Fluctuation-induced traffic congestion in heterogeneous networks

A. S. Stepanenko1, I. V. Yurkevich1,2, C. C. Constantino3 and I. V. Lerner2

1 School of Engineering and Applied Science, Aston University - Birmingham B4 7ET, UK, EU
2 School of Physics and Astronomy, University of Birmingham - Birmingham B15 2TT, UK, EU
3 School of Electronic, Electrical and Computer Engineering, University of Birmingham Birmingham B15 2TT, UK, EU

received 7 September 2012; accepted in final form 11 October 2012
published online 12 November 2012

PACS 64.60.Ht – Dynamic critical phenomena
PACS 89.75.Da – Systems obeying scaling laws
PACS 89.20.Hh – World Wide Web, Internet

Abstract – In studies of complex heterogeneous networks, particularly of the Internet, significant attention was paid to analyzing network failures caused by hardware faults or overload, where the network reaction was modeled as rerouting of traffic away from failed or congested elements. Here we model another type of the network reaction to congestion—a sharp reduction of the input traffic rate through congested routes which occurs on much shorter time scales. We consider the onset of congestion in the Internet where local mismatch between demand and capacity results in traffic losses and show that it can be described as a phase transition characterized by strong non-Gaussian loss fluctuations at a mesoscopic time scale. The fluctuations, caused by noise in input traffic, are exacerbated by the heterogeneous nature of the network manifested in a scale-free load distribution. They result in the network strongly overreacting to the first signs of congestion by significantly reducing input traffic along the communication paths where congestion is utterly negligible.

Introduction. – All Internet users are familiar with the feeling of frustration when their network connection slows down or halts. Barring cascading failures [1–9] which can shut down parts of the network, such a slowdown is a sign of network congestion which happens when the traffic load on some network elements exceeds their capacity [2–6]. For the Internet congestion can be quantified as a relative difference between the rates of sent and delivered data packets [6,10], with excess packets being eventually dropped. The first network reaction to a lost packet is a significant reduction of the transmission rate at the source followed by a slow recovery to the normal rate. When several loss events occur in quick succession, a multiplicative reduction drastically suppresses the transmission rate, which feels as congestion for the end user. If congestion persists for longer, the network eventually reroutes traffic away from congested links which may overload other links triggering a cascade of failures [4].

A surprising result is that transmission rates could be significantly reduced even when relative losses are utterly negligible. Such a reduction results from the existence of strong fluctuations of data losses along a typical communication path at the very onset of congestion. The loss fluctuations arise at each link at the threshold of its capacity due to noise in input traffic. Although such fluctuations exist only on shorter, mesoscopic time scales and will die out in time, we show that they might trigger a network overreaction to the first signs of losses by a feedback mechanism aggressively reducing traffic rates along the routes perceived as congested due to multiple loss events. The overreaction results from the probability of such events on the mesoscopic scale being much higher than the product of the single-event probabilities.

Network model. – The fluctuations are greatly magnified in heterogeneous networks characterized by a power law (PL) distribution of the link load since congestion on links with high load affects a disproportionately large number of communication paths, as illustrated in fig. 1. The link load in a network with a homogeneous traffic input distribution is proportional to the link betweenness $B_i$ (roughly, the number of shortest paths through link $i$) [11–13]. Many heterogeneous networks, including the Internet, fall into the category of scale-free (SF) networks characterized by the PL distribution of node degree [14–18]. Load distribution in SF networks also follows a (truncated) PL [18–24],

$$P_L(\ell) \propto \ell^{-2-\beta},$$

Where $\beta$ is the degree exponent.
with an almost universal exponent, $2 + \beta \sim 2.0 - 2.3$. The load distribution of the real Internet was found [18,25,26] to be in agreement with the above.

In our model we account for three sources of stochasticity: i) the spread (1) of the average link load, quenched on the time scales of interest; ii) the inevitable deviations from homogeneity of the input traffic, modeled as a quenched distribution of particular link congestion thresholds; iii) the dynamic noise in the input traffic. This noise is characterised by randomness in both packet arrival rate and packet length. Although the former might be heterogeneous and exhibiting bursty behaviour at the network edge, multiplexing many such input traffic streams makes the input traffic in the network core, i.e., on links with high betweenness, quasi-Markovian or even Poissonian [27,28]. This allows us to describe the input traffic noise as a Fokker-Planck process at a single link.

