# 1/f noise in computer network traffic



1994 J. Phys. A: Math. Gen. 27 L417

(http://iopscience.iop.org/0305-4470/27/12/004)

J. Phys. A: Math. Gen. 27 (1994) L417–L421. Printed in the UK

LETTER TO THE EDITOR

# $1/f$ noise in computer network traffic

István Csabai

Department of Atomic Physics, Eötvös University, Budapest, Puskin u. 5-7, H-1088, Hungary

Received 22 April 1994

**Abstract.** A huge number of computers connected together form an international network (Internet). As a result of their communication, trains of data packages travel through this network; due to the limited transfer rate and the behaviour of data source agents, the traffic fluctuates. Our measurements show that the power spectrum of the round-trip time between two points on Internet is $1/f$-like over a broad range. Models of collective phenomena, such as highway traffic models could be appropriate for describing this behaviour.

The main goal of this paper is to present our first results on Internet traffic measurements. Beside this, we would also like to give a rough picture of Internet, and to call to researchers' attention the close analogy between the basics of data traffic on Internet and vehicle traffic on highways.

The international network of computers (Internet) has gradually reached a size where methods of statistical physics, particularly of complex systems, can play a role. Up to October 1993 more than $2 \times 10^6$ hosts have been connected to Internet [1], and this number seems to double in slightly more than a year. The topology of the network connecting these machines is not a regular one, however, it can be regarded as a hierarchical, tree-like structure, although some loops can also occur. Hosts use this network to send and get data we use in our everyday work: computer mail, data transfer (FTP), newsgroups, remote logins, etc. At the vortices of the network there are special switches, called gateways, routers, bridges and nameservers. They help to route the data packets to their destinations. These devices have significant autonomy in finding the optimal way from source to destination, breaking up the messages into smaller datagrams, buffering the packets, etc. The interaction of nodes can produce collective phenomena which cannot be foreseen from an examination of individual parts of the system—statistical physics is often used in the understanding of such complex systems [2].

Let us consider a connection between two points on Internet, for example, I would like to transfer a file from a distant FTP site to my computer. We will not discuss the routing process and we suppose that the route connecting the two machines is not changing during this time, although this is not always true due to dynamical optimization of the connection speed [3]. This connection is not galvanic as in an old telephone system, where wires are directly connected at the crossbars. Messages go from gateway to gateway (figure 1). Since the lines connecting the gateways have different speeds and they are often not capable of transfering the data at the rate it is arriving at the gateway, the latter has to buffer the information. The capacity of the buffers is not infinite, so they can overflow and the gateway can get congested, so that the connection slows down, and data packets are 'dropped'.

0305-4470/94/120417+05$19.50  

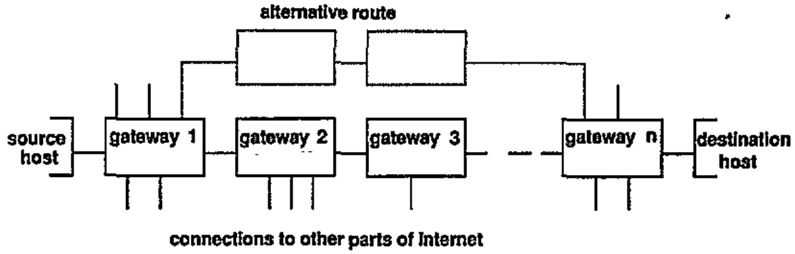


**Figure 1.** Schematic picture of a connection via Internet. Transfered data jumps from gateway to gateway, and can wait there if the output rate is not large enough.

To avoid the collapse of Internet several congestion control mechanisms have been suggested [3–5], such as source quench, random drop, fair queueing, etc. We will shortly discuss the most widely used source quench gateway protocol and the slow-start end-system congestion control policy [4]. If a gateway has to discard a datagram because of lack of buffer space, it sends the data back to its source, a so-called source-quench message. A destination host can send similar messages, if it is not fast enough to process the arriving data. This message is a request to the source to cut back the rate at which it is sending data. The strategy of the source host is the slow-start mechanism. It gradually increases its demand until a source-quench message arrives. In response to this the source host decreases the load and then begins an exploratory increase again. The hope is that this cycle will keep the total demand around the optimal level.

To investigate this we have performed measurements on the speed of the network between two points. We have measured the so-called round-trip time (RTT) between a workstation at Eötvös University, Hungary (*hercules.elte.hu*) and a distant FTP site (*funet.fi*) in Finland. The two machines can be regarded as distant, not geographically, but in network metrics. There are 15 gateways between the two hosts, and this number is not usually larger for a host in a physically more distant place, e.g. in Australia. We used the *ping* command of the Unix operating system, which sends a short datagram every second to the given destination, and reports the return time measured in milliseconds. Our method was able to collect RTTs shorter than 2000 ms. This RTT value can be regarded as the reciprocal of the actual transfer speed, measured in datagrams/second, so we get a time series of inverse speed values. The data were collected each second during a two week period. The total number of RTTs was 821 383.

