ISSN: 2776-0960 Volume 3, Issue 5 May, 2022

FRACTAL PROPERTIES OF NETWORK TRAFFIC IN A CLIENT-SERVER INFORMATION SYSTEM

Rakhmatullaev Doston
Tashkent University of Information Technologies named after Muhammad al-Khorazmiy, Uzbekistan

Abstract

This paper uses three statistical analyses methods, R/S statistic, variance-time plots and periodogram-based analysis, to estimate the Hurst parameter of the wireless network traffic. The fractal statistical characteristics of wireless network traffic are studied based on the Hurst parameter. Simulation results compare the parameter estimation values of fractal wireless network traffic by utilizing above three statistical approaches. The simulation results demonstrate that wireless network traffic exhibits fractal characteristics. These results can be very useful for the performance improvement and the system design optimization of wireless network.

Key words. teletraphic, exponentiality, traffic, SQL, fractal, modeling, investigation, probabilistic.

Introduction

The last decade has been marked by a significant achievement in the field of teletraphic theory – the discovery of self-similarity, or fractal properties of processes occurring in data transmission networks. There have been studies that revealed fractal properties of traffic generated by a variety of applications – video conferencing, file-server applications, access to Web resources, etc. At the same time, when optimizing the probabilistic-temporal characteristics of client-server information systems (IS), they most often proceed from the assumption of the exponentiality of event flows at all levels, including in the network.

The conducted research shows the presence in the network traffic of real SQL server applications of all the most important properties of self-similarity, which have a significant impact on the probabilistic characteristics of the system.

The concept and properties of self-similar traffic

Informally, a self-similar (fractal) process can be defined as a random process whose statistical characteristics exhibit scaling properties. A self-similar process does not significantly change its appearance when viewed at different scales on the time scale. In particular, unlike processes that do not have fractal properties, there is no rapid «smoothing» of the process when averaged over a time scale – the process retains a tendency to spikes.

When modeling network traffic, the values of Xk are interpreted as the number of packets (less often – as the total amount of data in bytes) that entered the channel or network during the k-th time interval. The initial process is already averaged. In some cases, when there is a need to avoid such an initial averaging, a point process or a stream of events is considered, i.e. a sequence of moments when single packets arrive in the network.

Long-term dependence is the cause of pronounced pulsations of the process, however, it allows us to talk about some predictability within a short time. From the point of view of queue theory, an important consequence of flow correlation is the unacceptability of estimates of queue parameters based on the assumption of the same and independent distribution of intervals in the incoming stream.

2. Slowly decreasing variance.

When averaging the process, the variance of the sample average decays more slowly than the inverse of the sample size, according to the law:

while for traditional stationary random processes

, i.e. decreases inversely proportional to the sample size.

The property of slowly decreasing variance suggests the possibility of significant «outliers» not smoothed by averaging in a random process, and connects self-similarity with such a concept as distributions with weighty tails. An important consequence of the property of slowly decaying variance is that in the case of classical statistical tests (for example, calculating confidence intervals), the generally accepted standard deviation measure is erroneous.

The most significant feature of a random variable having a distribution with a weighty tail is extreme variability. With a probability that is not negligible, a number of «very large» values may be present in the sample. Such distributions significantly reduce the accuracy of statistical estimates; for example, the final sample size leads to an underestimation of the mean and variance.

The presence of RVX in phenomena external to the processes under consideration is one of the reasons for the occurrence of self-similarity in the corresponding stochastic models.

Often, when considering self-similar processes, they talk about a complex of interrelated concepts: self-similarity, scaling, long-term dependence, RVX and power laws of statistical characteristics. This set of properties distinguishes processes called self-similar from classical random processes, for example, Poisson processes.

Traffic under investigation

Numerous measurements show the presence of substantially self-similar traffic properties in client–server systems of various architectures - from «classic» two-link, to multi-level with Web access and terminal [8, 9].

The fractal nature can be considered by the example of a traffic slice (Fig.1) obtained when working with a remote unit from several workstations with a DBMS server. The parameters of the traffic slice are shown in Table 1.