We focus on a critical regime at the onset of congestion in a typical communication path — naturally, it is a subset of paths rather than the entire network that might become congested. In this regime, at least one link in the path has a small imbalance,

$$\eta_i \equiv 1 - \tau_i r_i,$$

between the average packet arrival rate (load), $1/\tau_i$, and departure rate (capacity), $r_i$, at link $i$. The nodes in the Internet core are routers and the links are output memory buffers (with attached transmission lines). For $\eta_i > 0$ a memory buffer eventually becomes full and a newly arrived packet is dropped. On average, $\eta_i = \langle \Phi_i(T) \rangle$ where $\Phi_i(T)$ is the fraction of packets dropped during an observation time window $T$. Shifting this window in time causes $\Phi_i$ to fluctuate due to inevitable flow variations [29]. We show the loss fluctuations to be crucial for network transport at a certain mesoscopic time scale; for large enough $T$ they die out and $\Phi_i$ equals $\eta_i$. Positive $\eta_i$ play the role of a local congestion parameter: their sum defines the network congestion parameter [6,10].

The main quantity of interest in this letter is relative loss, $\Phi$, along a typical communication path. This is governed by losses in the comprising links and fluctuates due to both the noise in each link and uncertainty in the betweenness of each link comprising a randomly chosen path. The probability of a randomly picked link to be in the path is proportional to its betweenness. When $\Phi$ is small (which is the case at the onset of congestion), it is given by $\Phi \approx \sum_i \Phi_i$, where the summation is carried over the links making up this path. Hence, in a network with average link betweenness $\bar{B}$, a small loss along a path with $a$ links is given by

$$\Phi = \sum_{i=1}^a \ell_i \Phi_i, \quad \ell_i \equiv B_i / \bar{B}.$$  \hfill (2)

The quenched distribution of the relative load $\ell_i$ is given by the truncated PL (1), cut from below by $\ell \sim 1$.

**Single-link model.** — To describe noise in packet arrivals at link $i$ we assume, without loss of generality, that the inter-arrival time is random, with average $\tau_i$; while packets have a fixed length $l_0$ (see footnote 1). Arriving packets join the queue in the memory buffer. The queue length, $x_i(t)$ (measured in $l_0$), performs a random walk bounded by a buffer size $c_i$. The probability density, $w_i$, of diffusion from $x'$ to $x$ over time $t$ obeys the Fokker-Planck equation with diffusion and advection coefficients $D_i \equiv 1/\tau_i$ and $V_i \equiv \eta_i/\tau_i$ together with the standard probability-conservation boundary conditions:

$$\partial_t w_i(x, x'; t) = \left[ -V_i \partial_x + D_i \partial^2_x \right] w_i(x, x'; t).$$  \hfill (3)

In the critical regime the queue hovers at the upper boundary. A newly arriving packet is dropped every time when the queue length $x_i(t)$ overflows reaching the boundary layer, $c_i - 1 \leq x_i(t) \leq c_i$. Thus the fraction of packets lost over an observation time $T \gg \tau_i$ is

$$\Phi_i(T) = \frac{1}{T} \int_0^T dt \theta[x_i(t) - c_i + 1] = \frac{1}{K_i} \sum_{k=1}^{K_i} \sigma_k.$$  \hfill (4)

Here $\theta[y] = 1$ for $y > 1$ and 0 otherwise, $K_i \equiv [T/\tau_i] \gg 1$ is the number of packets arrived over time $T$, and $\sigma_k$ equals 1 if $x_i(t)$ reached the boundary layer at the $k$-th step, or 0 otherwise.

