EUROPHYSICS LETTERS 15 April 2003

Europhys. Lett., **62** (2), pp. 189–195 (2003)

Similarities between communication dynamics in the Internet and the autonomic nervous system

K. Fukuda[1,2], L. A. Nunes Amaral[3] and H. E. Stanley[1]

[1] *Center for Polymer Studies and Department of Physics, Boston University Boston, MA 02215, USA*

[2] *NTT Network Innovation Laboratories - Tokyo, 180-8585, Japan*

[3] *Department of Chemical Engineering, Northwestern University Evanston, IL 60208, USA*

(received 18 December 2002; accepted 21 February 2003)

PACS. `05.45.-a` – Nonlinear dynamics and nonlinear dynamical systems.
PACS. `05.45.Tp` – Time series analysis.

Abstract. – The Internet is a world-wide communication network, whose optimization depends on the knowledge of the statistical characterization of the aggregated traffic flow. Internet traffic is dependent on a number of factors, including communication protocols, network topology, and human behavior. Using a recently proposed segmentation algorithm, we find a surprising analogy between the nonstationarity and the correlations in the communication dynamics in the Internet and in another communication network of great interest: the autonomic nervous system (ANS). The ANS controls involuntary muscle motion, secreting glands, and the heart, hence, we surmise that the time interval between successive heartbeats —an easily measured physiological signal— provides a probe of the communication dynamics for the ANS. We find quantitative similarities between the statistical properties of i) healthy heart rate variability and non-congested Internet traffic, and ii) diseased heart rate variability and congested Internet traffic. Our findings suggest that the understanding of the mechanisms underlying the "human-made" Internet could help to understand the "natural" network that controls the heart.

Introduction. – The Internet [1, 2] —a large communication network that now connects more than 10^8 hosts— is a prime example of a self-organizing complex system [3–9], having grown mostly in the absence of centralized control or direction. In this network, information is transferred in the form of packets from the sender to the receiver via routers, computers which are specialized to transfer packets to another router "closer" to the receiver (fig. 1a). A router decides the route of the packet using only local information obtained from its interaction with neighboring routers, not by following instructions from a centralized server. Thus, it can be viewed as a simple and autonomous entity. A router stores packets in its finite queue and processes them sequentially. However, if the queue overflows due to excess demand, the router will discard incoming packets, a situation corresponding to congestion. A router can only control incoming traffic by discarding arriving packets, so that in order to know and

