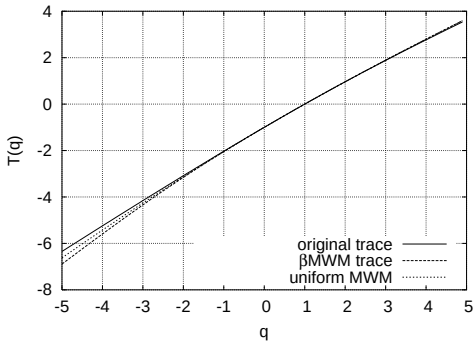


Figure 5.3: The partition functions of the original and the two kinds of MWM traces

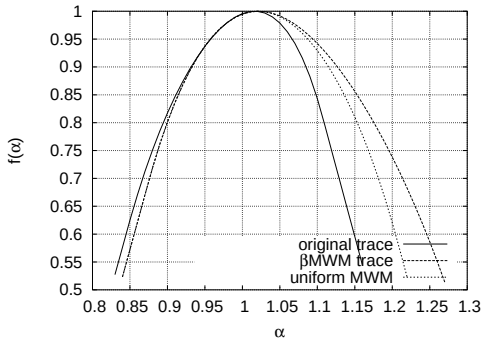


Figure 5.4: The Legendre transform of the partition functions in Figure 5.3

As one can see the Legendre spectra 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 we examine only the *richness* of the *multifractal spectrum* and all the three curves depicted in Figure 5.4 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 5.5 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 4 values when $j = 18$.)

The $E[d_j^2]$ parameters of the original and the approximating traces 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.

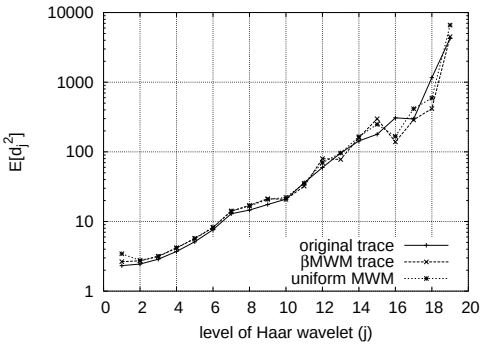


Figure 5.5: The Haar wavelet parameters of the original and the two kinds of MWM traces vs. j

Stationary 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.

We have depicted the queue length distribution of the system with four values of ρ in Figure 5.6.

Taking a look at Figure 5.6 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.

Transient queue length behaviour

This test is used to examine the queue length behaviour in the time domain, i.e. the burstiness of the trace, thus the diagrams depicts the queue length

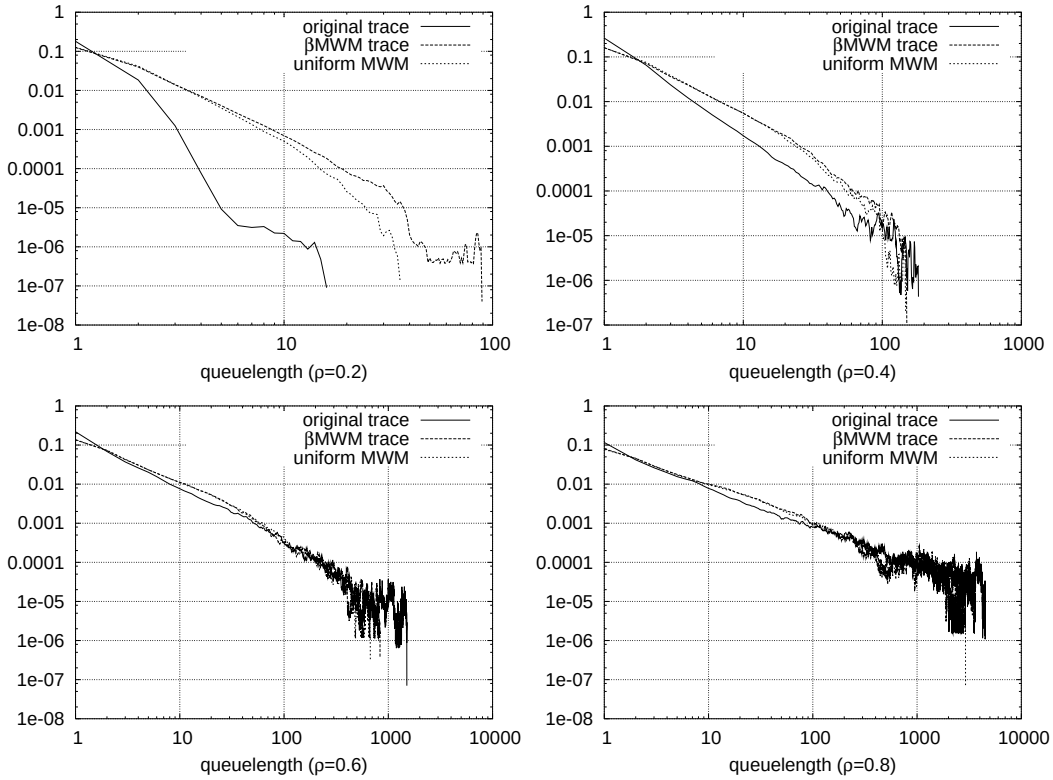


Figure 5.6: The probabilities vs. the queue length for the original and the MWM traces with several utilizations

versus time. To perform the test we assumed the same deterministic service time.

As one can see in Figure 5.7 the original trace has only one burst period near $t = 1000000$ as told in Section 3.4, however the approximating traces have several peaks larger than the one in the graph of the original BC trace. As a consequence we can say that these two models, the original and the modified Multifractal Wavelet Model, is not good enough from this point of view.

Although we have jumped to this conclusion we could not done anything to capture better the temporary queue length behavior of the original trace with the MWM.

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

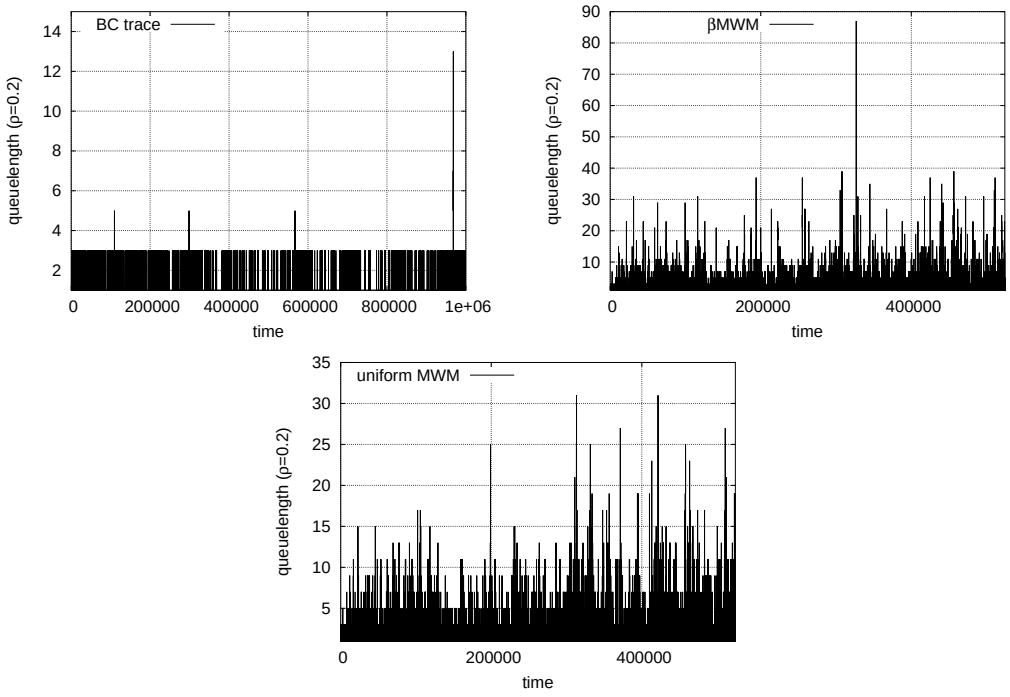


Figure 5.7: The temporary queue length of the BC trace and of its two MWM approximates

ccdf and moments are calculated as it is introduced in Section 3.5.

These properties of the Bellcore trace and of its Multifractal Wavelet Model approximations can be seen in Figure 5.8 and 5.9.

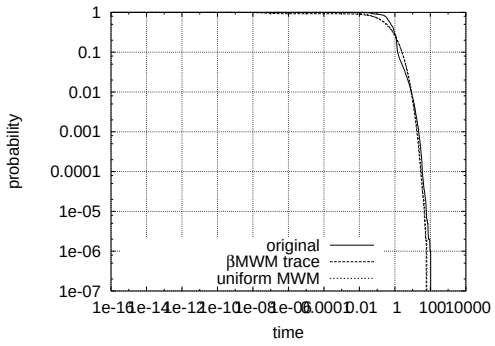


Figure 5.8: The complementary cumulative distribution functions of the three traces

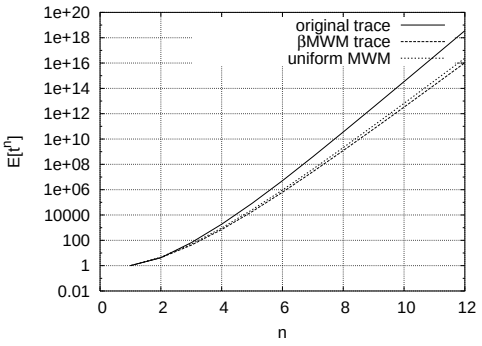


Figure 5.9: The first twelve moments of the three observed traces

Based on Figure 5.8 the marginal distributions looks very similar, but

the differences are more visible from the series of moments on Figure 5.9 (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 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 (see Figure 4.1).

Lag-k correlation

As told in Section 3.6 the approximation of the lag coefficinets of the trace were not in our primary goals, but for curiosity and for comparison we have observed our models with this test too.

The results are not bad. Both the original and the modified MWM approximates the original trace well for the first few ks apart from the oscillation (see Figure 5.10), but for larger ks the tendency declines as shown in Figure 5.11.

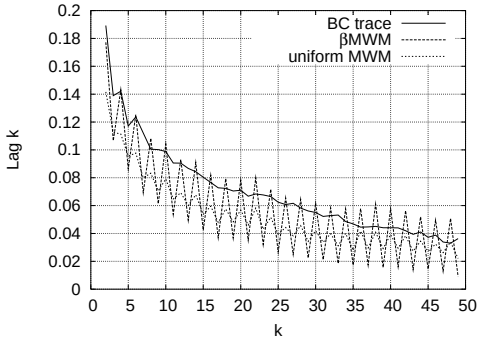


Figure 5.10: The lag-k correlation of the BC trace and of its two MWM approximations until $k = 50$

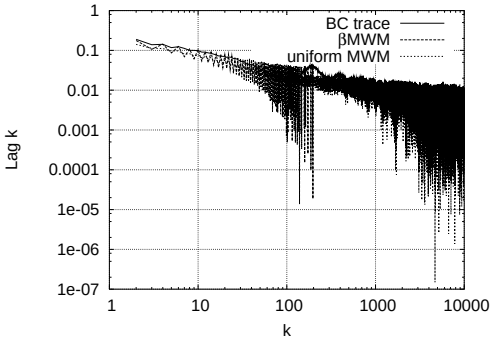


Figure 5.11: The lag-k correlation of the BC trace and of its two MWM approximations until $k = 10000$

There is higher correlation for even lag-k correlation in Figure 5.10 and 5.11. The oscillation is a result of the MWM simulation method. The correlation of two samples, X_n and X_{k+n}, when k is even is characterized by one Haar parameter more than the correlation of X_n and X_{k+n-1}.

Dependence on the initial seed value

In [15] 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 5.12 and 5.13.

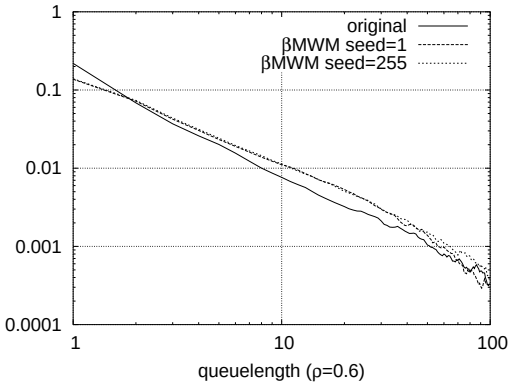


Figure 5.12: The probabilities vs. the queue length with different seeds ($\rho = 0.6$)

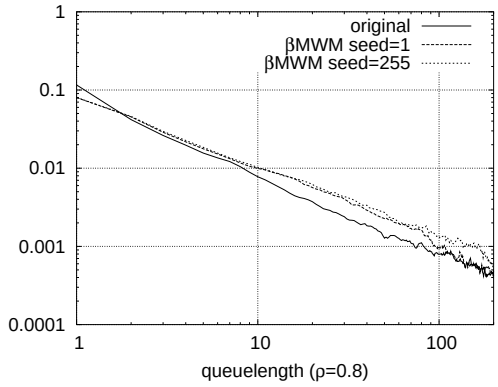


Figure 5.13: The probabilities vs. the queue length with different seeds ($\rho = 0.8$)

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 4.1.2. 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 4.1.2 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 4.1.2 we introduced a level independent asymmetry parameters setting, i.e. $p_i = p, \forall i$, but 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 get longer. 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 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 significantly from the original trace and from the trace generated by the symmetric MMPP.

