A Multifractal Wavelet Model of Network Traffic
Project Report for ENSC802

Jinyun Ren
School of Engineering Science, Simon Fraser University
Email: jren@cs.sfu.ca

Abstract

In this paper, a new Multifractal Wavelet Model (MWM) is studied and simulation results are presented. We conclude that MWM has captured most of the properties of network traffic and the computational complexity of MWM is only O(N). The efficiency of MWM makes it possible in the real-time applications.

1.Introduction

Network traffic modeling is one of the key topics in the application of network and multimedia. Taking the multimedia download as an example, a well-designed model will help make use of the network efficiently and speed up the transmission of data. One of key issues of multimedia download is how to model the communication channels. This is because after we get the accurate information, like distribution of the network delay, about channel, we may optimally distribute the multimedia data and fetch them by optimal download strategy. The more accurate the information is, the better we may find the strategy.

Here we concern the accuracy of the network information in a statistical sense. What we like to know is how to predict the real-time delay of each link over which the data is transmitting, which has the close relation with network traffic modeling. Lots of research that have been done is based on measurement data upon a specific Internet connection [13][14]. They tried to use currently existed models to match measured data, such as AR model, Bernulli model, 2-state Markov chain model, and k-th order Markov chain model. Although these methods give some solutions to prediction of network delay, nobody has concluded whether or not these methods can be generalized to any kinds of Internet situations. Actually we know the Internet channel is so complex that these conventional statistical methods really cannot capture the rich properties of Internet [3].

Nonetheless, a bunch of researchers in Rice University are doing the challenging jobs with excellent insight [1]. By introducing the concepts of multifractal into Internet traffic model, they build a multifractal model that matches the traffic properties, such as Long-range dependency, self-similarity and burstiness, very well. Wavelets are used as a natural tool to analyze and synthesize the model due to its inherent multiscaling property. Based on these ideas, a
“Multifractal Wavelets Model (MWM)” is presented.

Based on MWM, V. Ribeiro etc. also develop a novel algorithm to estimate multifractal cross-traffic [2], which is equivalent to delay distribution [3]. The algorithm is adaptive, effective and requires no a priori traffic statistics, which greatly proves the effectiveness and efficiency of MWM and provides new method to the application in the multimedia area.

In section 2, we give the tutorial about the fractal, multifractal and wavelets, and some general ideas how and why these concepts can be related to network traffic model. In section 3, we provide the performance criterions for network traffic models. After we show how to extract the information of these criterions from traffic data, we build the multifractal wavelet model following these criterions. Section 4 presents the simulation results and gives the comments to MWM based on the criterions we provided. It can be shown that MWM capture most the properties of network traffic, such as self-similarity, long-range dependence, burstiness and multiscaling behaviors, with computational complexity of only O(N). We also show a simple approach to improve the performance of this model by simulation results. We closed this paper with conclusion and future work.

2. Fractal, Multifractal And Wavelets

2.1 From Fractal To Multifractal

By introduction of idea of fractal, researchers have made a great progress in modeling network traffic compared with classical models. A carefully designed fractal model can to great extent capture the fractal properties of network traffic, such as self-similarity, burstiness and Long-Range Dependence.

2.1.1 Fractal And Fractal Random Process

Fractal is a natural phenomenon that is of the self-similarity on finer and finer level behind the extremely irregular shape, such as coastline, snowflake. A fractal object is measured with a dimension that is fractional [5].

For network traffic, we cannot capture its rich properties with those classical models dealing with only time and sample value. A fractal random process has so irregular sample values that its graph has an “effective” dimension that exceeds its topological dimension of unity. Typically, the dimension of a fractal random process is between 1 and 2. [6]. This irregular property of fractal signal model makes it possible to be used to model network traffic.

2.1.2 Fractional Brownian Motion

Fractional Brownian motion (fBm) has been broadly used to model fractal random process [1]. It is a generalized version of Brownian motion or Wiener process by making the increments of
Wiener process normally distributed but no longer independent [7]. It preserves the self-similarity of Wiener process. The pdf of its increment is [7]

\[ P(\Delta B(t)) = P(B(t + \varepsilon) - B(t)) = \frac{1}{\sqrt{2\pi \varepsilon^H}} \int_{-\infty}^{\infty} \exp\left(-\frac{u^2}{2\varepsilon^{2H}}\right)du, \quad 0 < H < 1 \]  

(1)