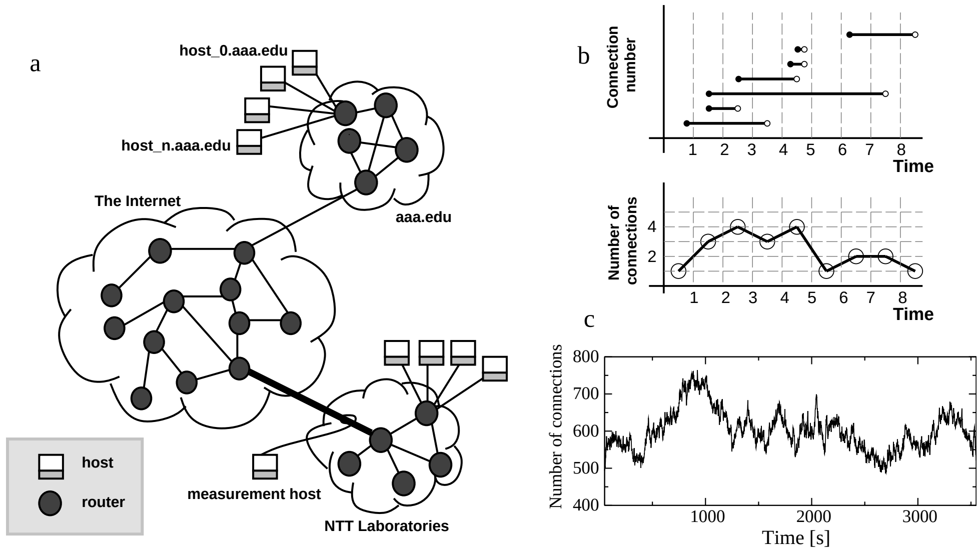

Fig. 1 – Measuring Internet traffic. a) Schematic representation of Internet topology. The Internet comprises links and nodes (routers and hosts), which are connected in a scale-free structure [10, 11]. Suppose that one downloads a file on a web page in a host at `aaa.edu` from a host in NTT laboratories. In this case, a connection is established between the two hosts, and the packets encoding the file travel from the source host to the destination host via routers. The connection is based on a feedback control, so that the duration of the connection for transferring the file strongly depends on the current network traffic conditions; if the network is congested, it takes longer to finish the transfer. Here, we report on measurements of the number of connections at the link between NTT laboratories and the Internet (thick line). There is a firewall host in NTT labs, so that a connection initiated from the Internet is restricted. Namely, our data reflects the daily activity pattern of the NTT employees. b) Measuring the number of connections. In the top panel, each line represents a connection passing through the observation link. For example, the first connection starts at 0.5 s (black circle) and finishes at 3.5 s (open circle). We count the number of connections within a one second interval and obtain a time series of the number of connections (bottom panel). c) A typical data set, showing the number of connections for each of the 3600 one second intervals between 13:00 and 14:00 on July 19, 2001.

adjust to the current network traffic condition, each host has the ability to control its traffic by using a feedback-based flow control [12] for the communication between the sender and the receiver.

Even though these rules controlling traffic flow were programmed by humans, the dynamics of Internet traffic [13–20] are difficult to predict due to the complex interactions between routers, the flow control mechanisms, and the diversity of applications running in the Internet. Moreover, the traffic flow is also highly correlated with human activity, as most users of the Internet are human. Internet traffic fluctuations are reported to be statistically self-similar [13–15]. It has also been reported that Internet traffic displays two separate phases, congested and non-congested, and that traffic fluctuations are characterized by $1/f$-behavior at the point separating the two phases [17–20].

Data. – We record every packet flowing through the link between NTT laboratories in Tokyo and the Internet during 4 days in July 2001 (figs. 1b and c). We form data sets,

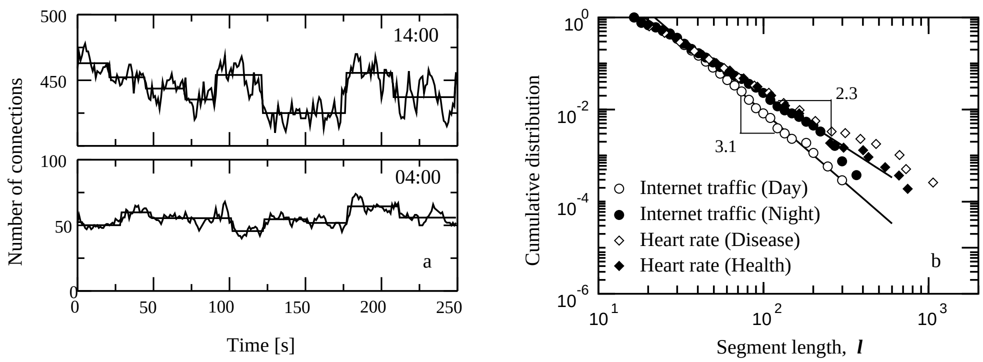

Fig. 2 – Analysis of the number of connections between the NTT laboratories and the Internet. a) Number of connections for periods during the day (four minutes starting at 14:00) and the night (four minutes starting at 4:00). The solid lines display the output of the segmentation algorithm [21]. b) Cumulative distribution of segment lengths of stationary periods in the number of connections and the heartbeat variation. We find that the distributions are characterized by power law decays for both night and day (implying there is no typical size of the stationary duration) but with different exponents: $\gamma = 2.3 \pm 0.3$ for night, and $\gamma = 3.1 \pm 0.2$ for day. The distributions of the heartbeat interval fluctuations are also approximately power law with exponent $\gamma \approx 2.2$ for both the healthy and diseased cases.

each consisting of recordings of the number of connections between a host at NTT and some other host in the Internet in each second of one hour (3600 points each). We analyze 80 such one-hour time series.