There were inserted the distance of the first five moments of the original and the simulated traces in the modified set of parameters, but the optimization failed. The complexity and the accuracy of the optimization declined.

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 5.14 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 plot 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.

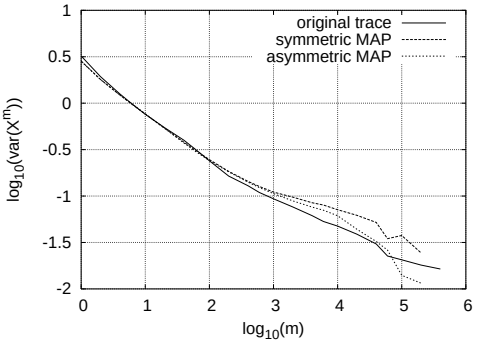


Figure 5.14: The Variance-time plot for the original and for the two different MMPP traces

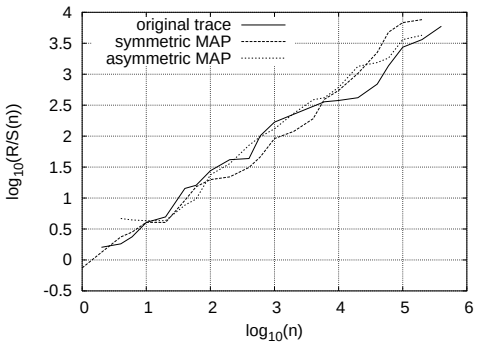


Figure 5.15: The R/S plot for the original and for the two different MMPP traces

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 5.15. 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. 5.16) and the Legendre spectra (Fig. 5.17) are very close to each other and to the original trace too. All three traces shows *rich multifractal spectrum.*

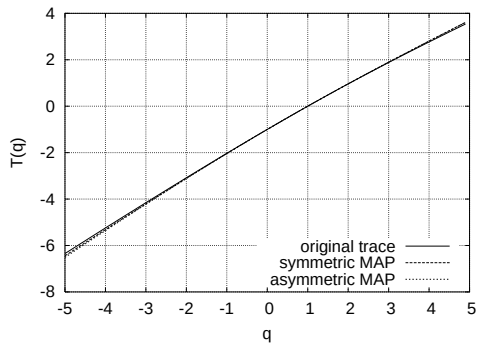


Figure 5.16: The partition functions of the original and the two kinds of MAP traces

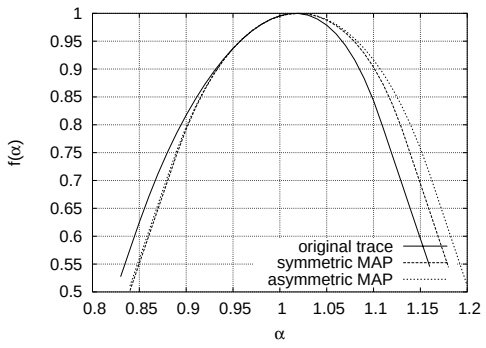


Figure 5.17: The Legendre transform of the partition functions in Figure 5.16

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 5.18.

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 there are finitely many samples of random variables. Based on Figure 5.18, we can say

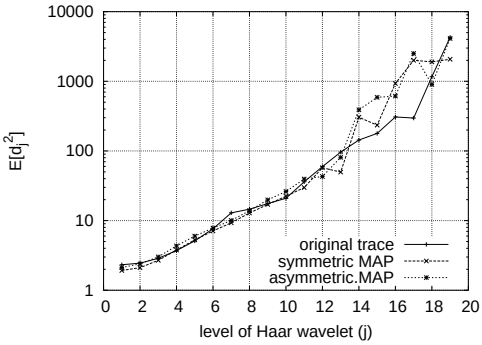


Figure 5.18: The Haar wavelet parameters of the original and the two kinds of MAP traces vs. j

that the two MAP based simulation models approximate the Haar wavelet parameters reasonable well.

Stationary 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 0.2, 0.4, 0.6 and 0.8 are depicted in Figure 5.19.

In the case of MAP based simulations the synthetic traces closely approximate the original one even with low utilization level. This is the reason why we have chosen this model as a base for our further researches on the queue length behaviour of the simulation models.

Transient queue length behaviour

Similar to the observation in Section 5.1 we have observed the burstiness of the original MAP and the asymmetric MAP model in comparison with the original BC trace.

As shown in Figure 5.20 the graph of the original BC trace is the same as in Figure 3.10 and in Figure 5.7. The MAP approximates differs less from the original one than the MWM.

For the symmetric MAP the value of the peaks are approximately the same as in the case of the original one, but there are much more of them. According to this the body of the queue length distribution for this model is larger than for the original BC trace as shown in Figure 5.19 at utilization level $\rho = 0.2$.

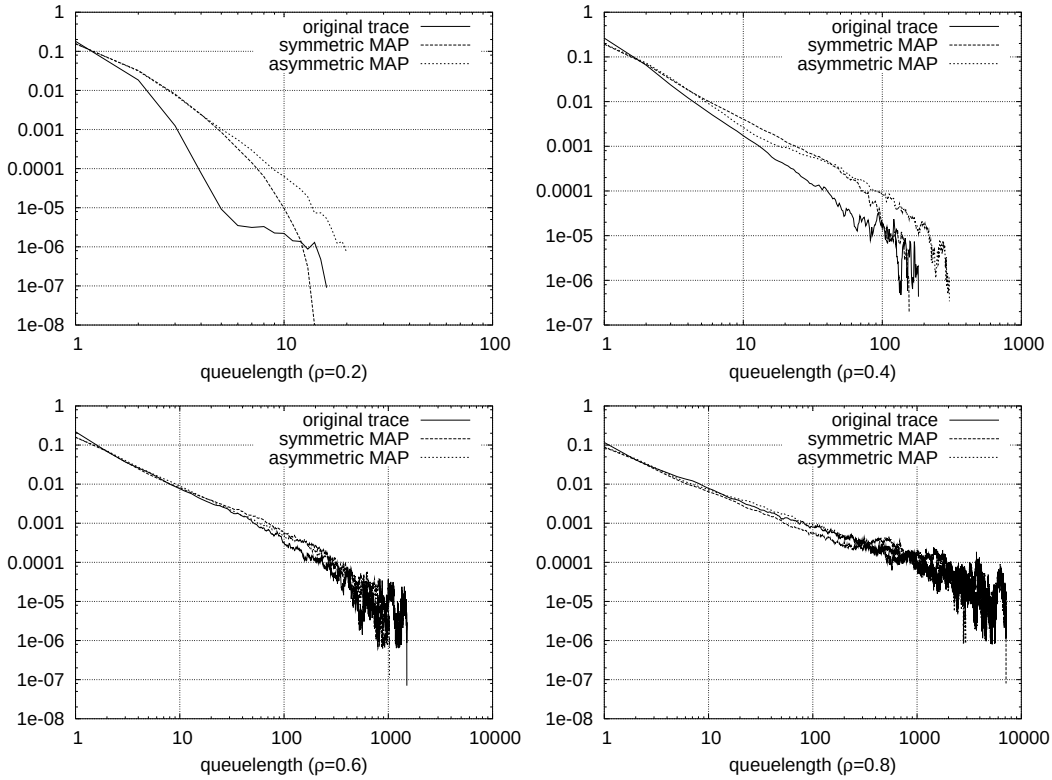


Figure 5.19: The probabilities vs. the queue length for the original and the MAP traces with several utilizations

For the asymmetric MAP model both the number of peaks and the value of them are larger. This results in a queue length distribution with longer and larger body than the original (see Figure 5.19 at $\rho = 0.2$).

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 distribution 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 with asymmetry parameter setting $\{0.5, 0.5, 0.15, 0.5, 0.5\}$ are depicted in Figure 5.21 and 5.22.

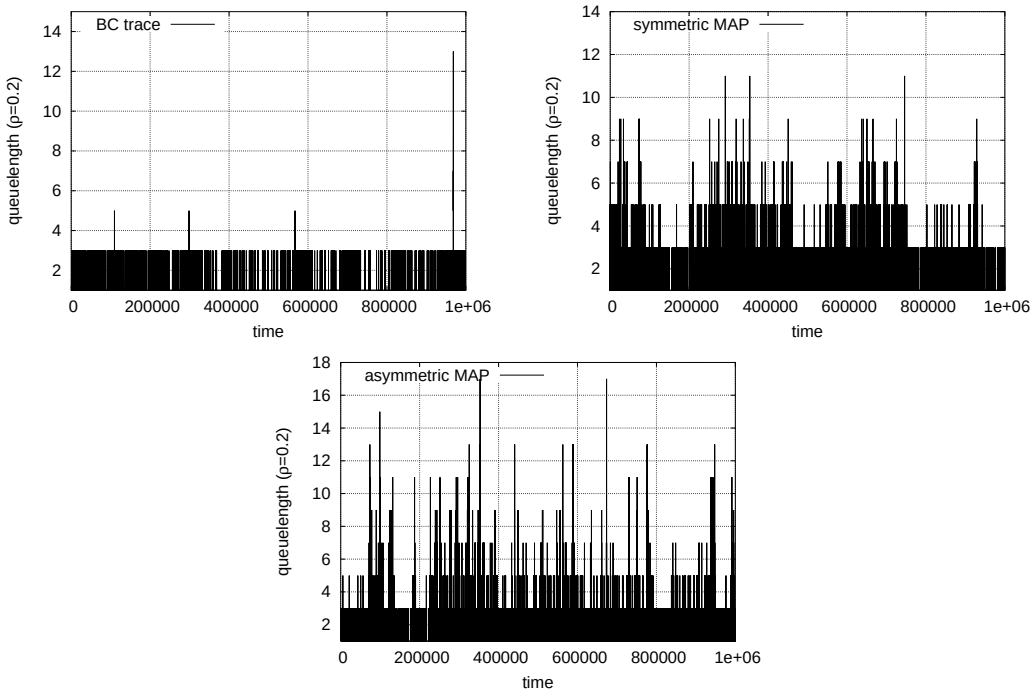


Figure 5.20: The temporary queue length of the BC trace and of its two MAP approximates

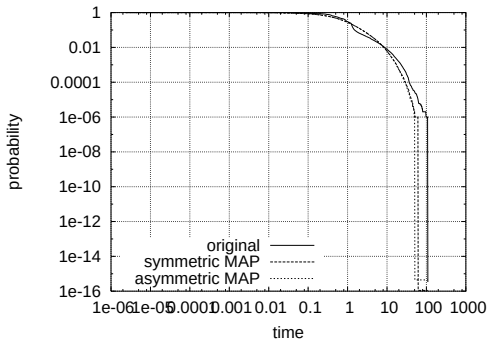


Figure 5.21: The complementary cumulative distribution functions of the three traces

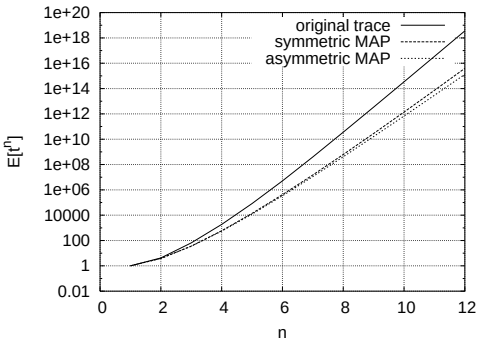


Figure 5.22: The first twelve moments of the three observed traces

Lag-k correlation

In the case of the MAP simulation models we have also compared the lag-k correlation of the models with the original trace however we have not fitted the lag coefficients.

The results was fairly good. The lag coefficients of the simulated traces are close to the originals for the more important small ks (see Figure 5.23) and also for larger ks (see Figure 5.24).

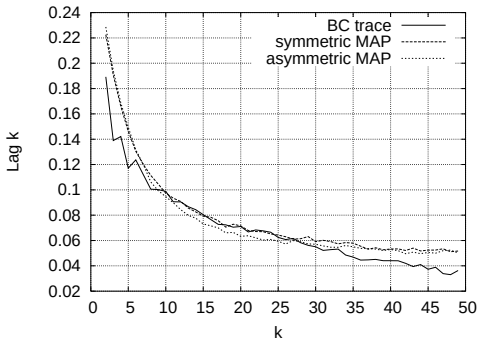


Figure 5.23: The lag-k correlation of the BC trace and of its two MAP approximations until $k = 50$

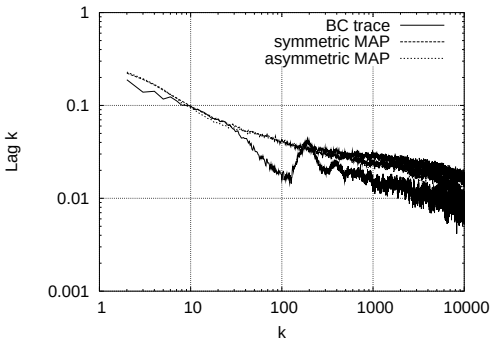


Figure 5.24: The lag-k correlation of the BC trace and of its two MAP approximations until $k = 10000$

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.