The data was obtained by intercepting frames on the FastEthernet interface of the DBMS server using the tcpdump program. Based on the assumption of the duplexity of the channel, traffic of one direction is accepted for consideration – outgoing in relation to workstations. The DBMS server is the Oracle8i server; the application (the billing system of the telecom operator) is implemented according to the classical two-link client-server scheme, i.e. network interaction takes place on the basis of TNS / SQL*net over TCP. Since the bandwidth of the channel under consideration (100 Mbit/s FastEthernet) significantly exceeds the total traffic, and the third–party load on the day of research is negligible, the slice can be considered «free traffic», in the terminology introduced by I. Norros [5] - i.e., traffic completely determined by its source and not affected by the network.

Figure 1 clearly demonstrates the highly pulsating nature of traffic with significant variance, the presence of sharp spikes, grouping into «bundles».

Table 1.

Date	26.11.2003
tcpdump output file	s24-26nov.td
Duration, s	21 936
Duration, h	6.093
Number of packages	688108
Intensity λ,s-1	31.368
Average packet size, bytes	193.2
Jobs	25
Hearst Parameter (IDC)	0.729
Hearstan parameter (autocorrelation)	0.724

ResearchJet Journal of Analysis and Inventions
https://reserchjet.academiascience.org

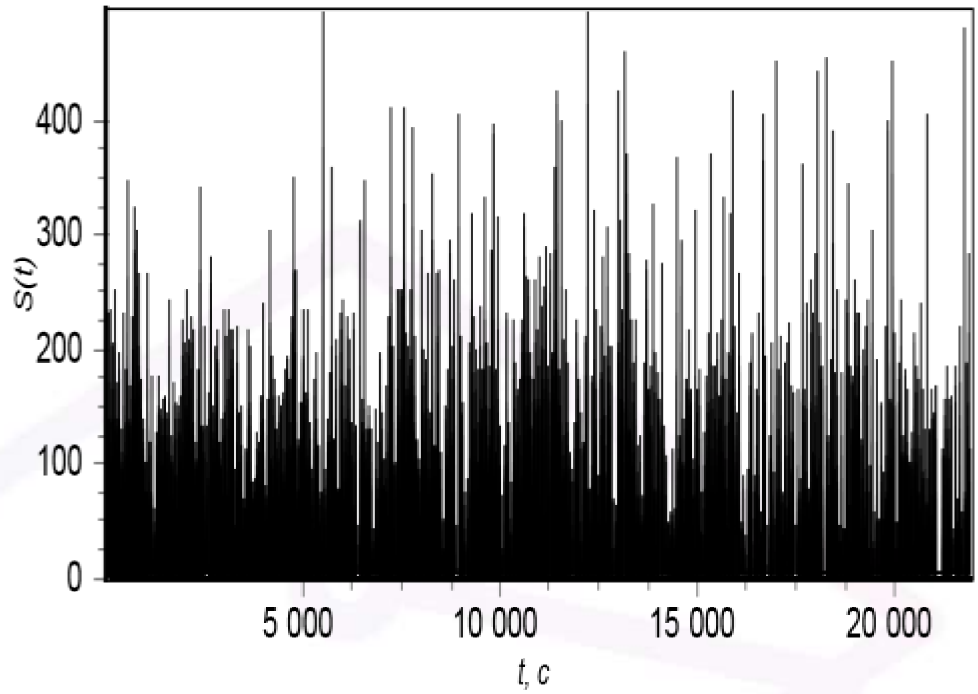

Fig. 1. Traffic intensity, packets per second.

Correlation Structure

The graphical representation of the correlation coefficient allows you to visually verify that the traffic under study has a long-term dependence.