Figure 2a displays the number of connections for periods during night (04:00) and day (14:00). It is visually apparent that both time series are nonstationary, and are characterized by different mean values depending on the observation period; the mean number of connections is ≈ 50 connections/s during the night, and ≈ 700 connections/s during the day. In order to quantitatively characterize such a nonstationary time series, we will use an analysis method that we describe in the following.

The segmentation algorithm. – We define stationarity here as the property that the mean value of a time series does not change under time translation. In order to divide a nonstationary time series into stationary segments, we perform the following procedure [21]: we move a sliding pointer from left to right along the time series. At each position of the pointer, we compute the mean of the subset of the signal to the left of the pointer (μ_{left}) and to the right (μ_{right}). To measure the significance of the difference between μ_{left} and μ_{right}, we compute the statistic

$$t \equiv \left| \frac{\mu_{\text{left}} - \mu_{\text{right}}}{S_D} \right|, \tag{1}$$

where

$$S_D = \left(\frac{(N_{\text{left}} - 1)s_{\text{left}}^2 + (N_{\text{right}} - 1)s_{\text{right}}^2}{N_{\text{left}} + N_{\text{right}} - 2} \right)^{1/2} \left(\frac{1}{N_{\text{left}}} + \frac{1}{N_{\text{right}}} \right)^{1/2} \tag{2}$$

is the pooled variance, s_{left}, s_{right} are the standard deviations of the data to the left and to the right of the pointer, respectively, and N_{left} and N_{right} are the number of points to the left and to the right of the pointer. Next, we calculate the statistical significance $P(\tau)$ of the

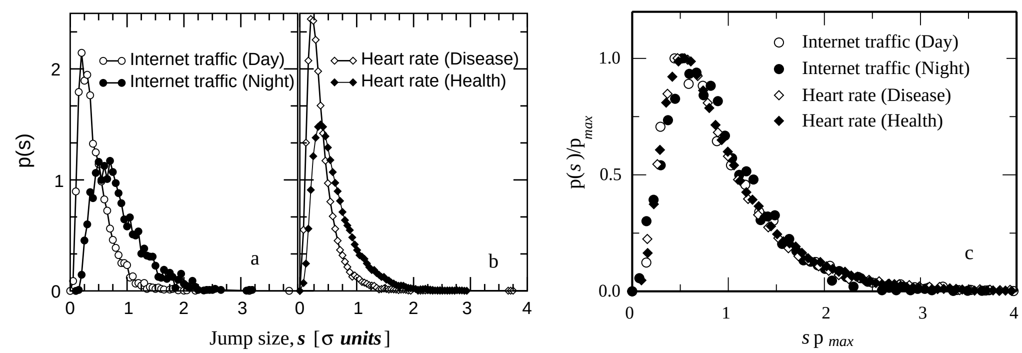

Fig. 3 – Probability density distributions of the jump sizes for a) Internet traffic fluctuations, and b) heartbeat interval fluctuations. For the Internet traffic, the day data peak at $s \approx 0.2$, while the night data peak at $s \approx 0.7$. Also, for the heartbeat intervals, the diseased heart data peak at $s \approx 0.2$, and the healthy heart data peak at $s \approx 0.4$. c) Scaled distributions of jump sizes. The distributions are scaled by the maximum value of the probability density $p_{\max}$, and s scaled by $1/p_{\max}$. It is visually apparent that the distributions are similar for night and day. Moreover, these distributions are similar to the distributions found in the analysis of heartbeat intervals.

position of pointer having the largest t value. $P(\tau)$ is numerically approximated as [21]

$$P(\tau) \approx \left\{1 - I_{[\nu/(\nu+\tau^2)]} \times (\delta\nu, \delta)\right\}^{\gamma}, \tag{3}$$

where $\gamma = 4.19 \ln N - 11.54$, $\delta = 0.40$, N is the size of the time series to be split, $\nu = N - 2$, and $I_x(a, b)$ is the incomplete beta-function. If the significance value is larger than a threshold (typically 0.95), one decides that the left and right parts of the time series have statistically significant differences in mean, and one cuts the time series into two time series. Otherwise, the time series remains uncut. If the time series is cut, we repeatedly iterate this algorithm in each fraction until the significance value is smaller than the threshold or the length of the time series is smaller than a minimum size ℓ_0 [21]. In our analysis, we set $\ell_0 = 18\,\mathrm{s}$.