We can see that its increment process, which is known as fractional Gaussian noise (fGn), has normal distribution with mean zero and variance $\varepsilon^{2H}$. When $a=1/2$, it is the Brownian motion case. From pdf, we can show the correlation functions of fBm and fGn are

\[ E[B(t_1)B(t_2)] = \frac{1}{2}(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}) \]  

(2)

\[ E[\Delta B(t + \tau)\Delta B(t)] = E[(B(t + \varepsilon + \tau) - B(t + \tau))(B(t + \varepsilon) - B(t))] 
= \frac{1}{2}( |\tau + \varepsilon|^{2H} + |\tau - \varepsilon|^{2H} - 2|\tau|^{2H} ) \]  

(3)

From these correlation functions, we know fBm is not stationary, but fGn is stationary.

From the pdf of (1), we see that

\[ B(t) = a^{-H}B(at) \quad a \text{ is constant} \]  

(4)

Where = is in the statistical sense, which means $B(t)$ and $B(at)$ is of the same distribution and is statistical self-similar.

Although fBm and fGn offers a simple and good way to model nonstationary random processes, and some of authors has done a good simulation with Wavelet [8], but they have some limitations in modeling network traffic [1][6]:

1. fBm and fGn are models based on Gaussian distribution, which means there must exist some negative signal. However, the data of network traffic should not be less than zeros.
2. Network traffic will display short-term correlation as well as well-known Long-Range dependence. It is inconsistent with self-similarity fGn and fBm provide.
3. Unlike the fBm and fGn, which have a constant scaling value at different moment orders, the scaling behaviour of moments of network traffic is not homogeneous.

2.1.3 Multifractal Measure

It is natural to think about the question “can we partition network traffic into multiple number of fractal sets such that on each individual partition, the measure is homogeneous or unifractal?” The answer is yes and this is also the motivation of multifractal. From the motivation, we know multifractal refers to a measure method [5].

3
2.2 Multifractal Theory

There are many special definitions in multifractal theory [5]. We give some general ideas related to the modeling of network traffic.

2.2.1 Local behavior---Multifractal spectra

Define for an increasing process $Y$ at time $t$ [1]

$$\alpha(t) = \lim_{k_n \to t} \alpha_n$$

(5)

with

$$\alpha_n = \frac{1}{\log_2 \Delta_n[Y]}$$

(6)

$$\Delta_n[Y] = \left| Y((k_n + 1)2^{-n}) - Y(k_n 2^{-n}) \right|$$

(7)

$k_n = 0, 1, \cdots 2^n - 1$

The smaller the $\alpha(t)$, the faster $Y$ grows at $t$.

Define

$$f_G(\alpha) := \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \log_2 \left\{ \alpha_n \in (\alpha - \varepsilon, \alpha + \varepsilon) \right\}$$

(8)

as (grained) multifractal spectrum, which is the frequency measure of occurrence of a given strength $\alpha$.

Here are some physical interpretations of $\alpha$ and $f_G(\alpha)$:

$\alpha > 1$, means $Y$ has small increment, $\alpha < 1$ means $Y$ has “instant growth” called burstiness if applied to network traffic.

$f_G(\alpha) = 1$ means most of $Y$ increments is nearly equal to $2^{-n\alpha}$. $f_G(a) < f_G(b)$ means the probability of a is higher than b in process $Y$.

If we know the multifractal spectrum $f_G(\alpha)$, we can give a good interpretation to local behaviors of network traffic. The definition (8), however, is difficult to get with two limits in one equation. It will be shown that we can find an easy way to compute $f_G(\alpha)$ with the help of partition function and Legendre transform.

2.2.2 Global behavior---partition function
Define partition function as \[1\] 
\[\text{T}(q) := \lim_{n \to \infty} \frac{1}{n} \log_2 \mathbb{E} \left[ \sum_{k=0}^{n-1} (\Delta_k^n [Y])^q \right] \] 
where \( q \) is the moment of \( Y \) increments at different partitions. The \( \text{T}(q) \) could be interpreted as the extent to which the measure deviates from a measure uniformly distributed on its support, which can equivalently be used to interpret the global behavior of network traffic. When \( q > 0 \), it describes the measure of relatively high increment values, that is, businesse. When \( q < 0 \), it describes the measure containing relatively small but not-zero increments\[5\]. Specially, when \( q = 0 \), \( \text{T}(q) = -1 \); for a positive increment process, when \( q = 1 \), \( \text{T}(q) = 0 \).[1]