5.3 Application of the structurally modified MAP model

The structurally modified MAP is a heuristic extension of the MAP model discussed in Section 4.1.2. This means that during the fitting method we tried to find a solution which fit almost perfectly both the tail and the body of the queue length distribution at low utilization level.

The considered properties were those reacting the queue length distribution the fitted time scales and the burstiness of the trace. After the first attempts to set the asymmetry parameters in the MAP model, we modified the structure. A new state was added to the existing ones. This state is responsible for the burst period in the temporary queue length, i.e. for the tail of the queue length distribution. According to these assumptions the setting responsible for the significant properties was the following:

- the arrival intensity generated by the new state was set to $\delta^+ = N * v$, where
- $v = 1$ is the transition rate from the new state to the "old" part of the Markov chain and
- $N = 22$ is the highest queue length wanted.
- The stationary probability of the new state is set to $p = 10^{-6}$ (the process moves most probably only once in this state during the simulation run),
- the transition rates to the new state are $\frac{1}{10^6-1}$ to the set the cycle time between the "old" and new states to 10^6.

Finally, according to the conclusions in Sections 4.3.2 and 4.3.1 the base of the new structure was the MMPP model with $n = 3$. These conclusions were that the body of the queue length distribution of the trace behaves like any random variable at low utilization levels and the correspondence between the considered time scales and the queue length distribution at different utilization levels. This is why the setting of the MMPP model was

- $n = 3$,
- $\gamma = 2$,
- $T_M = \frac{1}{\lambda_0}$, where λ_0 is the mean of the trace.

Variance-time plot

In the case of the structurally modified MAP we dispensed with the fact that at high aggregation levels the plot differed very much from the theoretical straight line. We assumed that this is caused by the error of the variance of few samples.

Apart from this the empirical Hurst parameter was 0.52 based on the Variance-time plot and this is confirmed in Figure 5.25. Since $0.5 < H < 1$ means that the trace is *pseudo* self-similar the structurally modified MAP is "almost non LRD".

R/S plot

Based on the R/S plot we can say that the trace is a border-line case from *pseudo* self-similarity point of view. We have plotted the logarithm of the $R/S(n)$ of the trace, fitted a line on it and the slope of it, i.e. the Hurst parameter was rounded up to $H = 0.5$. This is depicted in Figure 5.26.

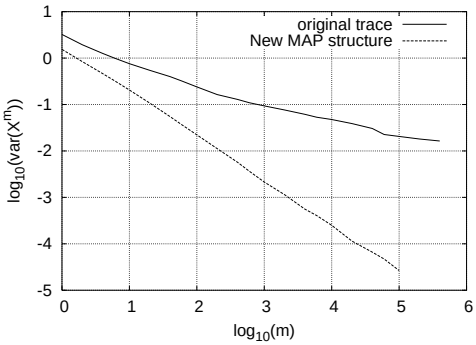


Figure 5.25: The Variance-time plot for the original and for the new MAP simulated trace

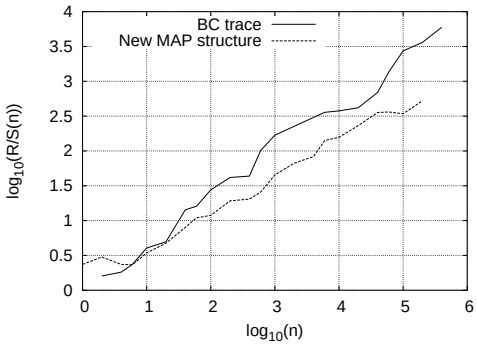


Figure 5.26: The R/S plot for the original and for the new MAP simulated trace

Haar wavelet

In the case of the structurally modified MAP model the multifractal behaviour is analysed only with the Haar wavelet transform.

At first glance one can say that the fitting of the Haar parameters is poor based on Figure 5.27, but taking into consideration the parameters of the background Markov chain the opinion is changed.

As told in Section 4.1.2 the lower Haar parameters "are responsible" for the fitting of the smaller time scales and for small js ($j \in [1, 3]$) the lines are fairly close to each other in Figure 5.27.

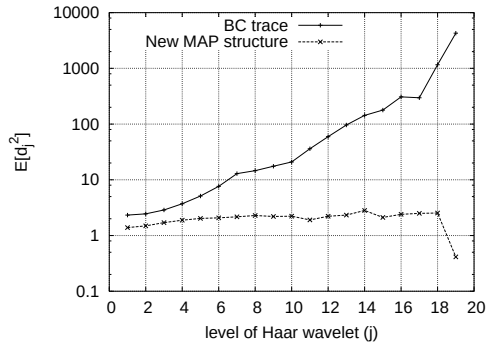


Figure 5.27: The Haar wavelet parameters of the original and the new MAP simulated trace vs. j

Stationary queue length behaviour

Although the modification of the MAP model was performed based on the queue length distribution at low utilization level, the observation at high utilizations is instructive. This is why the trace is analysed at all usual utilizations.

In Figure 5.28 at $\rho = 0.2$ one can see that the fitting was successful. Both the body and the tail of the queue length distribution was captured.

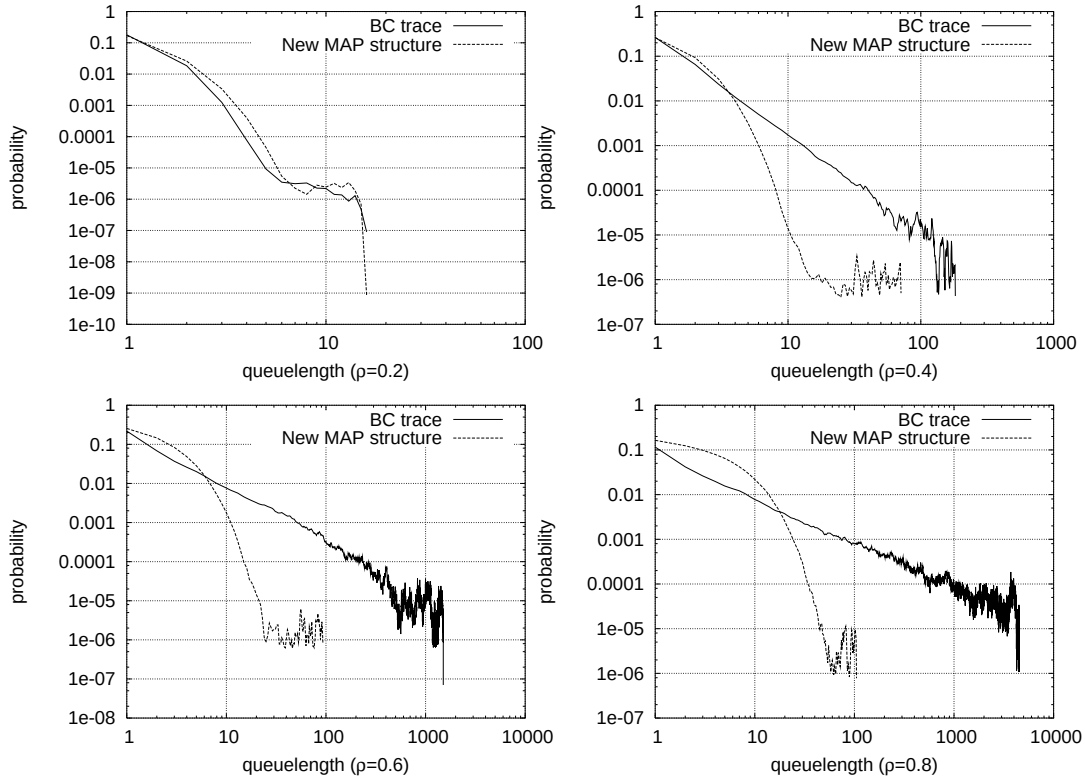


Figure 5.28: The probabilities vs. the queue length for the original and the new MAP trace with several utilizations

At other utilization levels the conclusion is the same as in Section 4.1.2. We can observe it again that the smaller time scales are responsible for the queue length at low utilization and at high utilizations the system fed with this trace results in a shorter queue length distribution. As one can see the tail, caused by the burst period, is getting shorter with the growth of the utilization. This is because the new state generated only a short time burst period which resulted in the tail of the graph at $\rho = 0.2$.

Transient queue length behaviour

After the previous analysis of the trace generated by the new MAP model our expectation is that the diagram illustrating the temporary queue length of the simulation results in a graph with one peak.

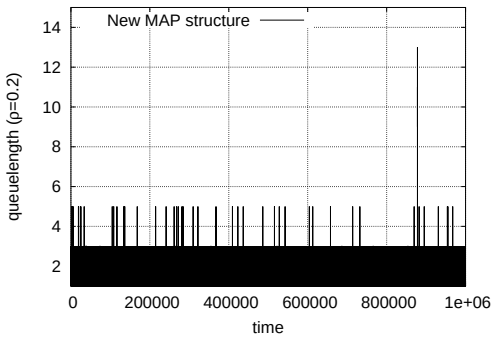


Figure 5.29: Transient queue of the new MAP simulated trace at utilization $\rho = 0.2$

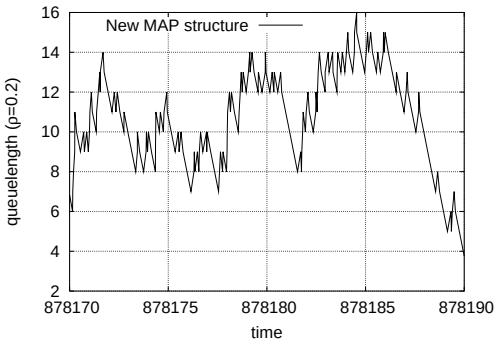


Figure 5.30: The burst in the same trace at utilization $\rho = 0.2$

Observing the diagram in Figure 5.29 it meets our expectation. There is only one peak with the required height. This is enlarged in Figure 5.30. This burst is responsible for the tail in Figure 5.28 at utilization 0.2.

Marginal distribution

After the attempts to capture the marginal distribution of the original trace presented in Sections 4.1.2 and 4.1.1 we can say that our modification on the original simulation model was not proper to perform a good approximation. Although we make the comparison of the original trace and the model to make a wide variety of observations on it.

As one can see in Figure 5.31 the tendency of the two traces are the same, but for all random variables the ccdf have almost the same look hence we have depicted the moments of the trace. This property shows the significant differences between the two data sets, i.e. the marginal distribution of the original trace was not captured in this case too.

Lag-k correlation

As told in Section 3.6 the lag-k correlation is derived from the autocorrelation function of the sample series. According to this and that the scrambling of the blocks of the trace eliminates the dependencies longer than the block,

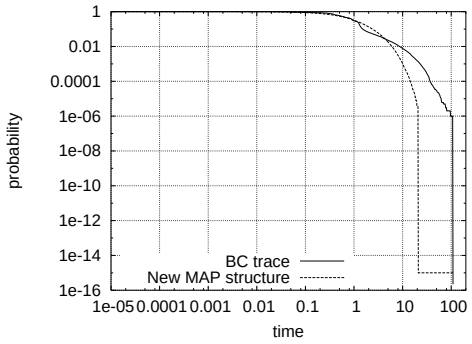


Figure 5.31: The complementary cumulative distribution functions of the two traces

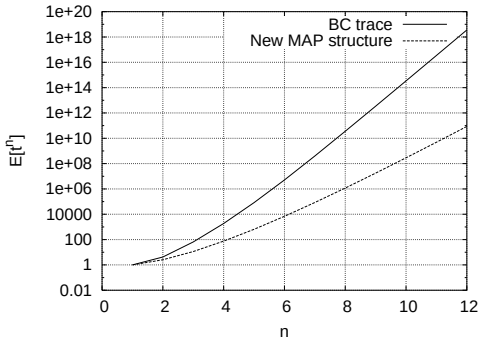


Figure 5.32: The first twelve moments of the two observed traces

the looks of the Lag-k in Figure 5.33 is evident. A trace generated by a structure with three dimensions ($n = 3$) responsible for the smallest three time scales ($\gamma = 2, T_M = 1/\lambda_0$), results in a flat lag-k correlation close to 0 except the first few ks. This is the reason of the looks in Figure 5.34, where the logarithm of the values near 0 and smaller than 0 were calculated. The omitted points in the graph are the logarithm of the negative values.