Results. – We divide each of the 80 time series in our data records into "stationary segments" using the recently proposed segmentation algorithm of Bernaola-Galván *et al.* [21]. The solid lines in fig. 2a illustrate the output of the segmentation algorithm. We also apply the Kolgomorov-Smirnov (KS) test to two distributions of segment lengths for all one-hour traces, and obtain two empirical distributions: the night [01:00-06:00] and the day [10:00-16:00]. In fig. 2b we show the cumulative distribution $P(\ell)$ of the durations ℓ of the stationary periods. We find power law distributions,

$$P(\ell) \sim \ell^{-\gamma}, \tag{4}$$

for both night and day, but with different exponent values: $\gamma = 2.3 \pm 0.3$ for night and $\gamma = 3.1 \pm 0.2$ for day. A power law distribution indicates that there is no typical scale for the duration of the stationary segments in Internet traffic data.

To characterize the changes between stationary segments, we study the jump size s, which we define as the absolute value of the difference of the mean number of connections between neighboring stationary segments (fig. 3a). The distribution of the day data peaks at $s \approx 0.2$, and the distribution of the night peaks at $s \approx 0.7$, indicating that the variations of the day traffic are smaller than the variations of the night traffic.

Table I – *Quantitative comparison between Internet traffic data and heartbeat intervals data. The parameters γ, τ, and α are defined in eqs.* (4), (5), *and* (6), *respectively.*

	Internet traffic (night)	Heart rate (healthy)	Internet traffic (day)	Heart rate (failure)
α	1.0 ± 0.1	1.0 ± 0.1	1.3 ± 0.1	1.3 ± 0.2
γ	2.3 ± 0.3	2.2 ± 0.2	3.1 ± 0.2	2.1 ± 0.1
τ	0.5 ± 0.1	0.6 ± 0.1	0.6 ± 0.2	0.6 ± 0.1

In order to compare the functional form of the distribution, we scale the probability densities $p(s)$ by the maximum probability density $p_{\max}$, and we scale s by $1/p_{\max}$. We find that these scaled curves display the same functional form for both night and day (fig. 3c); both decay exponentially for $sp_{\max} > 0.55$ (table I),

$$\frac{p(s)}{p_{\max}} \sim e^{-sp_{\max}/\tau}. \tag{5}$$

This exponential decay suggests that there is a well-defined scale for the changes in the mean number of connections between stationary segments.

We also analyze the auto-correlations in the number of connections by using the DFA method [22]. We find a power law relation between the average magnitude of the fluctuations $F(n)$ and the observation scale n, eq. (6), where the scaling exponent α quantifies the nature of the correlations,

$$F(n) \sim n^{\alpha}\,. \tag{6}$$

Figure 4a displays the results of the DFA method for day and night data. Both curves are characterized by a power law dependence, implying that there is no typical time scale for the decay of the correlations in the number of connections. We find a significant difference in the type of correlations for less active (night) and more active (day) periods: $\alpha = 1.0 \pm 0.1$ for

