From Poisson Processes to Self-Similarity: a Survey of Network Traffic Models

Michela Becchi, *mbecchi@wustl.edu*

Abstract
The paper provides a survey of network traffic models. It starts from the description of the Poisson model, born in the context of telephony, and highlights the main reasons for its inadequacy to describe data traffic in LANs and WANs. It then details two models which have been conceived to overcome the Poisson model's limitations. In particular, the discussion focuses on the packet train model, validated in a Token Ring LAN, and on the self-similar model, used to capture traffic burstiness at several times scales in both Ethernet LANs and WANs. The discussion closes with some examples of usage of those models in LAN and WAN environments.

Keywords: Traffic models, Poisson processes, stochastic processes, compound processes, renewal processes, packet trains, self-similarity, fractals.

Table of Contents

1. Introduction

One important research area in the context of networking focuses on developing traffic models which can be applied to the Internet and, more generally, to any communication network. The interest towards such models is two-fold. First, traffic models are needed as input in network simulations. In turn, these simulations must be performed in order to study and validate algorithms and protocols to be applied to real traffic, and to analyze how traffic reacts to particular network conditions (e.g.: congestion, etc.). Thus, it is essential that the assumed models reflect as much as possible the relevant characteristics of the traffic it is supposed to represent. Second, a good traffic model may lead to a better understanding of the characteristics of the network traffic itself. This, in turn, can help designing routers and devices which handle network traffic. If, for instance, a model which has been well validated shows some correlation between traffic arrivals, this information can be used in order to conceive ad hoc packet handling strategies.

The first traffic model, based on Poisson processes, was born in the context of telephony, where call arrivals could be considered independent and identically distributed and "holding times" followed an exponential distribution. Although initially successful and analytically simple, the Poisson model has proven not suitable to describe data traffic in modern LANs and WANs, where batch arrivals, event correlations and traffic burstiness are important factors. The use of heavy tailed distributions and of self-similarity has become more and more predominant.

The goal of this paper is to point out the most important concepts at the core of the basic traffic models in use, and to show how these models are applied to LANs and WANs.

The remainder of the paper is organized as follows. Section 2 introduces basic concepts about traffic modeling. Section 3 describes the Poisson model and its limitations. Section 4 presents the Packet trains model, intended to overcome some of the Poisson model's limitations and to capture correlation and locality among packet arrivals. Section 5 introduces a mathematical description of self-similarity and explains how the self-similar model differs from traditional ones and allows capturing traffic burstiness at different time scales. Section 6 lists other traffic models in used. Sections 7 show how the models presented in the paper have been applied to traffic in LANs and WANs. Finally, section 8 closes the discussion with concluding remarks.

Back to Table of Contents

---

2. Traffic modeling: basic concepts

Internet traffic can be modeled as a sequence of arrivals of discrete entities, such as packets, cells, etc. Mathematically, this leads to the usage of two equivalent representations: *counting processes* and *interarrival time processes*. A counting process $\{N(t)\}_{t=0..\infty}$ is a continuous-time, integer-valued stochastic process, where N(t) expresses the number of arrivals in the time interval *(0,t]*. An interarrival time process is a non-negative random sequence $\{A_n\}$, where $A_n=T_n-T_{n-1}$ indicates the length of the interval separating arrivals *n-1* and *n*. The two kind of processes are related through the following equation:

$$\{N(t) = n\}=\{T_n \leq t < T_{n+1}\} = \left\{\sum_{k=1}^{n} A_k \leq t < \sum_{k=1}^{n+1} A_k \right\} \quad (1)$$

In case of ***compound*** traffic, arrivals may happen in *batches*, that is, several arrivals can happen at the same instant T_n. This fact can be modeled by using an additional non-negative random sequence $\{B_n\}_{n=1..\infty}$, where B_n is the cardinality of the *n-th* batch. The traffic model is largely defined by the nature of the stochastic processes {N(t)} and $\{A_n\}$ chosen, which will be analyzed in the remainder of this paper.

One important issue in the selection of the stochastic process is its ability to describe traffic ***burstiness***. In particular, a sequence of arrival times will be bursty if the T_n tend to form clusters, that is, if the corresponding $\{A_n\}$ sees a mix of relatively long and short interarrival times. Mathematically speaking, traffic burstiness is related to short-terms autocorrelations between the interarrival times. However, there is not a single widely accepted notion of burstiness [frost94traffic]; instead, several different measures are used, some of which ignore the effect of second order properties of the traffic. A first measure is the ratio of peak rate to mean rate, and has the drawback of being dependent upon the interval used to measure the rate. A second measure is the coefficient of variation $c_A=\sigma[A_n]/E[A_n]$ of the interarrival times. A metric considering second order properties of the traffic is the index of dispersion for counts (IDC). In particular, given an interval of time τ, $IDC(\tau)=Var[N(\tau)]/E[N(\tau)]$. Because of the relationship in Eq. (1), IDC includes in the numerator the effects of the autocorrelation between the A_n. Finally, as will be better detailed later, the Hurst parameter can be used as a measure of burstiness in case of self similar traffic.