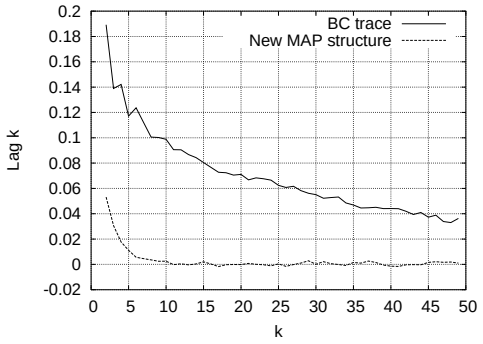


Figure 5.33: The lag-k correlation of the BC trace and of the new MAP approximation until $k = 50$

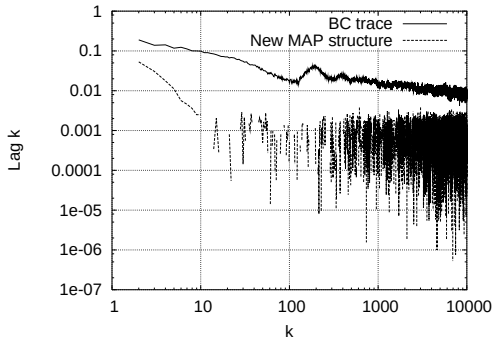


Figure 5.34: The lag-k correlation of the BC trace and of new MAP approximation until $k = 10000$

5.4 Application of the fractional Gaussian noise

We introduced the fractional Gaussian noise in Section 4.1.3, where we roughed out the applicable settings. These parameters and the applied settings are enumerated in the table below:

parameter	value
H	0.8
the mean of the trace	1
the peakedness of the trace	3.22
the length of the trace	$2^{20} \sim 10^6$.

We used the tool available at [13] to do the simulations. The Hurst parameter was set according to the estimation in 3.1.1, the mean of the trace after a normalization equals to 1, the peakedness was set according to the variance of the trace and the length was set to be almost the same as the original which is 10^6.

Variance-time plot

Based on the behaviour of the fractional Gaussian noise as told in Section 4.1.3 the Hurst parameter of the trace should be the same as the predefined setting. In our investigations the two value was not exactly the same because the simulation method resulted a finite sample set and the test, the Variance-time plot, estimates the Hurst parameter from this samples. However the result was fairly good. The Hurst parameters are the followings:

- $H_{\mathrm{BC}} = 0.8$,
- $H_{\mathrm{fGn}} = 0.77$.

This demonstrates the graph in Figure 5.35.

R/S plot

Although the simulation was based on the estimated result of the Variance-time plot the Hurst parameter of the fGn trace calculated with the R/S plot was also good (see Figure 5.36). Note that the estimated Hurst parameter of the original BC trace calculated with the R/S plot is 0.67 as told in Section 3.1.2.

- $H_{\mathrm{BC}} = 0.8$,
- $H_{\mathrm{fGn}} = 0.73$.

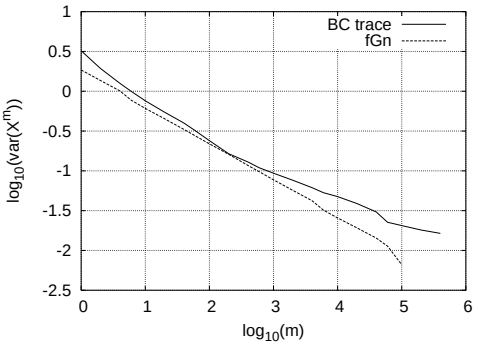


Figure 5.35: The Variance-time plot for the original BC and for the fGn trace

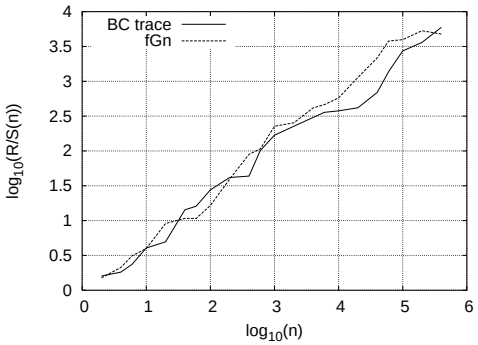


Figure 5.36: The R/S plot for the original BC and for the fGn trace

Haar wavelet

The statistical analysis via the Haar wavelet transform shows us that the monofractal behaviour of the trace generated by the fGn simulator appears as a straight line in the Haar wavelet graph. This graph in Figure 5.37 depicts the second moments of the d_j parameters, i.e. the different time scales. Observing this we can say that the only one parameter characterizing a monofractal is the slope of the line.

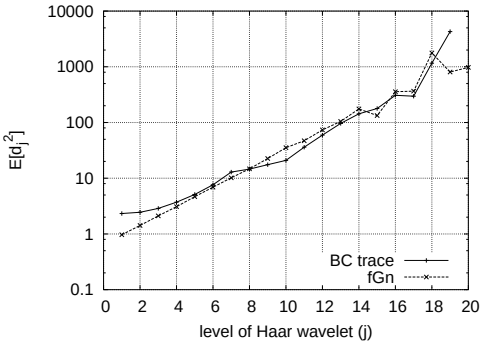


Figure 5.37: The Haar wavelet parameters of the original and the fGn trace vs. j

Legendre Spectrum

Generally the multifractal behaviour is observed via the Legendre spectrum. In this case we can analyse it with some restrictions originated in the fGn

simulation method. Normally he trace generated by the fGn simulator consists several samples smaller than 0, but an interarrival time trace consist positive data. The creator of the simulator changed these values to 0 to avoid the negative samples. This caused the Legendre spectrum not to be able to analyse the negative moments of the sums in (3.3). Accordingly the figures belonging to the Legendre spectrum have to be specially handled. The partition function is calculated and depicted only for ns greater or equal to 7 and the Legendre transform is made based on this partition function.

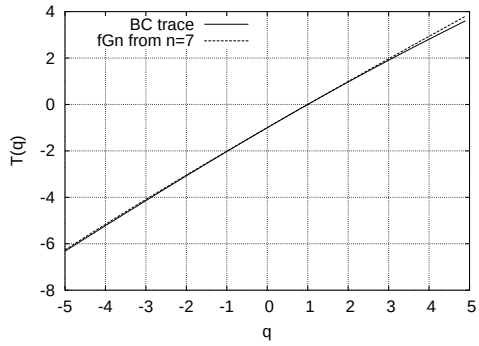


Figure 5.38: The partition functions of the original and the fGn trace

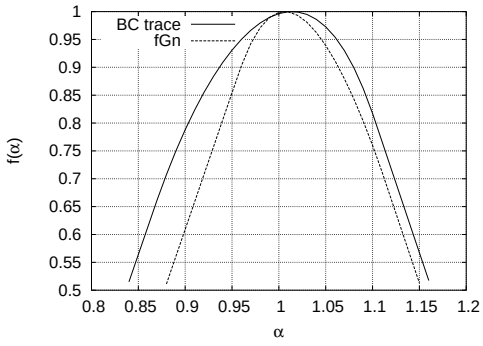


Figure 5.39: The Legendre transform of the partition functions in Figure 5.38

Taking into consideration these restrictions and observing the graphs depicted in Figure 5.38 and 5.39 we can say that the multifractal spectra of the original BC and of the fGn simulated trace differs for all αs.

Stationary queue length behaviour

As told in several previous sections (for example in Section 4.3) the better fitting of given Haar parameters results in the better approximation of the queue length distribution. When almost the whole range of $E[d_j^2]$ is fitted the queue length distribution is captured for all utilization levels. In case of fGn the original trace is assumed monofractal thus the Haar parameters are not fitted perfectly, i.e. the queue length distribution of the simulated trace differs from the original. It seems that the fGn trace with $H = 0.8$ approximates the queue length distribution fairly good at $\rho = 0.4$ utilization level. At other utilization levels the tendency is worse.

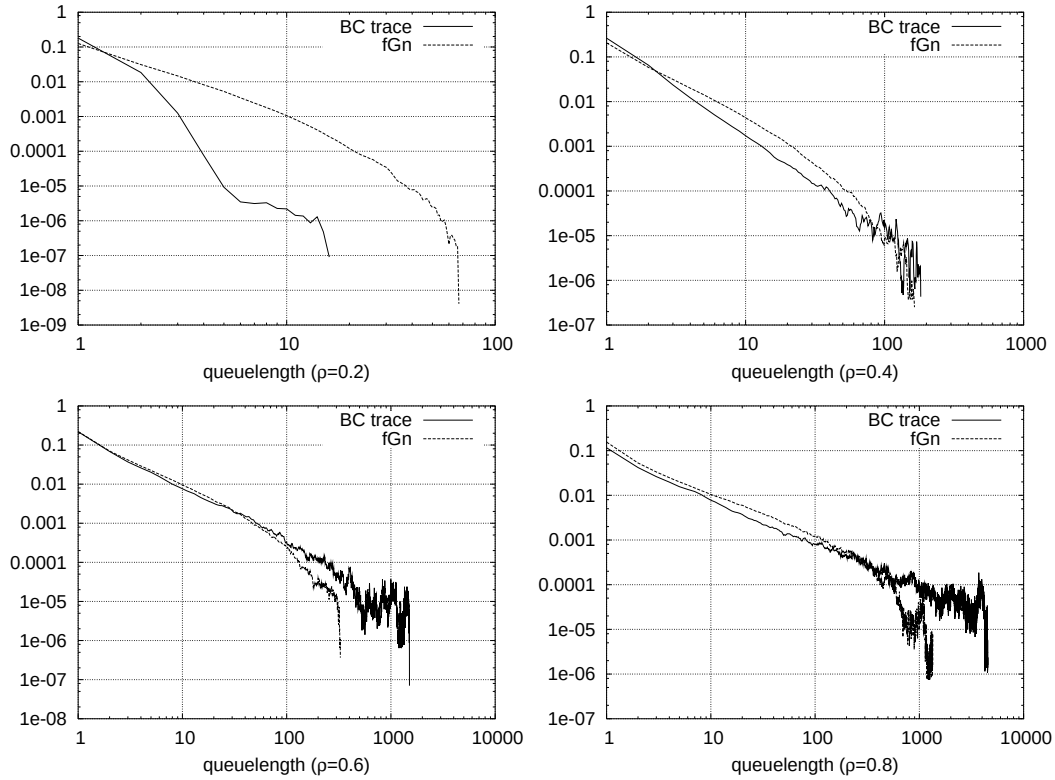


Figure 5.40: The probabilities vs. the queue length for the original and the fGn trace with several utilizations

Transient queue length behaviour

At this point we have the ability to deduce the transient queue length behaviour from the queue length distribution.

At low utilization level ($\rho = 0.2$) the transient queue length consists of several peaks larger or equal to the one in the original BC trace.

At utilization level $\rho = 0.4$ the queue behaves similarly to the original one, it has same number of peaks with the same value.

At higher utilizations the queue consists of fewer peaks with smaller values.

Marginal distribution

The fGn simulator is not prepared to capture the marginal distribution of the original trace and this can be seen on the results of test performed to show this property of the simulator.

Both the complementary cumulative distribution function (Figure 5.41)

and the moments of the trace (Figure 5.42) differs from the similar property of the original BC trace.

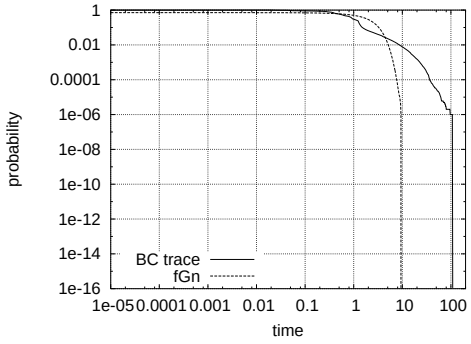


Figure 5.41: The complementary cumulative distribution functions of the two traces

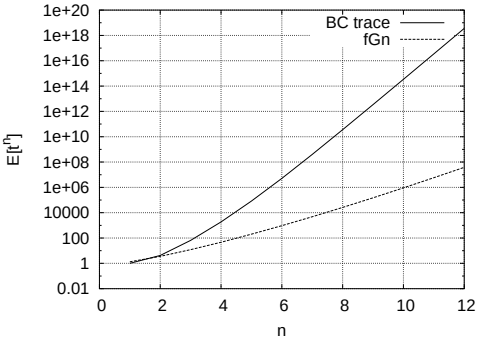


Figure 5.42: The first twelve moments of the two observed traces

Lag-k correlation

Although the fractional Gaussian noise generator was not optimized to fit the lag correlation structure of the trace the results are reasonably well. For the first ks there is a constant difference between the original and the simulated trace and for the higher ks the difference grows, as it can be seen in Figures 5.43 and 5.44, but the tendency is good.