To determine the probability density function (PDF) of $\Phi_i$, we begin with the characteristic function, $\chi_T(q_i) = \langle e^{i q_i \Lambda_i} \rangle$, of the distribution of cumulative loss, $\Lambda_i = \langle T/\tau_i \rangle \Phi_i(T)$. Using eqs. (3) and (4) we represent $\chi_T$ as the sum of time-ordered integrals:

$$\chi_T(q_i) = \sum_{n=0}^{\infty} \frac{(i q_i)^n}{n!} \langle \Lambda_i^n \rangle,$$

$$= \sum_{n=0}^{\infty} (i q_i)^n \int d^n t \mathcal{R}_i(t_n - t_{n-1}) \cdots \mathcal{R}_i(t_2 - t_1) p_i,$$

$^1$The fixed packet length is chosen to simplify the presentation. The real distribution of packet length in the Internet is almost flat between sharp peaks at lower and upper cutoffs [30], the latter being many orders of magnitude smaller than the buffer size. Introducing any such distribution of $l_0$ does not change the Fokker-Planck description of the single link [31].
running over the regions \(0 < t_1 < \cdots < t_n < T\). Here \(p_i \equiv w_{i}(c_i, \tau' ; t \to \infty) = \eta_i (1 - e^{-\eta_i c_i})^{-1}\) is the stationary probability density for the queue to be in the boundary layer and \(R_i(t) = w(c_i, \tau ; t)\) is the probability of the return to the boundary. We rewrite the expression for \(\chi_T\) as the integral equation
\[
\chi_T(q_i - 1) = i q_i \left\{ p_i T + \int_0^T dt \; R_i(T-t) \; \left[ \chi_T(q_i - 1) \right] \right\}. \tag{5}
\]
The inverse Fourier transform from \(q_i\) to \(\Lambda_i\) shows the PDF of \(\Lambda_i\) to be the sum \(F_T(\Lambda_i) + A_i \delta(\Lambda_i),\) with \(F\) describing losses \((\Lambda_i > 0)\) and \(A_i\) being the probability of no losses over the time \(T,\) with \(1 = A_i + \int_0^\infty d\Lambda \; F_T(\Lambda).\) The first part, \(F_T(\Lambda_i)\) obeys the following equation:
\[
F_T(\Lambda_i) + \int_0^T dt \; R_i(T-t) \; \partial_\Lambda_i F_T(\Lambda_i) = 0. \tag{6}
\]
Solving eq. (6) using the Laplace transform with respect to \(T\) gives
\[
\mathcal{F}_\xi(\Lambda_i) = \frac{p_i e^{-\Lambda_i \tau_i}}{\tau_i \varepsilon^2 \mathcal{R}_\xi}, \quad \mathcal{R}_\varepsilon = \sqrt{\frac{\Lambda_i^2 + 4\tau_i \varepsilon}{2\tau_i \varepsilon}} + \eta_i,
\tag{7}
\]
where \(\mathcal{F}_\xi\) and \(\mathcal{R}_\varepsilon\) are the Laplace transforms of \(F_T\) the return probability \(R_i(T),\) respectively.

For a perfectly designed network with fully utilized resources for homogeneous input traffic, \(c_i\) and \(\tau_i^{-1}\) are proportional to the relative link load \([19], \ c_i \equiv c \ell, \ \text{and} \ \tau_i^{-1} \equiv \ell \tau_i^{-1} \ (c, \tau_i^{-1} \text{are the parameters for a link with average betweenness, } B), \ \text{while the imbalance} \ \eta_i \equiv \ell_i\text{-independent. Then the congestion threshold } \eta_i = 0 \ \text{is reached simultaneously by all links. Naturally, such a complete utilization is impossible: design imperfections and local variations of demand cause the congestion thresholds to spread}[5]. \ We model such a (quenched) spread as a sharply peaked symmetric distribution of \(\eta_i\) with criticality width \(\gamma < 1\). This spread is analogous to the tolerance parameter in a finite-capacity model [2], albeit we focus here on the overload-induced losses. Realistic regimes are bounded by \(a \ll \gamma^{-1} \lesssim c.\) For \(\gamma^{-1} \gtrsim c\) congestion thresholds are still reached simultaneously for all links (the spread is too small). While a network with \(a \gamma \gtrsim 1\) would be permanently congested: (much) more than one link is congested along the typical path, i.e. the loss rate is not small.

