INFORMATIONAL THEORY APPROACH
FOR STUDYING SELF-SIMILAR
TRAFFIC

A. Asars, E. Petersons
Riga Technical University
Riga, Latvia
a.asars@inbox.lv

Existing analytical models are not adequate for a wide class of traffic scenarios in modern telecommunication networks with self-similar data flows. For purpose to define which traffic distribution should be utilized when modelling networks with these traffic patterns maximum entropy method is applied to derive unique solution for packet arrivals distribution. For testing self-similar properties of obtained distribution simulation results have been used.

Keywords: self-similar traffic, maximum entropy method.

1. INTRODUCTION

Data traffic observing in most communication networks exhibits self-similar or fractal properties. However network equipment parameters chosen in developing (implementing), was estimated using traditional (Poisson) traffic model, therefore whole transmission performance is decreased [1, 2, 3, 4].

The main motivation for our work is the fact that existing traffic models are not adequate for a wide class of traffic scenarios in modern telecommunication networks. In these networks, many independent sources carrying different media and content are multiplexed. Under strict performance requirements and with moderate buffer sizes, the utilization of the network is expected to be low, allowing high variability in the traffic. This type of traffic exhibits often long range dependence, which is caused by many possible reasons, such as multiplexing of multiple short-correlated sources. Recently proposed models are addressing the issues of self-similarity from two main directions in an ineffective way:

1) simple self-similar models, such as those based on the Fractional Gaussian Noise process, are proposed to model self-similarity, but are unable to model heavy-tailed traffic or traffic that has skewed marginal distribution;
2) heavy-tailed models, such as those based on the Pareto distribution or the $\alpha$-stable distribution are also proposed, but even though they are useful in modeling bursty traffic they are not useful in modeling self-similarity.

In practice, it has been noticed that network equipment do not perform in the field as predicted [5], and this is an indication of inaccurate traffic behavior assumptions. Inaccurate prediction of required bandwidths and buffer sizes leads either to low quality of service,
which becomes more perceived to the user when the losses are correlated, or to underutilization of the network resources that reduces profits for the network owners and operators.

The above mentioned conditions of heavy traffic, high variability and long-range dependence, require developing of new techniques of estimation network equipment parameters operating in environment with self-similar traffic. For purposes to derive distribution that capture desired properties maximum entropy approach is presented. This method is based on concept of entropy functional, introduced in information theory by Shannon [6]. Since maximum entropy corresponds to maximum disorder in the system, these solutions are the least biased of all solutions that satisfy the system constraints. More importantly maximum entropy solutions have been shown to be analytically tractable and can be utilized in queuing models.

In the following sections analysis of self-similar flow of arrivals is presented and generalized maximum entropy approach is directly applied to derive a unique distribution.

2. MAXIMUM ENTROPY SOLUTION FOR SELF-SIMILAR DATA TRAFFIC

In this section a generalized maximum entropy model is used to derive a unique solution for self-similar data flow. Consider situation, when someone observes data traffic on input interface of network equipment and records time intervals between packets. Our task to define a distribution function of time intervals between arrivals \( f(x) \), that in the best way fits to the collected statistics. We divided this task in two subtasks, those are to obtain \( f(x) \) if from measurements results is known:

1) mean value of interpackets time intervals;
2) variance of interpackets time intervals.

The maximum entropy solution for case A is formed by maximizing (1) subject to the constraints (2)–(3):

\[
H[f(x)] = - \sum_x f(x) \ln[f(x)],
\]

\[
\int f(x)dx = 1,
\]

\[
\int xf(x)dx = \bar{x}.
\]

The maximization of (1) can be carried out using Lagrange's method of underdetermined multipliers leading to solution

\[
f(x) = \frac{2\bar{x}}{x^2}.
\]

The maximum entropy solution for case B is formed by maximizing (1) subject to the constraints:

\[
\int f(x)dx = 1,
\]

\[
\int x^2 f(x)dx = \sigma.
\]
In similar way as above, Lagrange's method of underdetermined multipliers was used to derive following equation

\[ f(x) = 2 \sqrt{\frac{\pi}{\sigma}} \exp \left\{ -\frac{1}{4\sigma} x^2 \right\}. \]  \hspace{1cm} (7)

3. RESULTS AND ANALYSIS

To validate in Section 2 obtained results, consider that measured traffic is distributed according Pareto law. The Pareto distribution was chosen because it is one of the most popular distributions used to model the bursty network traffic. The Pareto distribution (8) is a power curve with two parameters, namely shape parameter \( \alpha \) and lower cutoff parameter \( k \) \((k, \alpha < 0)\)

\[ f(x) = \frac{\alpha}{k} \left( \frac{k}{x} \right)^{\alpha+1}. \]  \hspace{1cm} (8)

The parameter \( \alpha \) determines the mean and the variance in the following way:

- for \( \alpha < 1 \) the distribution has infinite mean;
- for \( 1 < \alpha < 2 \) the distribution has finite mean and infinite variance;
- for \( \alpha < 2 \) the distribution has infinite variance.