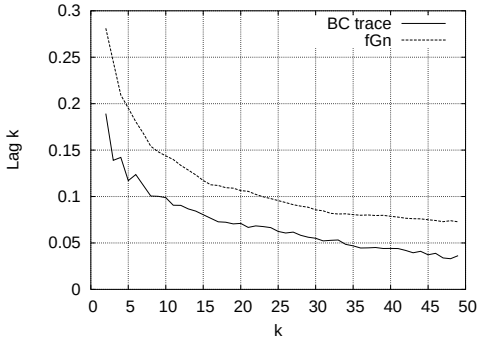


Figure 5.43: The lag-k correlation of the BC trace and of the fGn trace until $k = 50$

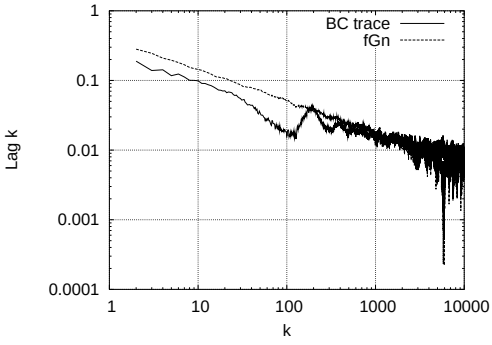


Figure 5.44: The lag-k correlation of the BC trace and of the fGn trace until $k = 10000$

5.5 PHFit

The tool presented in [7] is applicable to fit the marginal distribution of traces. The generated trace consist completely independent samples, i.e. there will be no long range dependency in the trace. Although we will analyse it to gain the goodness of a simulation model capturing the marginal distribution.

According to the introduced parameters in Section 4.2 the settings are the followings for us:

parameter	value
the time domain	continuous
the number of phases to fit the body	4
measure to fit the body	cdf area difference
the upper bound of fitting the body	0.05
the number of phases to fit the tail	2
the lower limit of fitting the tail	0.01
the upper limit of fitting the tail	0.1225
the distance of adjacent points	0.4.

The result was the row vector of the initial probabilities (α), the matrix containing the transition rates in the transient states ($\mathbf{B}$) and the column vector containing the transition rates from the transient states to the absorbing one ($\mathbf{b}$).

From these variables we calculate a MAP $\mathbf{D_0} = \mathbf{B}, \mathbf{D_1} = \mathbf{b} * \alpha$ with which we can generate a trace.

Variance-time plot

The Variance-time plot resulted in a line with the possibility to fit a straight line on it, but as one can see in Figure 5.45 the fitting of this parameter was poor. The slope of the two lines are significantly different and the estimated Hurst parameters differs too.

- $H_{\mathrm{BC}} = 0.8$,
- $H_{\mathrm{PHFit}} = 0.49$.

Based on this test we can say that the LRD of the original Bellcore trace was not captured and the simulated trace do not exhibits long range dependent behavior.

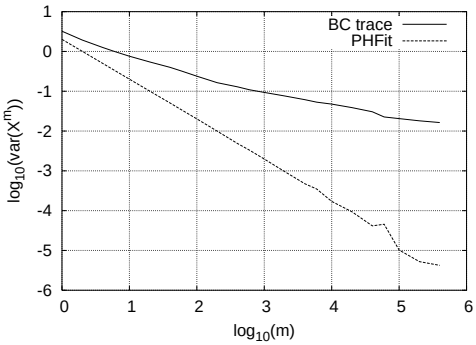


Figure 5.45: The Variance-time plot for the BC trace and for the PHFit trace

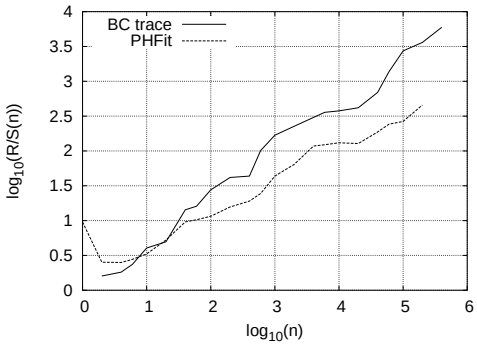


Figure 5.46: The R/S plot for the original and for the PHFit trace

R/S plot

Performing these test for long range dependency the result is the same as it was with the Variance-time plot. The PH fitted trace differs from the original in all the time scales (see Figure 5.46) and is not LRD. The estimated Hurst parameters are:

- $H_{\mathrm{BC}} = 0.8$,
- $H_{\mathrm{PHFit}} = 0.44$.

Haar wavelet

To examine the multirfactal behaviour we use the Haar wavelet transform. This resulted in a graph similar to those calculated from the Haar parameters of the Poisson process and the randomly scrambled trace. This is because the PHFit fits the marginal distribution of a trace with completely independent samples.

Stationary queue length behaviour

Based on the previous paragraph our expectations are that the fitting of the queue length behaviour is poor at high utilizations, but at low utilizations the body should be captured.

Taking a look at the graphs in Figure 5.48 the queue length distributions meet the expectations.

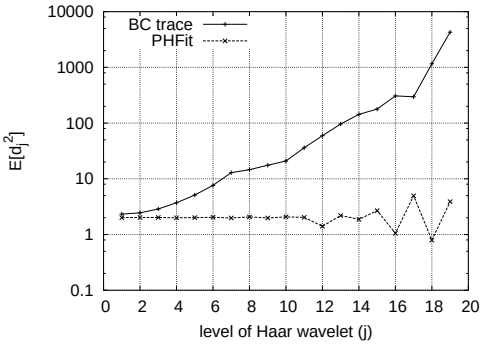


Figure 5.47: The Haar wavelet parameters of the original and the two kinds of MAP traces vs. j

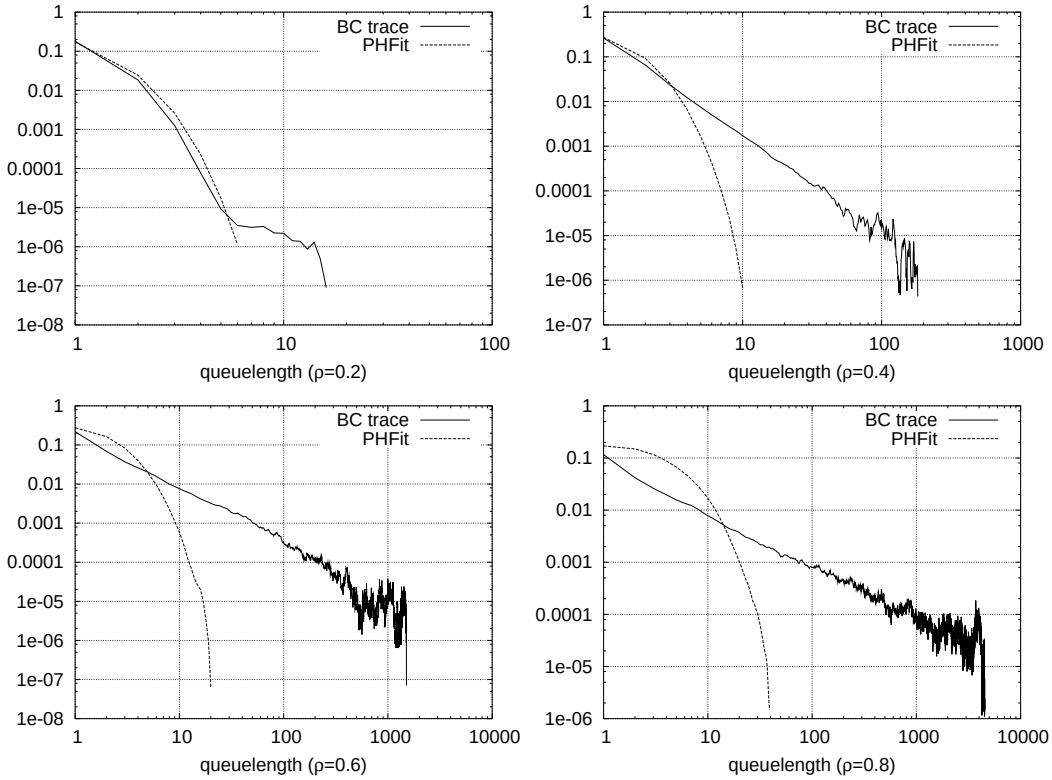


Figure 5.48: The probabilities vs. the queue length for the original and the PHFit trace with several utilizations

Marginal distribution

Observing the ccdf of the PH fitted trace in Figure 5.49 we see the perfect fitting of the marginal distribution of the original BC trace. This is confirmed

in Figure 5.50 where the moments of the two traces differs only a bit, but take into consideration that none of the previous simulation models was able to capture the moments of the trace like this one.

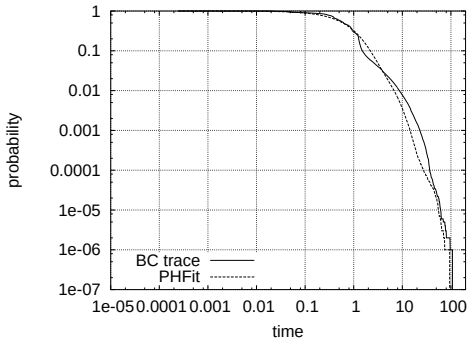


Figure 5.49: The complementary cumulative distribution functions of the two traces

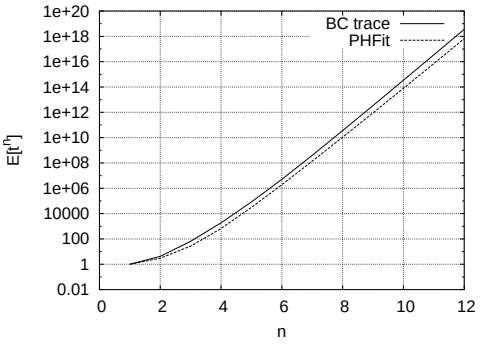


Figure 5.50: The first twelve moments of the two observed traces

Lag-k correlation

PH fits the marginal distribution, but do not care the correlation structure. This property can be recognized in Figures 5.51 and 5.52. There is no correlation for the whole range of k.

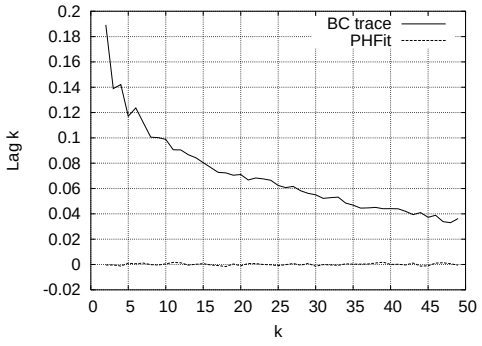


Figure 5.51: The lag-k correlation of the BC trace and of the PHFit trace until $k = 50$

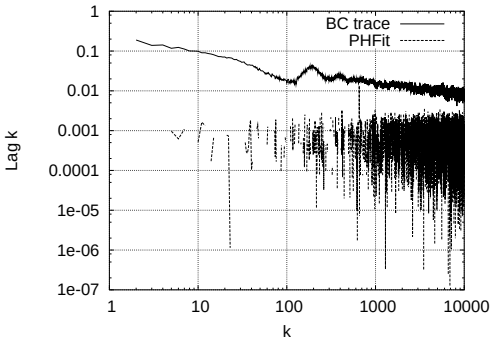


Figure 5.52: The lag-k correlation of the BC trace and of the PHFit trace until $k = 10000$

Chapter 6

Conclusion

Four traffic models were analysed, the fractional Gaussian noise, the PH renewal process, the Markov Arrival Process and the Multifractal Wavelet Model. The MAP model and the MWM were modified 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 two commonly available real data traces of the internet archive [16] with respect to a set of statistical tests, which contains the most popular tests in time series analysis.

Our experiments indicate that the MAP model and the MWM 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 of statistical errors due to the finite length of the traces.

Due to the poor approximation of the queue length distribution at low utilization level and the recognition of the correspondence between the fitting of the time scales and the queue length at different utilization levels the MAP model was modified again. The second modification of the MAP was to capture both the body and the tail of the queue length distribution. The approximation of the queue length distribution was successful, but the fitting of other statistical properties worsened.

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 using the considered MWM and MAP models. 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, the PH fitting and the trace generator method and helped in theirs use. A special thanks to Miklós Telek for his instructions and help.

Appendix A

Application of simulation models to the DEC traffic trace

To show the availability of the tools presented in previous sections we show the fitting of another traffic trace available at [16]. This is the TCP part of the DEC trace containing an hour's worth of all wide-area traffic between Digital Equipment Corporation and the rest of the world.

A.1 Application of the MWM

Since the modification of the Multifractal Wavelet Model was not effective from the marginal distribution point of view we compare only the original (β distributed A_j) MWM simulated trace with the original one. The fitting method is the same as with the Bellcore data set as described in Section 4.1.1.

Variance-time plot

One test for long range dependency is the Variance-time plot. Analysing the DEC-TCP trace with this we found that the trace is long range dependent with Hurst parameter $H = 0.81$.

Comparing the approximate MWM with the original trace in Figure A.1 shows that the MWM simulated trace has almost the same Hurst parameter:

- $H_{\text{DEC-TCP}} = 0.81$,
- $H_{\text{MWM}} = 0.80$.