The changing of RTT in time can be seen in figure 2. The data used in this plot is averaged over a 10 minute time period because of the resolution of the graphics. Since there is only one hour difference in time zone between the two endpoints, one can see the daily and even the weekly periodicity of the network speed.

We have tried some simple methods to analyse this time series. The embedding method [7] shows that there is no simple low-dimensional (up to seventh dimension) chaotic attractor behind the behaviour. To avoid possible errors due to missing datagram RTTs from our measurement (22 percent of datagrams had not returned), we have tried different methods [8] to calculate the power spectrum. All the methods gave nearly the same curve. Apart from the daily periodicity, the power spectrum of data shows $1/f$-like behaviour in the whole time domain, figure 3. The fit for data (without the daily periodicity peak) gives a slope of $-1.15$.

The description given in the previous paragraphs: packets running along the edges of a

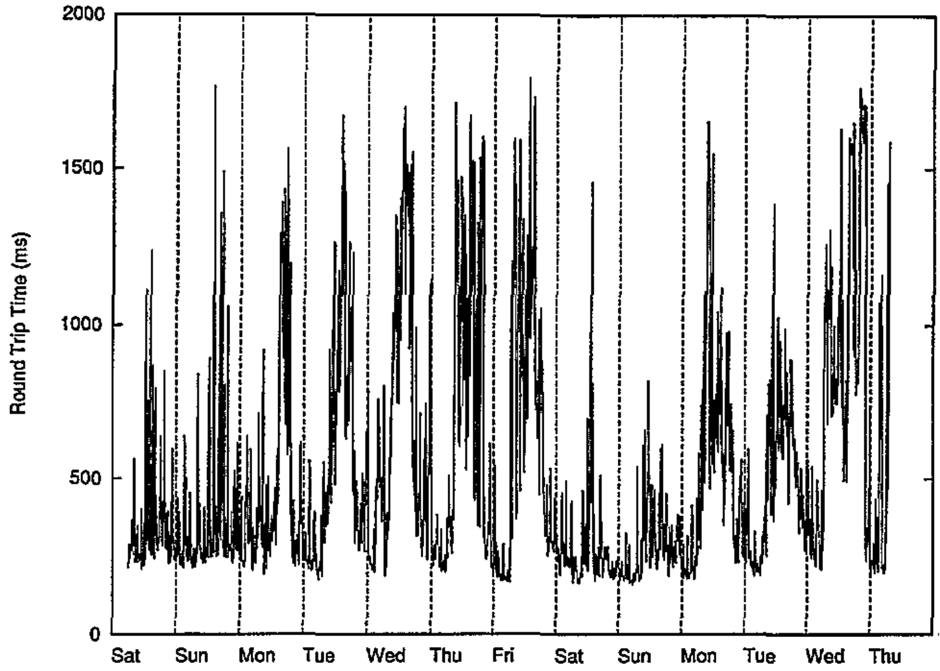


**Figure 2.** The changing of the round-trip time between two points of Internet from the 2nd to the 14th of October 1993. Vertical lines separate the days from each other. Daily and weekly periodicity can be easily recognized.

complex graph with crosspoints can evoke the picture of a road system, with cars stopping at crossings and moving towards their destination. Measurements of highway traffic [9] also gives $1/f$ spectra. Recently, several models [10, 11] of highway traffic were proposed. This analogy was used to model density waves in the flow of granular media [12] and it can also be used for computer networks, since the basics of packet sending, slowing and acceleration, are very similar to the behaviour of cars in these simple models.

The prototype for a one lane highway [11] is a one-dimensional cellular automaton. Each cell can have zero (no car) or some finite value $v$ (car with speed $v$). A vehicle with velocity $v$ is either accelerated ($v \to v+1$), when there are at least $v+1$ empty cells in front of it, or it is slowed down ($v \to \Delta x$) when there are only $\Delta x$ empty cells in front of it. After this velocity update, the car is advanced $v$ positions in one direction. The direction of movement is the same for all cars, and periodic boundary conditions are used in the model.

Let us now consider a heavily loaded computer network. We will discard all the side branches, and concentrate on the two-point connection described. Since we only want to model the internal traffic, not the behaviour of sources, the different nodes: gateways, sources and destinations are not distinguished. With the dictionary given below we can easily map computer network traffic to a highway traffic model:

speed $\equiv$ number of bytes sent per second

displacement (integral of speed over time) $\equiv$ the total amount of data sent

distance to next car $\equiv$ free space in next buffer.