Back to Table of Contents

---

3. The Poisson Model

The Poisson model is the oldest traffic model in use. Introduced in the context of telephony by A. K. Erlang, it shows some limitations when applied to Internet data traffic. In this section we first characterize the model and then point out the issues which make suitable the use of different frameworks.

3.1 Description of the model

Traffic is characterized by assuming that the packet arrivals A_n have the following characteristics:

1. they are *independent*.
2. they are *exponentially distributed* with rate parameter λ: $P\{A_n \leq t\} = 1-e^{-\lambda t}$.

Alternatively, this implies describing the traffic through a counting process satisfying the equation $P\{N(t)=n\} = e^{-\lambda t}(\lambda t)^n/n!$, where $N(t)$ is the number of arrivals at time t.
Poisson processes exhibit the following important analytical properties:

1. The superposition of independent Poisson processes with rates λ_1, λ_2, ... λ_n results in a new Poisson process with rate $\lambda_1+\lambda_2+...\lambda_n$.
2. The number of arrivals in disjoint intervals is statistically independent. This property is also referred to as *independent increments* property, and makes Poisson a *memoryless* process.
3. For an exponential distribution with parameter λ, not only the mean, but also the variance is equal to λ. This leads to a unitary coefficient of variation.
4. According to the Palm-Khintchine theorem, the multiplexing of independent traffic streams approximates a Poisson process if: (i) the traffic streams can be modeled as renewal processes (that is, interarrival times are independent and identically distributed), (ii) as the number of streams increases the individual rates decrease so as to keep the aggregate rate constant.

There are several ways to verify whether a particular arrival process is Poisson [jain86train]. An easy visual way consists in plotting the histogram of the interarrival times and verifying whether it is an exponentially decreasing function.
Alternatively, observing that if $p(t) = \lambda e^{-\lambda t}$ then $log(p(t)) = log(\lambda) - \lambda t$, one can plot the log histogram and check whether it is a linear function. In this case, the rate λ of the process can be easily extrapolated from the intersection with the y-axis (or from the slope of the line).

One special case of the Poisson model is represented by *time-dependent* Poisson processes. This representation is suitable for situations where the mean rate is not constant: in such cases, the rate parameter λ is expressed as a function of the time $\lambda(t)$.

3.2 Traffic burstiness: the limitations of the Poisson model

One basic limitation of the Poisson model is its inability to capture traffic burstiness which characterizes data traffic (as opposed to voice traffic in old telephone systems). Analytically, this can be explained as follows. In any renewal traffic process the autocorrelation function of the $\{A_n\}$ vanishes identically. But, as mentioned above, positive autocorrelation between the $\{A_n\}$ can explain, to a large extent, traffic burstiness. Thus, Poisson is not the appropriate model in case of bursty traffic, especially when traffic burstiness happens on multiple time scales.

Let us illustrate and analyze this concept more in depth. A way to have an intuition about traffic burstiness is to define it in terms of a *time scale* over which bursts occur. If, for instance, we consider a Poisson process with rate λ (e.g.: 100/sec), then the time scale for burstiness is 1/λ (e.g.: 10 msec), and periods of over- or lower-than average activity over much smaller or much larger time scales occur with rapidly decreasing probability. This is in general true in the case of telephone traffic, which is well modeled through Poisson processes. However, as will be detailed later, traffic bursts in data networks tend to happen on many different time scales [paxson95wide], [leland91high], [willinger98selfsimilarity], [leland93selfsimilar], [riedi00toward], which does not fit the Poisson model.

Figure 1, taken from [willinger98where], illustrates the failure of the Poisson model in capturing internet traffic burstiness. The plots on the right hand side represent the trace of traffic arrivals registered in 1995 on a network link connecting a large corporation to the Internet. The plots on the left hand side are obtained by fitting a simple Poisson-based model to the mean and variance of the measured samples. The different rows show distinct time scales: moving from one row to the subsequent the time scale is increased by a factor of 10. The black regions illustrate the area expanded in the previous row. Each point in the first row represents the number of packets during a 100 msec interval, in the second row the number of packets in a 1 sec interval, and so on. Notice that the scale on the y-axis is also varied. As can be observed, the Poisson traffic tends to become smoother as the time scale increases, whereas the original traffic is characterized by a bursty behavior on all the time scales.