2.2.3 multifractal formalism---Legendre Transform

It can be shown that the multifractal spectrum \( f_G(\alpha) \) and partition function \( \text{T}(q) \) are closely related by Legendre transform. The relationship is as follows \[1\][5]:
\[ f_G(\alpha) = q\alpha - \text{T}(q) \quad \text{and} \quad f_G'(\alpha) = q \quad \text{at} \quad \text{T}'(q) = \alpha \] 
With \( \text{T}(q) \) in hand, by finding the derivative of \( \text{T}(q) \), that is \( \alpha \), at every point \( q \), we can plot the relationship of \( f_G(\alpha) \) versus \( \alpha \). (The details may see \[1\][5]). This is illustrated in figure 1[1].

![Partition Function and Multifractal Spectrum](image)

\[\text{Fig. 1 Typical partition function and Multifractal Spectrum}\]

2.3 Wavelet Analysis

2.3.1 Why Wavelet?

It is natural to choose Wavelet as our analysis tool to multifractal process. As well as low computational complexity, wavelets have inherent similar properties with multifractal signal.
As we know, wavelet analysis acts as a mathematical microscope that allows us to zoom in (to find details) or zoom out (to find global behavior) a signal. When the signal exhibits self-similarity under different aggregation levels, wavelets are naturally able to reveal it by its scaling ability. Actually the iterating approximations of wavelet analysis at coarser and coarser resolutions are an implicit way of aggregating data, and evaluating wavelet or details coefficients is just a refined way of computing increments [3].

2.3.2 Orthogonal Wavelet Decomposition

The increasing process $Y(t)$ can be decomposed by orthogonal scaling function $\phi_{jk}(t)$ and wavelet function $\psi_{jk}(t)$ as follows [1] [8][9]:

$$Y(t) = \sum_{k} U_{jk} \phi_{jk}(t) + \sum_{j=J_0}^{2J-1} \sum_{k} W_{jk} \psi_{jk}(t)$$

with $W_{jk} = \int Y(t) \psi_{jk}(t) dt$ called wavelet or detail coefficients

$$U_{jk} = \int Y(t) \phi_{jk}(t) dt$$

called scaling or approximation coefficients.

Where $j$ represents the number of scales of wavelet decomposition. A larger $j$ corresponds to a higher resolution. $k$ is the number of scaling coefficients or wavelet coefficient at the coarsest level. If $k > 1$, it means that we do not keep decomposing the signal till the scale with only one number, and also means we have several sub-trees in hand with the interpretation of binary tree to wavelet decomposition.

Instead of evaluating wavelet and scaling coefficients by the inner product in the definition (11), it is possible to recursively compute them by cascaded discrete filter [9]. Taking Haar wavelet as an example [1][9], the wavelet and scaling coefficients can be computed by

$$U_{jk} = 2^{-1/2} (U_{j+1,2k} + U_{j+1,2k+1})$$

and

$$W_{jk} = 2^{-1/2} (U_{j+1,2k} - U_{j+1,2k+1})$$

We can see it is easy to analyze and synthesize the signal of interest.

3. Analysis and Synthesis

So far we have had some backgrounds for the motivation of a Multifractal Wavelet Model (MWM). Now we will show how to link wavelets and multifractal together and how to construct
a good model for network traffic in details.

3.1 Criterions for a good network traffic model

1). Positive Data: Back to limitations of fBm and fGn we mentioned before, the key of our model is to generate positive data.

2). Multifractal Scaling: Can the model capture the multifractal scaling behaviors given the real trace to be multifractal? Here the scaling behaviors focus on scaling of moment as the signal is aggregated with different level.

3). Multifractal Spectra: It is kind of similar to last criterion but with the different respective. We have discussed multifractal spectrum is one of main criterions for a multifractal process, by which we can find what is the probability of occurrence of burstiness of network traffic.

4). Long-Range Dependence (LRD): It is one of the most important properties of network traffic.

5). Queueing Behavior: The queueing behavior of network traffic is important because it will affect the network management algorithm such as admission control that supports certain QoS demands [1].

3.2 Trace Analysis

First of all, we need to know how to extract the performance criterions from the trace data.