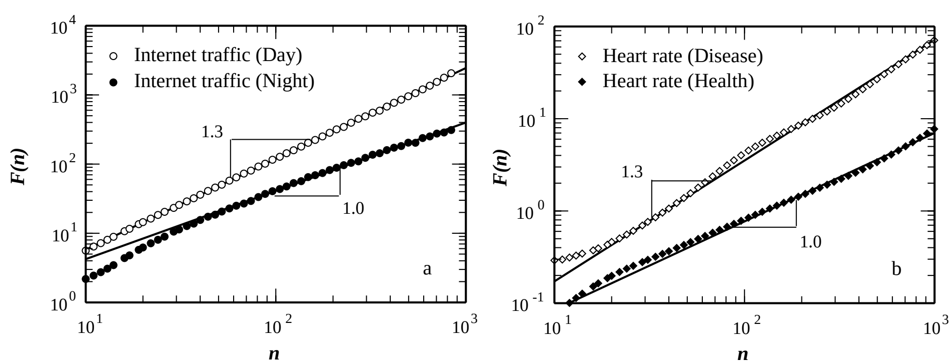

Fig. 4 – a) DFA estimation of auto-correlations of the number of connections. For both night and day, we find a power law scaling of the fluctuations but with different exponents; $\alpha = 1.0 \pm 0.1$ for night and $\alpha = 1.3 \pm 0.1$ for day. b) DFA estimation of auto-correlations of the heartbeat intervals. For the healthy heartbeat intervals, we find $\alpha = 1.0 \pm 0.1$, close to the exponent for the Internet night traffic data, and for the diseased heart data we find $\alpha = 1.3 \pm 0.2$, close to the exponent found for the Internet day traffic data.

$n > 50$, and $\alpha = 1.3 \pm 0.1$ for $n > 9$. These results indicate that the night dynamics are of the $1/f$-type, while the day dynamics are closer to Brownian noise. We tested these findings for traffic data at other links in the Internet(1) and obtained consistent results. Thus, it appears that our results are generic and not specific to the NTT labs. Note that these findings of $1/f$-behavior for fluctuations in number of *connections* do not contradict the findings of $1/f$-behavior for fluctuations in number of *bytes* per unit time reported in [18].

Discussion. – Our findings bear a striking quantitative similarity to those reported for heartbeat interval time series [21–27] (table I). Indeed, for heartbeat intervals, the distribution of durations of stationary segments also decays as a power law with an exponent $\gamma \approx 2.2$ [21], close to the exponent for night periods of the Internet traffic data. In the jump size distribution (fig. 3b), the distribution for the diseased heart peaks at $s \approx 0.2$, and the distribution for healthy heart peaks at $s \approx 0.4$, indicating a similarity between the diseased heart and the day traffic, and a similarity between the healthy heart and the night traffic. Moreover, as shown in fig. 3c and table I, the scaled probability densities of jump size display similar functional forms to those found for Internet traffic data [21, 24]. Further, the DFA analysis in fig. 4b shows that the correlation of the fluctuations of heartbeat intervals is also characterized by a power law [22]. The exponents for the healthy data ($\alpha \approx 1.0$) and for the disease data ($\alpha \approx 1.3$) are close to those for the Internet day data and for the Internet night data, respectively.

Our results show that if the network is not congested the information is transferred without disruptions, and we find traffic behavior that resembles healthy heartbeat dynamics. In contrast, when information transfer is blocked or delayed by congestion during periods of greater human activity, the network dynamics appear to resemble diseased heart dynamics. A question prompted by our empirical results is the reason for the similarity in the dynamics of these two systems, the Internet and the autonomic nervous system (ANS) [23]. Specifically, is this similarity due to chance or does it reveal an analogy between communication dynamics in the Internet and the ANS that regulates the heartbeat? In this regard, it is interesting to note that i) the ANS may be viewed as a communication network comprising a very large number of elements which connects and directs information between different physiologic systems, and ii) the sequence of the heartbeat intervals may be viewed as a probe of the activity of the ANS(2) [28].

It thus appears that the Internet, a human-made system for which the "inner" working of each unit and their connectivity is known, may provide a useful "model system" to investigate the mechanisms responsible for the dynamics of the ANS. This possibility is not implausible. Indeed, very simple models of very complex systems in many cases provide deep insights. For example, the Ising model and its simple variants such as the Heisenberg model are sufficient to quantitatively describe a wealth of very complex systems in regions of their respective phase diagrams where scale invariance is displayed. The principle of "universality" in chemistry and physics [29, 30], whereby diverse systems are described by the identical (simple) model, may have its counterpart in physiology. Specifically, simplified network control mechanisms in the Internet may provide useful insights into the dynamics of the ANS. For example, feedback control in each host, a mechanism to adapt traffic to the current network status, reproduces $1/f$-fluctuations arising from the network of interactions with other hosts [20].