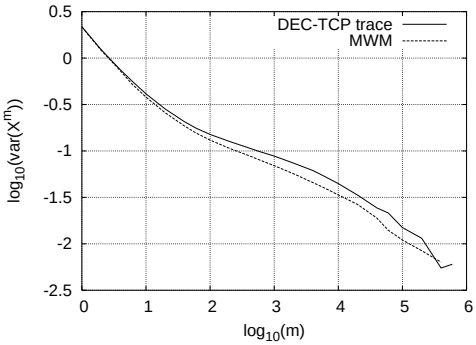


Figure A.1: The Variance-time plot for the DEC-TCP trace and for its MWM approximate

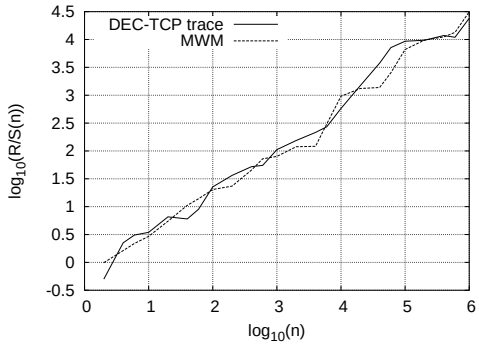


Figure A.2: The R/S plot for the DEC-TCP trace and for its MWM appriximate

R/S plot

The estimated Hurst parameters of both the original and its MWM approximate were in the range $[0.5 < 1]$, i.e. both of them are long range dependent based on this test too. The Hurst parameters are:

- $H_{\text{DEC-TCP}} = 0.80$,
- $H_{\text{MWM}} = 0.78$.

This can be checked in Figure A.2, where the R/S plots of them are very similar.

Haar wavelet

As introduced in Section 4.1.1 the Multifractal Wavelet Model simulates the trace as the result of a coarse to fine simulation method from up to down in the Haar wavelet tree. It follows that the Haar parameters of this simulation approximates well the original ones. This can be checked in Figure A.3.

The differences at large js come from the fact that the second moments depicted in Figure A.3 are calculated from finite sample series and if $j \geq 16$ the number of samples is less or equal to 32.

Stationary queue length behaviour

In the case of the DEC-TCP trace the tail at $\rho = 0.2$ utilization level do not appear, thus the fitting of the queue length distribution at this utilization is reasonably well with the MWM.

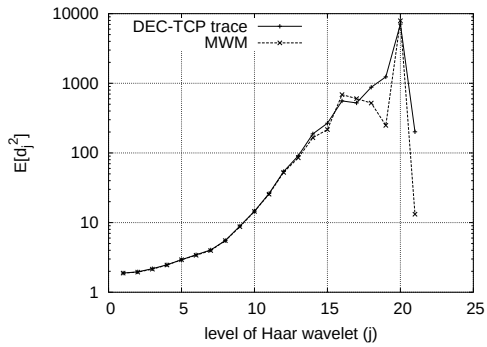


Figure A.3: The Haar wavelet parameters of the DEC-TCP trace and of the MWM traces vs. j

At medium utilization levels the queue length distribution of the DEC-TCP trace shows a long tail. This property is not captured by the MWM fitting method, which is not able to explore this behaviour.

See Figure A.4 for detailed comparison of the queue length distribution of the original and the MWM simulated trace.

Marginal distribution

The fitting of the marginal distribution of the DEC-TCP trace was unexpectedly successful. Both the fitting of the ccdf and the approximation of the moments of the trace was fairly good. This is depicted in Figure A.5 and A.6.

According to this observation we believe that the ccdf of a longer and much more ordinary trace than the Bellcore has a smooth ccdf. A ccdf like this is easy to approximate, i.e. to fit almost perfectly the moments of it.

Lag-k correlation

The tendency of the lag-k correlation function was fitted reasonably well (see Figure A.7 and A.8 for details). The oscillation is originated in the simulation method. The samples in the trace with the distances of the powers of 2 have a larger correlation. This is because the ith step in the simulation method is responsible for the 2^i long dependency and the random division of the time intervals causes the smaller correlation at odd ks.

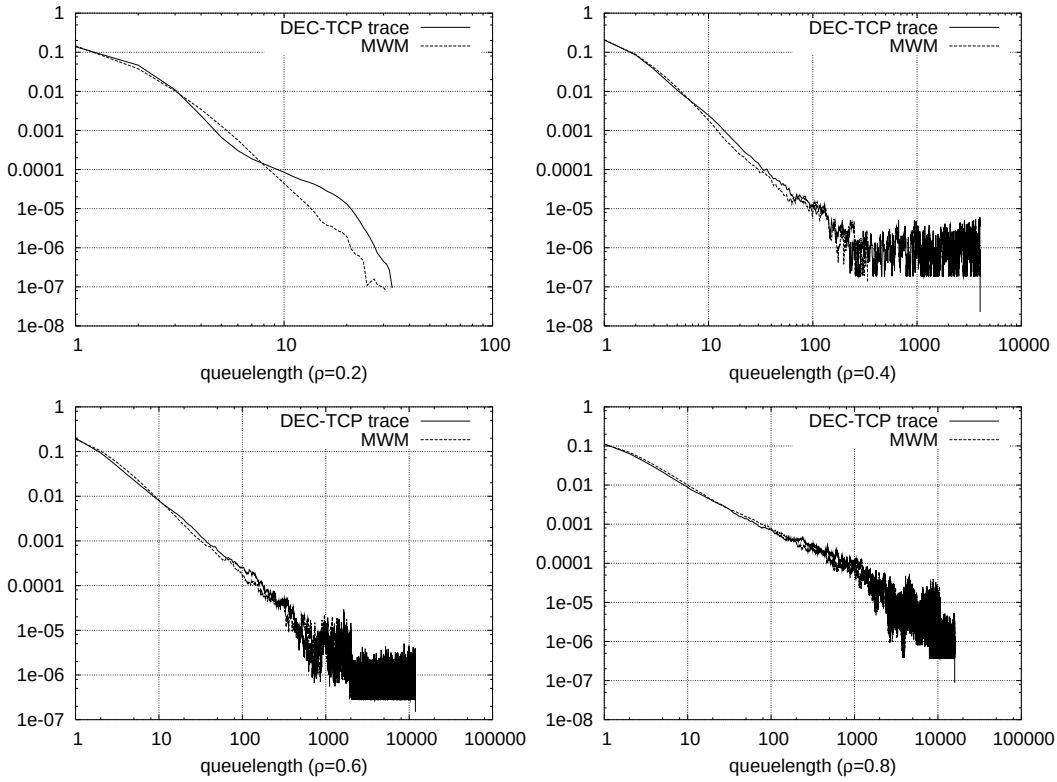


Figure A.4: The probabilities vs. the queue length for the DEC-TCP and the MWM traces with several utilizations

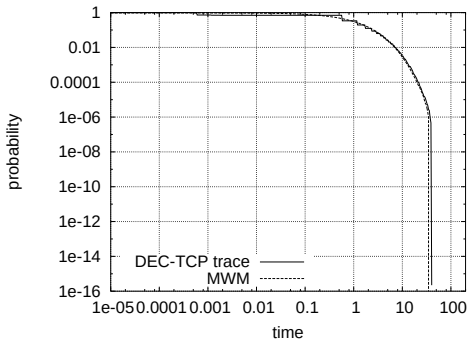


Figure A.5: The complementary cumulative distribution functions of the two traces

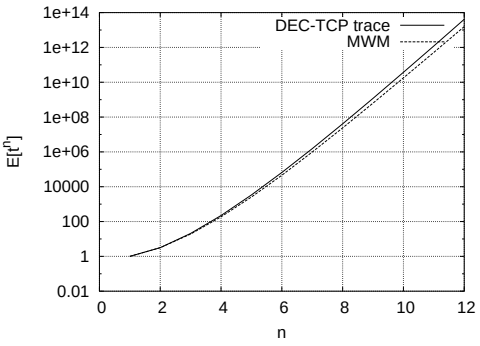


Figure A.6: The first twelve moments of the two observed traces

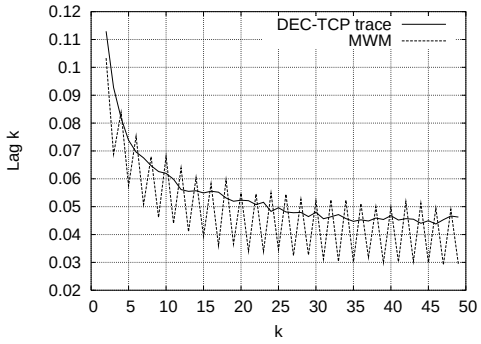


Figure A.7: The lag-k correlation of the DEC-TCP trace and of its MWM approximate until $k = 50$

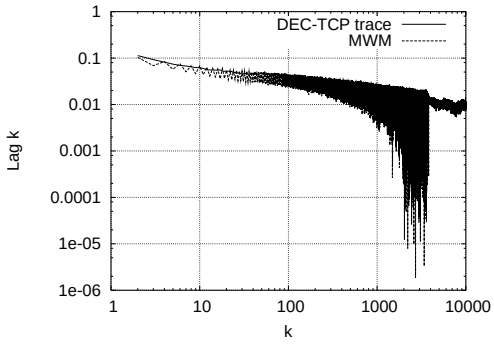


Figure A.8: The lag-k correlation of the DEC-TCP trace and of its MWM approximate until $k = 10000$

A.2 Application of the MMPP model

In the case of the MMPP model we also fitted the DEC-TCP only with the symmetric model. The settings of the model parameters were changed according to that the original DEC-TCP trace contains about two million arrivals. The parameters were:

- $n = 6$ dimensions,
- $T_M = \frac{1}{\lambda}$ was set to the mean of $N/4$ arrivals,
- $\gamma = 8$.

Variance-time plot

There were differences in the Variance-time plots of the original and the MAP simulated trace, but both of the estimated Hurst parameters was in the range $]0.5, 1[$.

- $H_{\text{DEC-TCP}} = 0.81$,
- $H_{\text{MAP}} = 0.7$.

R/S plot

Based on the R/S plot both of the traces are LRD, but a difference appeared in the plots. The Hurst parameters are:

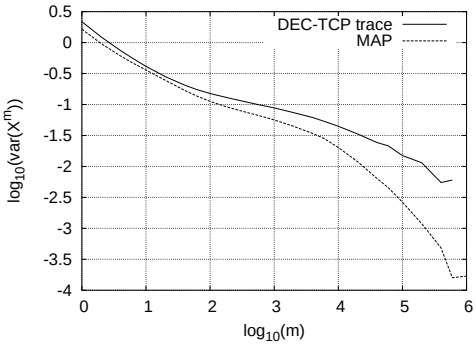


Figure A.9: The Variance-time plot for the DEC-TCP trace and for its MAP approximate

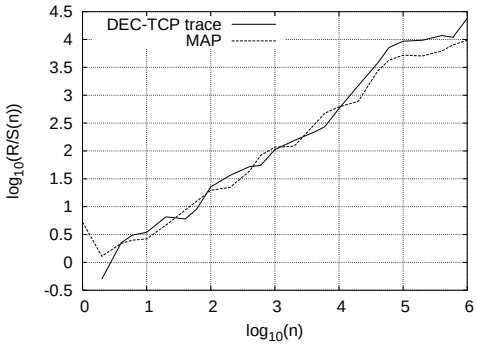


Figure A.10: The R/S plot for the DEC-TCP trace and for its MAP approximate

- $H_{\text{DEC-TCP}} = 0.80$,
- $H_{\text{MWM}} = 0.7$.

Haar wavelet

Although we have modified the goal function of the optimization for the distance of the Haar parameters the result differs from the original DEC-TCP trace. We optimized the distance of the first 16 Haar parameters instead of the first twelve. The result can be seen in Figure A.11.

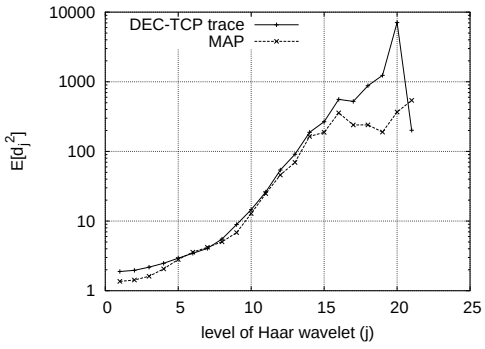


Figure A.11: The Haar wavelet parameters of the DEC-TCP trace and of the MAP traces vs. j

Stationary queue length behaviour

The queue length distribution was not captured at all utilization levels. The results are depicted in Figure A.12.