3.2.1 Long-Range Dependence

All processes with exact self-similarity exhibit LRD [3]. A process Y with property that its correlation is nonsummable is said to exhibit long-rang dependence (LRD). LRD can be equivalently characterized in terms of behavior of the aggregated processes [1]

\[ Y^{(m)}[n] = \frac{1}{m} \sum_{i=0}^{n} Y[i] \]  

(13)

here m is the number of data at a certain aggregation level. If Y is self-similar signal defined by equation (4), then (13) implies that \( Y[n] = m^{1-H} Y^{(m)}[n] \), n is the index of data, and

\[ \text{var}(Y[n]) = m^{2-2H} \text{var}(Y^{(m)}[n]) \]  

(14)

Taking log at both sides of (14), we get

\[ \log_2 \text{var}(Y^{(m)}[n]) = -(2-2H) \log_2 m + C \quad C = \log_2 (\text{var}(Y[n])) \]  

is constant  

(15)

If Y is of strict second-order scaling property, the log-log plot should be linear. By getting the slope p of the \( \log_2 \text{var}(Y^{(m)}[n]) \) as a function \( \log_2 m \), we can find
H=0.5*P+1
If 0.5<H<1, Y exhibits LRD property [3].

If we deal with the real data with Haar wavelet transform, by equations of (12), we know scaling coefficients at every level are just the aggregation of the process and m is equal to $2^j$ (scale j=0 represents the finest scale). To obtain a more accurate estimation of variance, we leave 15 numbers at the highest aggregation level, that is, the coarsest wavelet scale (see the routine “PlotVarianceTime.m”).

Network traffic is not strict second-order scaling process, but we can still use variance time plot as a tool to test its LRD [1].

3.2.2 Multifractal Scaling

We can get the information of multifractal scaling behaviors of network traffic from partition function. For convenience, we repeat its definition:

$$T(q) := \lim_{n \to \infty} \frac{1}{n} \log_2 E\left[ \sum_{k=0}^{2^n-1} (\Delta_n^k [Y])^q \right] = \lim_{n \to \infty} \frac{1}{n} \log_2 S_n(q)$$

With a larger number of measured data, we assume the process we are dealing with is ergodic [1]. Under this assumption, we can replace expectation with average of time or sample. After haar wavelet transform, $T_j(q)$ at every scale (aggregation level) can be shown as [1][3]

$$T_j(q) = \frac{1}{-j} \log_2 \sum_{k=0}^{2^j-1} 2^{-j/2} W_{jk}^q = \frac{1}{-j} \log_2 S_j(q)$$

$$\log_2 2^j T_j(q) = -\log_2 S_j(q) = -\left( -\frac{jq}{2} + \log_2 \sum_{k=0}^{2^j-2} |W_{jk}|^q \right)$$ (16)

If the process Y is of multifractal scaling behavior, a linear relation between j and $-\log_2(S_j(q))$ should be expected and the slope of these lines are the values of T(q) at point q’s. Here j is the scale of wavelet transform. A larger j represents a finer scale.

It should be aware that when q<0, T(q) is used to measure the data that is very small but nonzero. Here nonzero is of physical meaning, that is, when we deal with measured data, the resolution of measure instruments should be thought as zeros instead of real zeros.

3.2.3 Multifractal Spectrum

As shown in section 2.2.3, we can get every point of $f_G(\alpha)$ by $T_j(q)$ with the help of Legenre Transform

$$f_G(\alpha_j) = q_j \alpha_j - T(q_j) \quad \text{and} \quad T'(q_j) = \alpha_j$$ (17)
3.3 Building MWM

Now we discuss how to build a multifractal model with Haar wavelets transform.

3.3.1 Positive Data Guarantee

Wavelet-domain modeling of positive process is complicated due to the fact that the wavelet coefficients constraints required ensuring a positive output is not trivial. By equations (12) with Haar wavelet transform, we can get

\[ U_{j+1,2k} = 2^{-1/2} (U_{jk} + W_{jk}) \quad \text{and} \]
\[ U_{j+1,2k+1} = 2^{-1/2} (U_{jk} - W_{jk}) \]

As we know, all scaling coefficients \( U_{jk} \) of network traffic data will not be less than zero. Hence it is surprisingly easy to guarantee a positive data with Haar wavelet transform. All we need to do is let \( W_{jk} \leq U_{jk} \) \[1\].