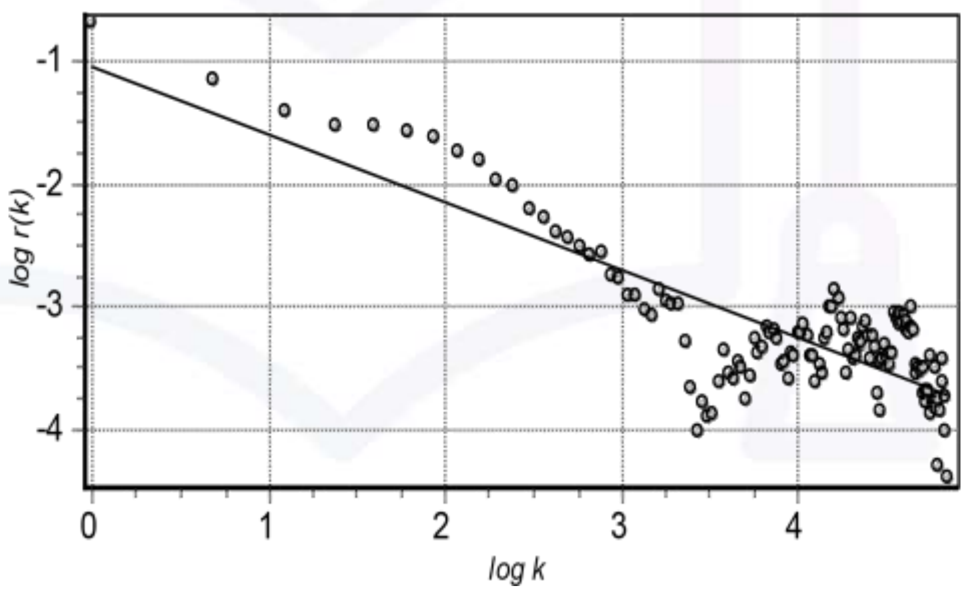

Fig.2 -Correlation coefficient values (logarithmic scale)

For comparison, a curve corresponding to the values of the correlation coefficient of a strictly self-similar process (2) is given. It is obvious that the sample estimates correspond quite accurately to the «ideal» curve, especially when the argument k is increased.

The essential role of long-term dependence in the initial process can also be identified based on the analysis of the so-called «mixed» process. Such a process is obtained from the original one by rearranging the intervals between packet arrivals in random order. In this process, the correlation decreases significantly faster, rather quickly tending to zero.

Hearst Parameter

Traditionally, self-similarity in a stochastic process is revealed by determining the Hurst parameter N. The fact that 0.5<H<1, i.e. the value of the Hurst parameter is different from 0.5, is considered sufficient reason to recognize the process as self-similar (at least asymptotically). It should be noted that the value of H, close to unity, may mean that the process is deterministic, i.e. not random: for a number of strictly deterministic processes, the structure is strictly repeated at any scale, which leads to a single value of the Hurst parameter.

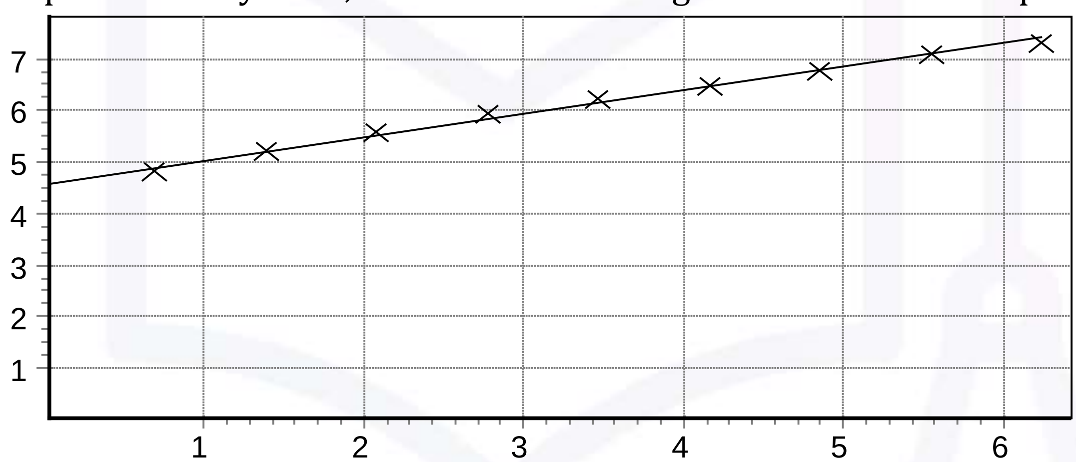

Fig. 3.-Determination of the Hurst parameter H by the values of IDC, or the Fano factor