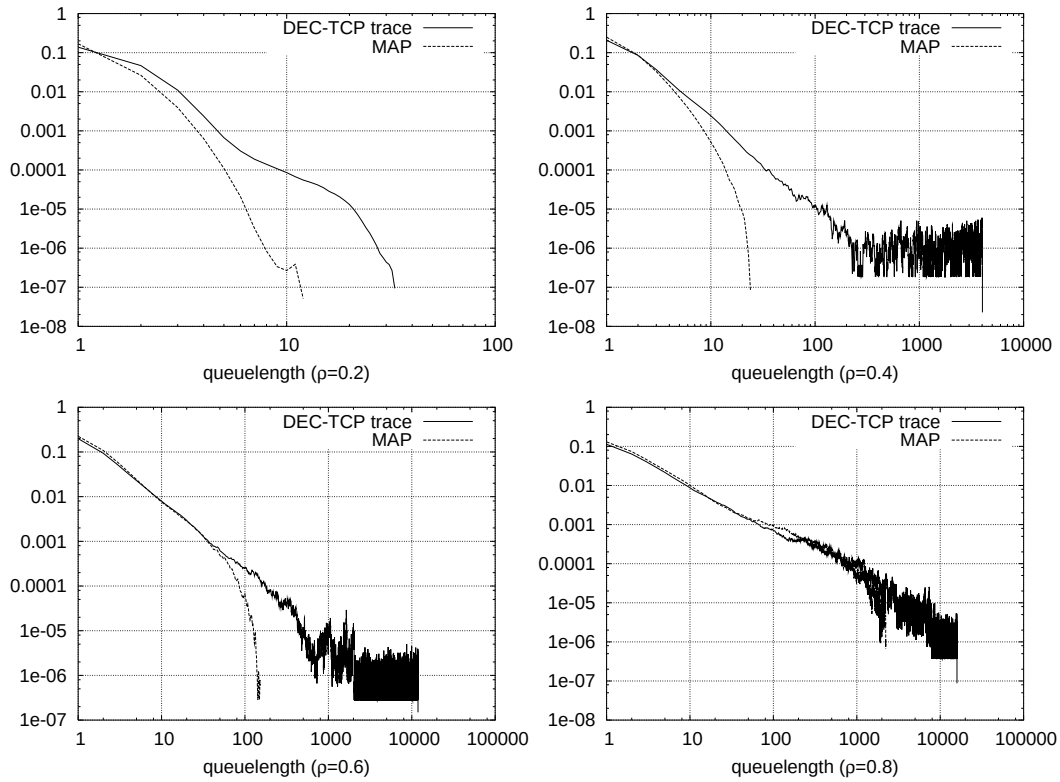


Figure A.12: The probabilities vs. the queue length for the DEC-TCP and the MAP traces with several utilizations

Marginal distribution

The fitting of the marginal distribution was reasonably good. This is due to the smoothness of the original ccdf.

Lag-k correlation

The fitting of the higher correlations was good, but the fitting was poor for the most important low ks. This can be seen in Figure A.15 and A.16.

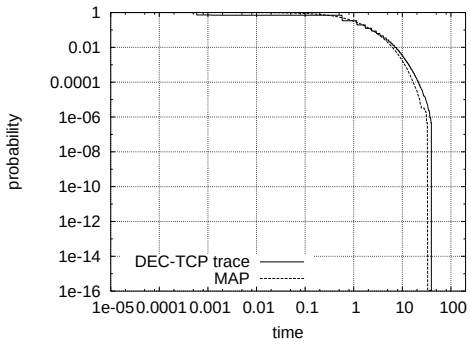


Figure A.13: The complementary cumulative distribution functions of the two traces

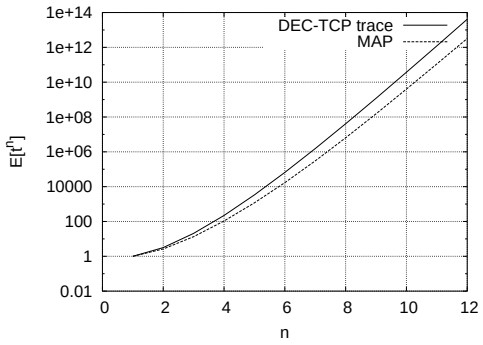


Figure A.14: The first twelve moments of the two observed traces

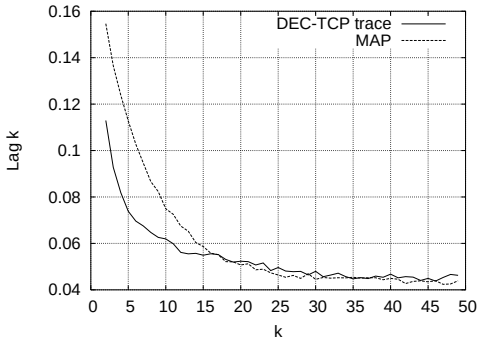


Figure A.15: The lag-k correlation of the DEC-TCP trace and of its MAP approximate until $k = 50$

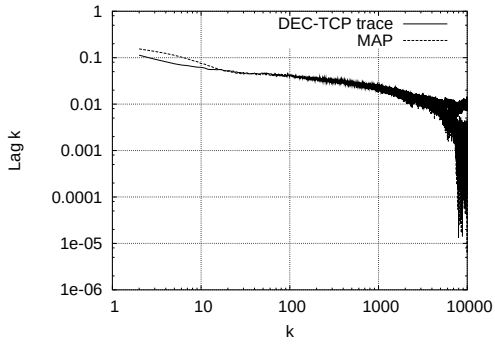


Figure A.16: The lag-k correlation of the DEC-TCP trace and of its MAP approximate until $k = 10000$

A.3 Application of the fractional Gaussian noise model

Variance-time plot

One of the three adjustable parameters of the fractional Gaussian noise is the Hurst parameter. This and the fact that the Variance-time plot estimates the self-similarity through the Hurst parameter suggests us that the aim of the almost perfect fitting of the long range dependency is not an excess with the fGn.

However the Hurst parameter setting of the fGn simulation was based on the estimation of the Variance-time plot, the result shows difference from the original trace. This is depicted in Figure A.17. The estimated Hurst parameters are the followings:

- $H_{\text{DEC-TCP}} = 0.81$,
- $H_{\text{fGn}} = 0.78$.

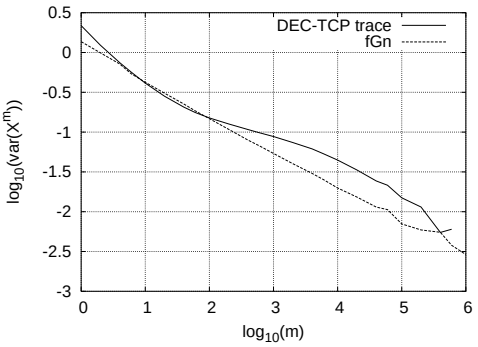


Figure A.17: The Variance-time plot for the DEC-TCP trace and for its fGn approximate

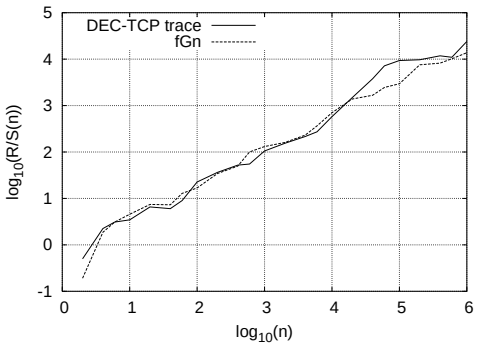


Figure A.18: The R/S plot for the DEC-TCP trace and for its fGn appriximate

R/S plot

The other test estimating the self-similarity is the R/S plot. Based on this analysing method we can say, that both the original and the approximating trace are long range dependent with almost the same Hurst parameter. For the results see Figure A.18, the estimated Hurst parameters are:

- $H_{\text{DEC-TCP}} = 0.80$,

- $H_{\text{fGn}} = 0.76$.

Haar wavelet

After the examination of the Haar parameters of the Bellcore and the DEC-TCP trace our observations about the monofractals are the followings.

The assumption in Section 2.2 was that the scaling behaviour of the absolute moments of the monofractal processes are characterized by a single number the Hurst parameter. In Section 4.1.3 the properties of the fGn was introduced and one important property is that the fGn is statistically self-similar, i.e. a monofractal. Its behaviour is given through its parameters, the mean, the Hurst parameter and the peakedness. Based on this and the examination of the fGn simulated trace in 5.4 we can say that the only one parameter of the trace describing its long range dependency is appearing in the slope of the straight line of its Haar parameter graph in Figure A.19.

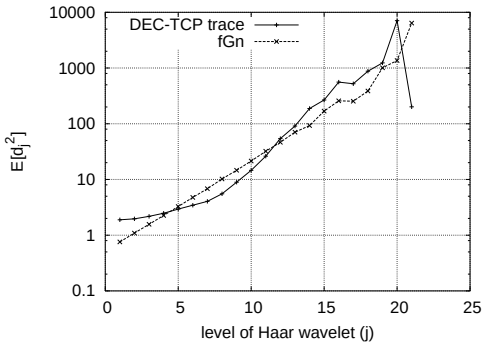


Figure A.19: The Haar wavelet parameters of the DEC-TCP trace and of the fGn traces vs. j

Stationary queue length behaviour

The stationary queue length behaviour of the fGn simulated trace matches our experiences in the case of the DEC-TCP trace.

There is only one queue length distribution captured. This is the one at $\rho = 0.4$ utilization level except the tail which cannot be captured with the fGn simulation, but the body is fitted well (see Figure A.20).

Marginal distribution

As in the case of the Bellcore data set the marginal distribution of the DEC-TCP trace is not fitted too. The fGn simulation is not prepared to approxi-

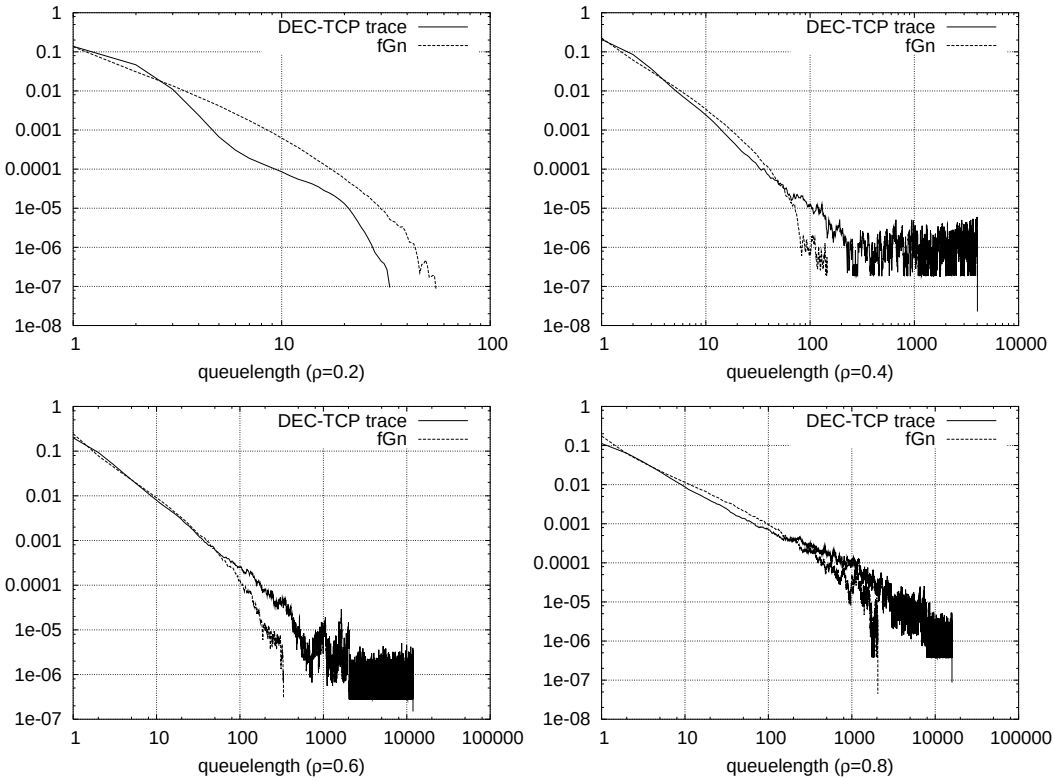


Figure A.20: The probabilities vs. the queue length for the DEC-TCP and the fGn traces with several utilizations

mate the marginal distribution of the original trace (see Figure A.21). Only the first two moments are captured, the first is set directly through the simulation and the second moment is captured through the peakedness of the trace the higher moments are different as one can see in Figure A.22.

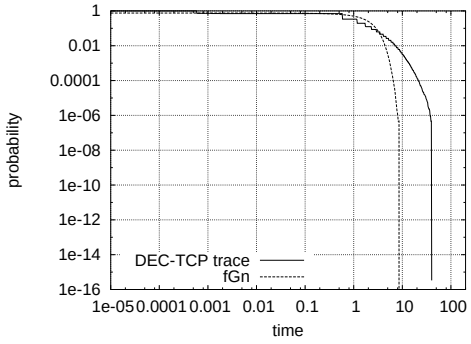


Figure A.21: The complementary cumulative distribution functions of the two traces

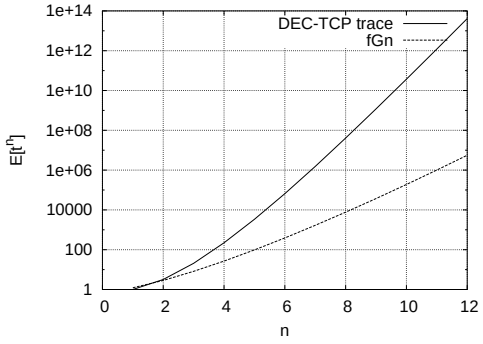


Figure A.22: The first twelve moments of the two observed traces

Lag-k correlation

Although the fGn simulation method is not prepared to fit the lag-k correlation structure, the higher lag coefficients were captured reasonably well (see Figure A.24), but the more important smaller coefficients differ from the originals (see Figure A.23).