(19)

To build a statistical model for \( W_{jk} \), we may find a random variable \( A \) supported on the interval [-1,1] and compute every wavelet coefficient by

\[ W_{jk} = A_{jk} U_{jk} \]

(20)

plug (20) into (18), we get

\[ U_{j+1,2k} = 2^{-1/2} (1 + A_{jk}) U_{jk} \quad \text{and} \]
\[ U_{j+1,2k+1} = 2^{-1/2} (1 - A_{jk}) U_{jk} \]

(21)

3.3.2 Modeling The LRD Property

Since we have applied a random variable \( A \) into model to guarantee a positive data, we can use the degrees of freedom in the pdf of this random variable to control the correlation and LRD property of modeling data \[1\][12].

Because of its simplicity and flexibility, we will use a symmetric beta distribution \( \beta(p,p) \) with pdf supported on [-1,1]

\[ f(a) = \frac{(1+a)^{p-1}(1-a)^{p-1}}{\beta(p,p)2^{2p-1}} \]

It can be shown \( f(a) \) exhibits different shapes with the change of values of \( p \), which is illustrated in figure 2 \[1\]. Its variance is given by

\[ \text{var}(A) = E[A^2] = \frac{1}{2p+1} \]

(22)

We already know the wavelet transform generates approximate decorrelated wavelet coefficients
for a LRD signal, so we can capture the correlation structure of modeling data by properly setting the energy of wavelet coefficients at each scale [1].

The energy of wavelet coefficients is computed by their second moment. With the symmetric structure of random variable A and wavelet coefficients at every scale with respect to origin, we know their means are both equal to zero. Hence we can find energy by calculating the variance.

To make our model begin with a little bit data, we should fix the energy at the coarsest scale and set ratios of energy for other scales. From (20) and (22), we have

$$E(W_{j0k}^2) = E(A_{j0k}^2)E(U_{j0k}^2) \Rightarrow \text{var}(W_{j0k}) = \frac{1}{2p_0 + 1} E(U_{j0k}^2)$$

then

$$p_0 = \frac{1}{2} \left( \frac{E(U_{j0k})}{\text{var}(W_{j0k})} - 1 \right)$$

(23)

Note that we cannot replace \(E(U_{j0k}^2)\) with \(\text{var}(U_{j0k})\) since the mean of \(U_{j0k}\) (scaling coefficients) is not equal to zero.

The energy ratios are calculated by

$$\eta_j = \frac{E(W_{j-1,k}^2)}{E(W_{j,k}^2)} = \frac{\text{Var}(W_{j-1,k})}{\text{Var}(W_{j,k})}$$

(24)

using (21), (22) and (24),

$$\eta_j = \frac{E(A_{j-1}^2)}{E(A_j^2)(1 + E[A_{j-1}^2])} = \frac{\text{Var}(A_{j-1})}{\text{Var}(A_j)(1 + \text{Var}(A_{j-1}))} = \frac{1/(2p_{j-1} + 1)}{(1/(2p_{j} + 1))(1 + 1/(2p_{j-1} + 1))}$$

Finally we get
\[ p_j = \frac{\eta_j}{2} \left( p_{j+1} + 1 \right)^{-1/2} \quad (25) \]

To summarize, at every level, we find \( p_j \) by \( \eta_j \) and \( p_{j-1} \), then by using the random number generator of beta function with parameter \( p_j \), we can generate some random numbers between \([-1,1]\) whose quantity is equal to the quantity of wavelet coefficients at \( j^{th} \) scale. When fitting the random numbers into equation (21), at the finest scale we get a series of scaling coefficients, which are just the realization of our model. We can see the algorithm only cost the time of \( O(N) \) to synthesize a \( N \)-point data.

### 3.3.3 Modeling The Root Scaling Coefficients

The theory of cascaded filters tells us the procedure we use to synthesize the data should begin from coarsest approximation (scaling coefficients) and add the details (wavelet coefficients) recursively. Therefore the job we still need to do thus far is to model the root scaling coefficients. If there are enough scales in wavelet transform, the scaling coefficients at coarsest scale can be seen as the sum of a large number of independent random variable. By Central Limit Theory we can expect the root scaling coefficients approximately follow Gaussian distribution [1], thus we can model them only through their mean and variance. For getting a better statistical information about the mean and variance, we have to leave enough coefficients at the coarsest scale.

However, negative data are unavoidable because of Gaussian distribution, which will ruin our whole model. Here the bottom line to this assumption is that the mean greatly outweighs the variance so that the probability of a negative value is very low.