So, node $i$ sends $v_i$ bytes per second to node $i+1$, and the total number of bytes sent successfully is $x_i$. In one timestep this means that $x_i \to x_i + v_i$. Data goes from node to

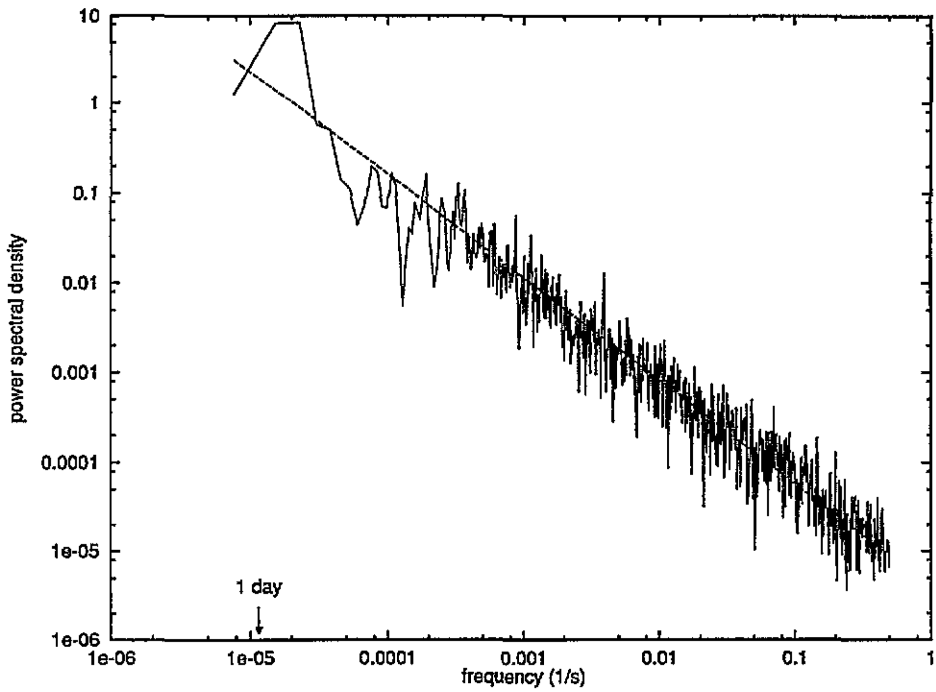


**Figure 3.** The power spectrum of RTT data. The fit (without the daily peak at the left side) gives a slope of $-1.15$.

node. Each of them has buffer size $b$, and $b_i^0$ of them are initially empty. We suppose, that this buffer is large compared to the outcoming data flow, so it is never empty. The sender always tries to increase its speed (up to a maximum):

$$v_i \to v_i + 1\,. \tag{1}$$

If there is no space in the buffer of node $i+1$, since more data is received than sent

$$x_{i+1} + b_{i+1}^0 < x_i + v_i + 1\,. \tag{2}$$

Source-quench messages reduce the speed of the sender to the number of empty spaces in the buffer $i+1$:

$$v_i \to x_{i+1} + b_{i+1}^0 - x_i\,. \tag{3}$$

This is exactly the same as the highway model mentioned above. Of course, to get better models one has to modify the above assumptions, distinguish between the behaviour of sources and gateways, handle buffers more realistically, make a network rather than single-lane model, etc.

Developing such models, more appropriate for Internet, including its structure, the behaviour of the data sources [13] and the different routing and congestion control methods is the subject of our future work. It could also be interesting to measure, for example, the occupancy versus flow diagram [10], which shows very good accordance between highway models and real measurements.

Designers of Internet devices usually take into account only the individual agents of the network, but no collective effects. The models of highway traffic clearly show that the formation of a traffic jam on an overloaded road is a collective phenomenon, not a direct effect of the behaviour of a single driver. We hope that future models—apart from the theoretical interest—can help designers to produce more suitable network devices.

## Acknowledgments

The author is indebted to I M Jánosi, A Czirók, T Geszti and T Vicsek for helpful discussions and careful reading of the manuscript. The work was supported partly by the Hungarian Academy of Sciences (OTKA F4437).

## References

[1] Lottor M 1993 *Internet Domain Survey* InterNIC Information Services
[2] Hogg T and Hubermann B A 1987 *Phys. Rep.* **56** 227
[3] Barden R and Postel J 1987 *Requirements for Internet Gateways (RFC-1009)* USC Information Science Institute
[4] Postel J 1981 *Internet Control Message Protocol (RFC-792)* USC Information Science Institute
[5] Nagle J 1984 *Congestion Control in IP/TCP Internetworks (RFC-896)* Ford Aerospace
[6] Mankin A and Ramakrishnan K 1991 *Gateway Congestion Control Survey (RFC-1254)* Digital Equipment Corporation
[7] Packard N H, Crutchfield J P, Farmer J D and Shaw R S 1980 *Phys. Rev. Lett.* **45** 712
[8] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes* (Cambridge: Cambridge University Press)
[9] Musha T and Higuchi H 1976 *Japan. J. Appl. Phys.* **15** 1271
[10] Nagel K and Schreckenberg M 1992 *J. Physique I* **2** 2221
[11] Nagel K and Herrmann H J 1993 *Physica* **199A** 254
[12] Leibig M 1994 *Phys. Rev.* E **49** 184
[13] Cáceres R, Danzig P B, Jamin S and Mitzel D J 1991 *Characteristics of Wide-area TCP/IP Conversations* ACM SIGCOMM