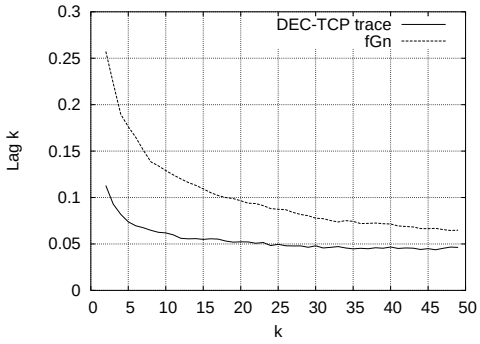


Figure A.23: The lag-k correlation of the DEC-TCP trace and of its fGn approximate until $k = 50$

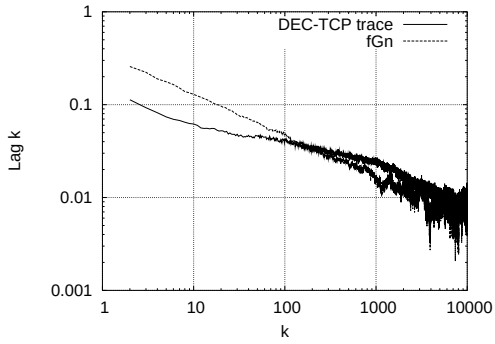


Figure A.24: The lag-k correlation of the DEC-TCP trace and of its fGn approximate until $k = 10000$

A.4 PHFit

The tool presented in [7] is also used to fit the DEC-TCP trace. The generated trace consist completely independent samples, i.e. there will be no long range dependency in it.

According to the introduced parameters in Section 4.2 in the case of the DEC-TCP trace the settings are the followings:

parameter	value
the time domain	continuous
the number of phases to fit the body	4
measure to fit the body	cdf area difference
the upper bound of fitting the body	0.05
the number of phases to fit the tail	2
the lower limit of fitting the tail	0.01
the upper limit of fitting the tail	0.525
the distance of adjacent points	0.4.

The result was the row vector of the initial probabilities (α), the matrix containing the transition rates in the transient states ($\mathbf{B}$) and the column vector containing the transition rates from the transient states to the absorbing one ($\mathbf{b}$).

From these variables we can calculate a MAP $\mathbf{D_0} = \mathbf{B}, \mathbf{D_1} = \mathbf{b} * \alpha$ with which we can generate a trace.

Variance-time plot

As told in the last section the PH fitted trace is not long range dependent. According to this the trace is only analysed with the Variance-time plot to observe its LRD instead of the two test introduced in Section 3.1.

As one can see in Figure A.25 the Variance-time plots differ and the estimated Hurst parameters are different too, the Hurst parameter of the PH simulated trace was rounded up to 0.5:

- $H_{\text{DEC-TCP}} = 0.81$,
- $H_{\text{PHFit}} = 0.50$,

i.e. the simulated trace is a border-line case from *pseudo* self-similarity point of view.

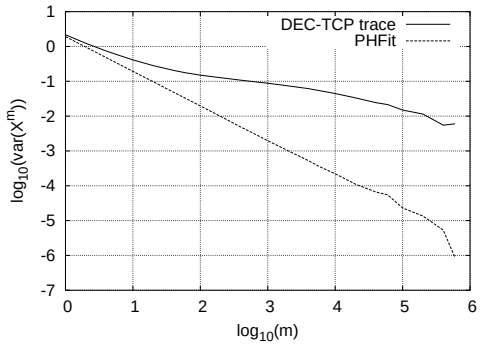


Figure A.25: The Variance-time plot for the DEC-TCP trace and for its PHFit trace

Haar wavelet

The Haar parameters of a trace containing completely independent samples results in a straight line. As told in Section 4.2 the PHFit tool approximates the marginal distribution of a trace through independent samples. Taking into account all these notes the Haar parameter graph of the PH simulated trace in Figure A.26 is confirmed.

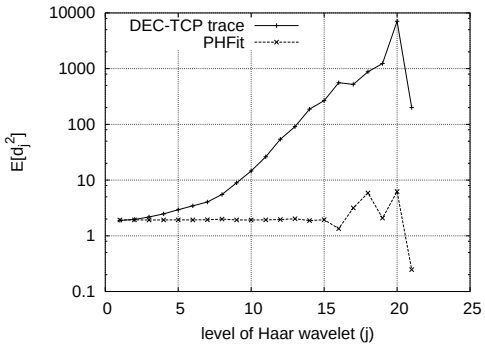


Figure A.26: The Haar wavelet parameters of the DEC-TCP trace and of the PHFit traces vs. j

Stationary queue length behaviour

The experienced correspondence between the Haar wavelet parameters and the queue length behaviour at different utilization levels (for details see Section 4.3) expounds the different graphs in Figure A.27.

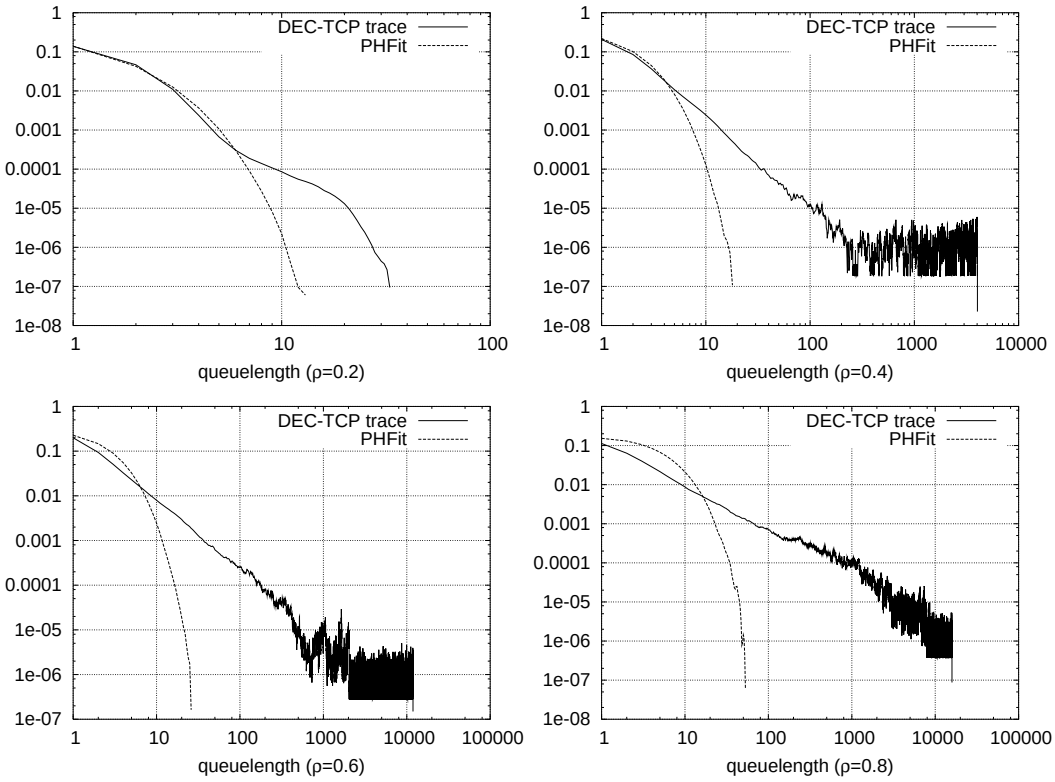


Figure A.27: The probabilities vs. the queue length for the DEC-TCP and the PHFit traces with several utilizations

Marginal distribution

The PHFit tool was presented in [17] and described in [7] to estimate the complementary cumulative distribution function of traffic traces. Although the goal of this tool is to fit the ccdf in the case of the DEC-TCP data set the MWM and the MMPP model approximated better the marginal distribution than the PHFit tool. Apart from this we can say that this tool had the same accuracy in the case of Bellcore trace. The conclusion is that the PHFit method can capture both a rough and a completely "ordinary" marginal distribution with the same goodness.

The results of the fitting are depicted in Figure A.28 for the ccdf and in Figure A.29 for the moments of the marginal distribution.

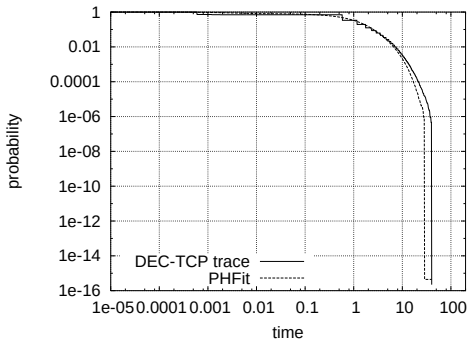


Figure A.28: The complementary cumulative distribution functions of the two traces

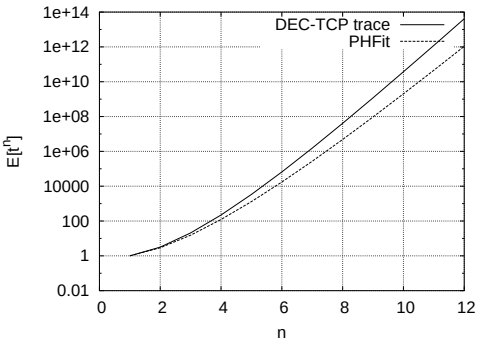


Figure A.29: The first twelve moments of the two observed traces

Lag-k correlation

According to the fact described in the introduction in Section A.4 the correlation between the samples is zero, this can be checked in Figures A.30 and A.31. The negative values can not be depicted in the logarithmic scale in Figure A.31.

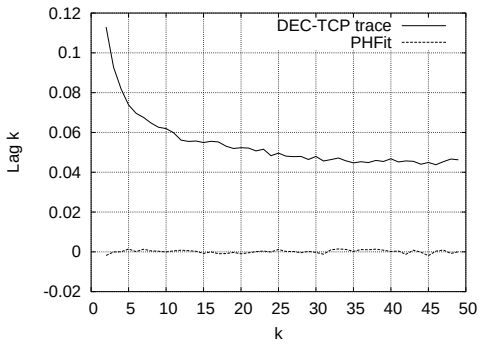


Figure A.30: The lag-k correlation of the DEC-TCP trace and of its PHFit approximate until $k = 50$

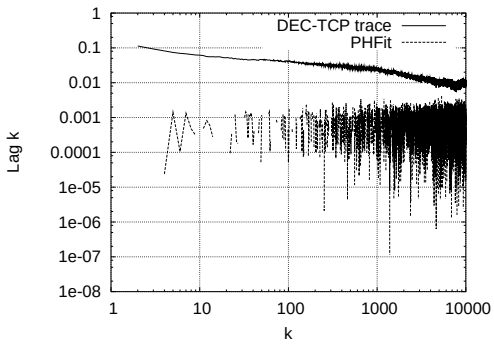


Figure A.31: The lag-k correlation of the DEC-TCP trace and of its PHFit approximate until $k = 10000$

Appendix B

Working environment

To make someone know the computation complexity and usability of the presented methods we describe either the simulation times and the problems arisen during the simulations.

The architecture which the simulations ran on was the following:

- Intel Celeron 768 Mhz processor
- 256 Mb memory
- 256 Mb swap size
- Linux operating system with 2.4.29 kernel
- Gnu Octave, version 2.0.16
- The c programs were compiled with gcc 3.2.2.

The running times were in the order of minutes except some special simulations. The exceptions were the random scrambling of the BC trace and the MAP fitting of the DEC-TCP trace. The time needed for the random scrambling is $\mathcal{O}(n^2)$ and in the BC case $n \sim 10^6$. The MAP simulation of the DEC ran on an Intel Pentium based machine in the order of hours.

Another problem was with the fractional Gaussian noise simulation, when the length of the simulated trace was larger than two million. In this case the memory and the swap space in the computer were small. After some attempts on an Intel Pentium based computer with 2 Gb of swap space the simulation was successfully completed.

Bibliography

[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] A. Horváth and M. Telek. Phfit: A general phase-type fitting tool. In *Proc. of 12th Performance TOOLS*, Imperial College, London, April 2002.

[8] G. Horváth, P. Buchholz, and M. Telek. A map fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation. May 2005.

[9] 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.

[10] 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.

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

[12] 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.

[13] Francesco Potortì. Software page. http://fly.isti.cnr.it/software/software.html#fractals.

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

[15] 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.

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

[17] Performance tools. http://webspn.hit.bme.hu/~telek/tools.html.