We model the root scaling coefficients as follows: control the number of scales of wavelet transform such that there are enough scales used and enough data left at the coarsest scale (16 scales used and 15 scaling coefficients left in my simulation); compute the mean and variance of these data; generate the Gaussian random numbers with parameter mean and variance whose quantity is equal to the quantity of scaling coefficients; check and make sure all the data is positive;

### 4. Simulation and Results

#### 4.1 Real Trace Data

The real trace we used is pAug89 that recorded the first 1 million arrivals (about 3142.82 seconds) of the daylong trace started at 11:25 a.m., 29 August 1989 on the "purple cable (the nickname of an Ethernet cable at the Bellcore Morristown Research and Engineering facility,
building MRE-2). The actual accuracy is roughly 10 microseconds [10].

We chose interarrival times as our training data because interarrival times, being continuous-valued, are most natural for our Multifractal Wavelet Model. Besides, unlike the data of packet inter-arrival times, byte-per-time, the analysis of interarrival times avoids the problem of choosing the time interval [1].

We extract the interarrival times from pAug89 and plot them in figure 3 grouped by 100, 10 and 1 packet. In fact it is a process of zooming in.

The data plotted of the next subfigure correspond the last 10% of the last subfigure. That means, from left to right, (1:1000000) grouped by 100 packets, (900000:1000000) grouped by 10 packets and (990000:1000000) grouped by 1 packet are plotted, respectively.

From the figures we get that the interarrival times should fall into the category of fractal process because it is irregular and self-similar no matter how fine scale we see it.

4.2 Test Multifractal Scaling Behavior of Real Trace

Using equation (16) we plot the relation of \( j \sim \log_2 S_j(q) \) in figure 4 (a). We can see they are of linear relations. We also plot \( j \) against \( \log_2(S_j(q)) \) in figure 4 (b) to show the linear relation more clearly.

From figure 4, we can see the excellent scaling property at most scales (aggregation levels) that shows the real trace pAug89 is multifractal [11][1].

We note that the two figures in figure 4 are kind of different from the figure 8 of paper [1]. It is because the y axis in figure 4(a) is actually \( -\log_2(S_j(q)) \) that is consistent with equation (16). It does not affect the final result of partition function and multifractal spectra which will be seen soon.
Fig 4 Multifractal Scaling Behavior. (a) scale \( j \) versus \(-\log_2(S_j(q))\) (b) scale \( j \) versus \( \Delta (\log_2(S_j(q)))\)

Scale \( j \) from 1 to 19 and \( j=1 \) is the coarsest scale. In both figures, from top to bottom, each line represents different moments \( q = [2.5, 2.0, 1.5, 1.0, 0.8, 0.4, 0, -0.3, -0.5, -0.8, -1, -1.3] \). In figure (a), the right triangles represent computed values and the green lines through the right triangles are the fitted lines in the sense of least square.

4.3 Generating MWM Data With \( \beta \) Random Variable

Based upon the results presented in section 3.3 “Building MWM” (refer to the routine “GnerateMwmData.m”), we can generate one set of data by MWM. We repeat the routines twice due to the un-deterministic property of the algorithm. We denote the two realizations of MWM as “Generated Trace 1” and “Generated Trace 2” and pAug89 as “Real Trace”.

4.4 Test Model By Criterions

4.4.1 Positive Data

We plot generated traces grouped by 100, 10, and 1 packet in figure 5 with the same approach as the real trace.

It is clear that all the data generated by MWM is positive. It also has the irregular shape similar to real trace plotted in figure 1.

4.4.2 Long Range Dependence

We plot the relation of variance and time By equation (15) to test the LRD property of real trace and generated traces (refer the routine “PlotVarianceTime.m”) in figure 6.

From the relation of variance and time of real trace, we find interarrival times are not strict second-order scaling. However, we can conclude the generated traces are of LRD property. We
can see the generated traces capture much of the correlation structure except at coarser aggregation levels (m=12~16). All H is between [0.5,1] that means they exhibit LRD property.
4.4.3 Multifractal Scaling

The relation of \( j \sim -\log_2 S_j(q) \) and \( j \sim \Delta(-\log_2(S_j(q))) \), for generated traces, are shown in figure 7. We can see that the generated traces exhibit similar multifractal scaling behavior to real trace except the moment \( q \) strongly negative, which can been seen more clearly by partition functions \( T(q) \) in figure 8. The values of \( T(q) \) are calculated with the slope of these lines got from fitting \( j \sim -\log_2 S_j(q) \) to be linear in least square sense.