In this case, the values of the Hurst parameter determined from the type of the correlation coefficient curve and by analyzing the IDC (0.724 and 0.729, respectively) are practically the same. At the same time, the value indicates significantly pronounced fractal properties.

Figure 3 shows the values of ln[F(T)-1] depending on lnT. According to (7), the points form a straight line that allows unambiguously determining the Hurst parameter. The straight line obtained by linear regression has an angular coefficient of 0.459, which corresponds to H=0.729.

For a process with randomly mixed, the value is significantly less: H=0.563. This fact, together with the type of correlation structure of the original and «mixed» processes, suggests that for the traffic in question, self–similarity is not so much in the «heavy» distribution of intervals, as in the long-term dependence - grouping of short intervals in bundles.

Queue Modeling

Self-similar load in queuing systems (QMS) leads to significantly greater delays at the same load value than the «classic» load [2, 7, 8, 9], due to the strong grouping of applications and significant variance.

Simulation modeling of a CFR with an unlimited buffer and an input stream, which is determined by the traffic in question, shows the queue behavior characteristic of fractal traffic. At the same time, as the simulation results show, the increase in the queue and the delay time is influenced not so much by the distribution of intervals as by the correlation structure of the process, i.e. long-term dependence. Randomly mixed traffic, while maintaining the distribution of intervals between arrivals and packet volumes, demonstrates a significantly less sharp increase in the queue.

In conclusion, it should be noted that for a full description of traffic in client–server ICS, not only an analysis of statistical characteristics is required, but also the choice of an adequate mathematical model - a self-similar process. Fractal shot process (FSNDP) [6, 8, 9] or α-stable fractal processes can act as such a model.

Literature

1. Leland W.E., Taqqu M.S., Willinger W. and Wilson D.W. On the self-similar nature of Ethernet traffic // Computer Communications Review, 1993:23, p.193-193.
2. Shelukhin O.I., Tenyakshev A.M., Osin A.V. Fractal processes in telecommunications. Monograph / Edited by O.I.Shelukhin - M.: Radio Engineering, 2003.- 480 p.
3. Tsybakov B.S. Model of teletraphy based on a self-similar random process. // «Radio Engineering», 1999, No. 5, pp. 24-31.

4. Neiman V.I. A new direction in the theory of teletraphy. // «Telecommunication», 1998, No. 7, pp. 27-30.
5. Norros I. The Management of Large Flows of Connectionless Traffic on the Basis of Self-Similar Modeling // ICC '95, IEEE International Conference on Communications. – Seattle, 1995.
6. Ryu B.K. Fractal Network Traffic: From Understanding to Implications. Ph.D. thesis. - Columbia University, 1996. – 143 p.
7. M.N.Petrov, D.Y. Ponomarev. Self-similarity in queuing systems with a limited buffer. // «Telecommunication», 2002, No. 2, pp. 35-39
8. Sokolov D.E., Trenogin N.G. The nature of network traffic on the client side of a distributed client-server system. // Materials of the International Scientific and Technical Conference «Informatics and problems of telecommunications». Novosibirsk: SibGUTI, 2001, pp.34-35.
9. Sokolov D.E. Load modeling in client-server systems based on fractal processes. // Materials of the interuniversity scientific and technical conference «Control and computing systems. New technologies». – Vologda, 2001, pp.59-60.
10. Terekhov V.I., Sokolov D.E., Trenogin N.G. Monitoring and analysis of traffic in IP networks: an integrated approach. // Materials of the International Scientific and Technical Conference «Informatics and problems of telecommunications». – Novosibirsk: SibGUTI, 2001, p.36.