(1)The measurement point is a link between Auckland University in NZ and the Internet, and the data record can be obtained from `http://pma.nlanr.net/Traces/long/auck4.html`.

(2)Parasympathetic blockade in healthy subjects leads to dynamics similar to those observed for diseased heart patients.

* * *

We thank T. MORI for providing the NTT data traces, and the WAND research group at Waikato University for providing the Auckland traces. We also thank NIH/NCRR (P41 RR13622) and NSF for support.

REFERENCES

[1] TANENBAUM A., *Computer Networks*, 3rd edition (Prentice Hall, Englewood Cliffs) 1996.
[2] COMER D., *Internetworking with TCP/IP*, Vol. **1**, 4th edition (Prentice Hall, Englewood Cliffs) 2000.
[3] MONTROLL E. W. and SHLESINGER M. F., in *Nonequilibrium Phenomena II. From Stochastics to Hydrodynamics*, edited by LEBOWITZ J. L. and MONTROLL E. W. (North-Holland, Amsterdam) 1984, pp. 1-121.
[4] BUNDE A., KROPP J. and SCHELLNHUBER H. J., *The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes* (Springer-Verlag, Berlin) 2002.
[5] ABBOTT D. and KISH L. B. (Editors), *Unsolved Problem of Noise 1999, AIP Conf. Proc.*, Vol. **511** (American Institute of Physics, Melville) 2000.
[6] WATTS D. J. and STROGATZ S. H., *Nature*, **393** (1998) 440.
[7] BARABÁSI A.-L. and ALBERT R., *Science*, **286** (1999) 509.
[8] STROGATZ S. H., *Nature*, **410** (2001) 268.
[9] ALBERT R. and BARABÁSI A.-L., *Rev. Mod. Phys.*, **74** (2002) 47.
[10] FALOUTSOS M., FALOUTSOS P. and FALOUTSOS C., *Comp. Commun. Rev.*, **29** (1999) 251.
[11] PASTOR-SATORRAS R., VÁZQUEZ A. and VESPIGNANI A., *Phys. Rev. Lett.*, **87** (2001) 258701.
[12] POSTEL J., *Transmission Control Protocol*, RFC 793 (1981).
[13] LELAND W. E., TAQQU M. S., WILLINGER W. and WILSON D. V., *IEEE/ACM Trans. Network*, **2** (1994) 1.
[14] PAXSON V. and FLOYD S., *IEEE/ACM Trans. Network*, **3** (1995) 226.
[15] CROVELLA M. and BESTAVROS A., *IEEE/ACM Trans. Network*, **5** (1997) 835.
[16] CSABAI I., *J. Phys. A*, **27** (1994) 417.
[17] TAKAYASU M. *et al.*, *Physica A*, **233** (1996) 924.
[18] TAKAYASU M. *et al.*, *Physica A*, **277** (2000) 248.
[19] FUKUDA K. *et al.*, *Physica A*, **287** (2000) 289.
[20] FUKUDA K. *et al.*, in *Traffic and Granular Flow 2001*, edited by SUGIYAMA Y. *et al.* (Springer, Heidelberg) 2002.
[21] BERNAOLA-GALVÁN P. *et al.*, *Phys. Rev. Lett.*, **87** (2001) 168105.
[22] PENG C.-K. *et al.*, *Chaos*, **5** (1995) 82.
[23] GREEN J. H., *An Introduction to Human Physiology* (Oxford University Press, London) 1972.
[24] IVANOV P. CH. *et al.*, *Nature*, **383** (1996) 323.
[25] IVANOV P. CH. *et al.*, *Nature*, **399** (1999) 461.
[26] GOLDBERGER A. L. *et al.*, *Proc. Natl. Acad. Sci. USA, Suppl. 1*, **99** (2002) 2466.
[27] BUNDE A. *et al.*, *Phys. Rev. Lett.*, **85** (2000) 3736.
[28] AMARAL L. A. N. *et al.*, *Phys. Rev. Lett.*, **86** (2001) 6026.
[29] STANLEY H. E., *Nature*, **378** (1995) 554.
[30] STANLEY H. E., *Rev. Mod. Phys.*, **71** (1999) S358.