![Generated Trace 1](image1)

![Generated Trace 2](image2)

Fig 7 multifractal scaling Behavior of generated traces

Scale \( j \): from 1 to 19 and \( j=1 \) is the coarsest scale. In both figures, from top to bottom, each line represents different moments \( q \) =\{2.5, 2.0, 1.5, 1.0, 0.8, 0.4, 0, -0.3, -0.5, -0.8, -1\}. The right triangles represent computed values and the green lines through the right triangles are the fitted lines in the sense of least square.
We can see generated traces match real trace very well when moment $q$ is from $-0.5$ to $2$. Why does it happen? As we have discussed in section 3.2.2, when $q<0$ in partition function, we are trying to measure the data that is very small but nonzero that is of physical meaning. In real trace, we treat $10^{-5}$ as its zero due to the resolution of measure instruments, but for generated trace, we treat real zero as its zero because it is generated by Matlab. In fact we find the minimum order of generated trace is $10^{-12}$, which is the reason why more small values are found in generated trace, and why the partition functions of generated traces deviate more from uniform distribution when $q<0$. The mismatch when $q>2$ is because we did not modeling the scaling behavior of higher moments of real trace when building our MWM model. It will be improved when add these constraints into our model.

4.4.4 Mutifractal Spectra

Multifractal spectrum has similar interpretation with partition function because we compute them by partition functions with the help of Legendre transform. By equation (17), finding the tangent of $T(q)$ at every point, we plot the multifractal spectra of generated traces and real trace in figure 9.

From figure 9, we see that the spectra of generated traces closely match real trace when $f(\alpha)$ near to 1, which means the generated trace capture much of the most often occurred events. The similarity on the left of two spectra ($\alpha<1$) indicates generated traces also capture the property of burstiness of real trace. The divergence of the spectra of generated traces from real trace on the right indicates that the frequency of observing large $\alpha$ is kind of too high. In multifractal spectrum, larger $\alpha$ means more small increment, which corresponds to the part of $q<0$ in partition function, then we have the same interpretation as before.
4.5 A Simple Approach To Improve The Performance Of MWM

In [1] authors discussed about how to improve the mismatches mentioned above. They replace the beta distribution with point mass distribution with two degrees of freedom in those scales of mismatch. Rather than the complex method, we present a simple solution. The divergence of spectra of generated traces from real trace is mainly because the data resolution of generated traces is smaller than the one of real trace. If we truncate all generated data to the resolution of real trace, we get a set of data with the same resolution as real trace. We used the same criterions to test the truncated data and the simulation results are shown in figures 10.

From figure 10, we can see after the data of generated trace are truncated to the same resolution as real trace, the LRD property has almost no change, but the multifractal scaling, partition function and multifractal spectrum have the improvement to some extent. Especially when q<0 in partition function and on the right of spectrum, the divergence from real trace is greatly reduced.

Therefore, it can be used as a simple way to improve the performance of MWM by truncating the generated data to the same resolution as the real trace before we use generated data to the application.
Fig 10 Simulation Results For Truncated Data Of Generated Trace 2
5. Conclusion

Based on performance criterions, we develop a Multifractal Wavelet Model (MWM) and show its effectiveness and efficiency by simulation results.

The Multifractal Wavelet Model can capture most of the properties of network traffic, such as self-similarity, burstiness Long-Range dependence (LRD) and multiscaling behavior, which greatly overweighs other classical stochastic models.

By using the Haar wavelet transform, the total cost of building the model is O(N) for an N-point output, which is more suitable to be used to online applications, such as some ongoing research “distributed storage and retrieval for multimedia [4]”.

6. Future Work

As it has been shown in section 3.1, queuing behavior is one of important performance criterions of network traffic models. It should be added into our simulation to show if MWM works well.

From the mismatches of multifractal spectra and partition functions between generated traces and real trace (figure 8 and figure 9), we see we need to add more constraints to control the behavior of generated trace when moments are beyond –0.5 and 2.

As we know, network traffic behaves differently in different time and places. It is necessary to test the effectiveness of MWM with traffic trace of nowadays because the trace we used in this paper is kind of old even though it is typical.

The low computational complexity of MWM provides some hope to real time applications, such as the online network delay prediction based on network traffic model, which is what we are continue working on.

7. References

Communications”, An NESC Strategic Grant Application, unsubmitted.