So in this case, measured traffic mean is equal to

\[ E(X) = \frac{\alpha}{\alpha - 1} k, (\alpha > 1). \]  \hspace{1cm} (9)

Since, the maximum entropy solution according (4) is:

\[ f(x) = \frac{2\alpha x_0}{x^2(\alpha - 1)}. \]  \hspace{1cm} (10)

For testing self-similar properties of distribution (10) simulation in which packet was generated according to this equation had been used. One of the most popular methods typically used for verifying self-similarity is aggregated variance plot. This method visualizes the variance of the self-similar process. The logarithm of the variance of an aggregated self-similar process, \( X_t \), decreases linearly with the logarithm of the aggregation size \([1], m\):

\[ \log[Var(X_t^{(m)})] = \log[Var(X_t)] - \beta \log[m]. \]  \hspace{1cm} (11)

A log-log plot which fits a line to the aggregated variance points allows us to verify the self-similar nature of the process, as well as estimate the degree of self-similarity \( H = 1 - \beta/2 \), where \( \beta \) is the slope of the line. The degree of self-similarity of a process is typically described by the Hurst parameter \( (H) \) \([1]\). \( H \) is between 0.5 and 1, where 0.5 represents non self-similar behavior and the closer to 1, the more long-range dependent the process is. From the figure 1(a) it is clear that the slope of the variance plot line formed by the generated data is much less than 1, therefore \( H > 0.5 \) for this data, i.e., the process is self-similar. To achieve more precise results curve in figure 1(a) was approximated by linear regression and value was estimated. So, for maximum entropy distribution (4) parameter \( H = 0.7842 \). For comparing results similar procedure was performed with Pareto
Fig 1  Simulation results for maximum entropy (a) and Pareto (b) distributions

Fig 2  Simulation results for maximum entropy distribution with $\alpha = 2; 3, 4$

Fig 3  Hurst parameter dependence on $\alpha$ (a) and variance (b) for maximum entropy distribution
distribution: packet traffic was simulated and parameter $H$ estimated (figure 1(b)). For this distribution parameter $H = 0.945$.

Studying function (4) behavior depending on parameter $\alpha$, can see that Hurst parameter doesn’t change its value at different values of parameter $\alpha$. Simulation showed that at $\alpha = 2$, $\alpha = 3$ and $\alpha = 4$ appropriate Hurst parameters were $H = 0.83, 0.85, 0.83$. Graphically these results are represented in figure 2 and in more detailed manner in figure 3.

From above mentioned results can be concluded that analytically obtained function (4) fits to self-similar properties and could be used for studying and modeling performance of data network elements with self-similar input traffic.

4. CONCLUSIONS

Future high-speed networks will carry traffic from many different sources, due to convergence of voice and data into a single backbone medium and the emergence of new applications. They will also provide different levels of service depending on the application and the demands of the customer. At the same time, the investment and the operation costs should be economically viable for companies and customers. Therefore, good engineering design and prediction of their performance relative to the cost and quality demands are of paramount importance.

Under heavy traffic conditions and with strict quality of service requirements, the mean utilization of backbone networks is expected to be low, which will allow for high burstiness. Strong correlation, existent in many time scales, is also observed often in the traffic. There are many possible causes for the existence of long correlations, such as multiplexing of short-range correlated sources or the existence of on/off types of sources with long-tailed distributed on and off durations. These two unique properties, namely burstiness and long-range dependence, were not taken into consideration in most traditional models even though they are of great importance in performance prediction.

We introduced a new model for heavy-traffic modeling based on the Maximum Entropy approach. The significance of the maximum entropy analysis is that it allows to build closed form analytical expression for arrival distribution self-similar data flow. In further work is planned to use more constraints, derived from physical backgrounds of self-similar processes.

Currently, no other model has such desirable properties for modeling aggregated traffic in high-speed networks. There are many areas that the work presented in this paper can be extended in terms of theory and applications. For example effective bandwidth extensions for heavy-tailed processes: An interesting topic for research is the extension of the effective bandwidth theory to include allocation schemes for long-tailed sources. The general assumption is that the overflow probability of a large buffer always decreases asymptotically in an exponential fashion with the buffer size. What would happen if we do not assume exponential decay but some other decay with a more general function, such as a hyperbolic function? Given a QoS description can we find a general bandwidth allocation scheme in this case?
We hope that the proposed model can be employed to answer these questions and could be successfully used to simulate self-similar traffic, thus enable network expert to analyze traffic loads, predict and avoid congestion, and test offered quality of service.

REFERENCES