The PDF of \(\Phi_i\) has the same form as that of \(\Lambda_i,\) namely \(\Lambda_i \delta(\Phi_i) + P_T(\Phi_i; \ell_i),\) where its lossy part, \(P_T(\Phi_i; \ell_i),\) is found for fixed \(\ell_i\) and \(\eta_i\) by rescaling \(F_T,\) the inverse Laplace transform of eq. (7), as \((\ell_i / \varphi_0^2) F(\ell_i / \varphi_0^2),\) where \(\varphi_0^2 \equiv \ell_i / \tau_i / T.\) In order to perform averaging over \(\eta_i\) and \(\ell_i\) we need explicit expressions for the long and short time limits of \(P_T(\Phi_i; \ell_i).\) The long \(T\) limit, \(\tau_i / T \equiv \varphi_0^2 / \ell_i \ll \eta_i^2,\) corresponds to \(\tau_i \ll \eta_i^2\) when \(F_T(\Lambda_i) \approx (\tau_i / \eta_i) \exp(-\Lambda_i \tau_i / \eta_i + \Lambda_i \tau_i^2 \eta_i^2 / 2)\) which results in \(P_T(\Phi_i; \ell_i)\) being a Gaussian distribution with average \(\eta_i\) and width \(\Phi_0 / \sqrt{T} \ll \eta_i.\) In this limit the packet loss is fully described by its average.

In the limit of relatively short times, \(\ell_i \eta_i^2 \ll \varphi_0^2 \ll 1,\) we have \(F_T(\Lambda_i) \approx (\tau_i / \eta_i) \exp(-\Lambda_i \tau_i / \eta_i),\) and the inverse Laplace transform can be expressed in terms of the complementary error function, but for our consideration it is sufficient to note that it is a constant of order \(\eta_i\) for \(\Phi_i \ll \ell_i / \Lambda_i^2.\) Thus, we conclude that \(P_T(\Phi_i; \ell_i)\) has a plateau of height \(\sim \ell_i \eta_i / \varphi_0^2\) from 0 to \(\Phi_i \ll \varphi_0 / \sqrt{\tau_i},\) followed by a fast decay \(\sim e^{-\Phi_i^2 / \varphi_0^2}.\) It also has a 6-like contribution at \(\Phi_i = 0\) (no losses) since the probability of loosing a packet is \(1 - A_i - \eta_i \sqrt{\tau_i / \varphi_0} \ll 1.\)

Now we perform the averaging of the PDF over \(\eta_i\) using the Gaussian distribution centred at \(\bar{\eta} = 0\) of width \(\gamma\) with \(1/c_i \ll \gamma \ll 1.\) We note that only a positive relative imbalance, \(\eta_i > 0,\) has the meaning of the congestion parameter while \(\eta_i < 0\) corresponds to the free-flow regime. Since negative \(\eta_i\) do not contribute to losses and the distribution of \(\eta_i\) is relatively sharp in the short time limit, \(\approx \ell_i \ll \varphi_0^2\), we find that the averaged PDF is obtained by substituting \(\gamma\) for \(\eta_i\) into the above results for \(P_T(\Phi_i; \ell_i).\)

In the long time limit, \(\ell_i \gg \Phi_0^2\), the averaged PDF has the form \(\sim \gamma^{-1} \exp(-\Phi_i^2 / 2\gamma^2),\) with the probability of not loosing a packet being \(A_i = 1 / 2\) due to the contribution of links in the free-flow regime.

**Results for network congestion.** – The PDF of losses along path (2) is a convolution of PDFs of statistically independent \(\ell_i \Phi_i,\) averaged over the load distribution (1). It still has the structure \(\delta(\Phi_i) + P_T(\Phi_i)\) with \(A = \prod_{i=1}^{n} (A_i)\). For \(\varphi_0 \ll \gamma\) losses in each link in (2) are on the macroscopic time scale, light-shaded (green online) in fig. 2. Then \(A = 2^{-a} \ll 1\) for \(a > 1\) and, by the law of large numbers, \(P_T(\Phi)\) is a normal distribution with average \((\Phi)_A \sim a \gamma A / \sqrt{\tau_i}.\) In the plateau in this rescaled PDF is stretched up to \(\varphi_0 \ell_i^{1/2}\) with height \(a \gamma / \varphi_0^2.\) For \(\Phi_i \ll \varphi_0\) averaging over the distribution (1) leaves the plateau intact. For \(\Phi_i \gtrsim \varphi_0\) the plateau exists only for links with load \(\ell_i (\Phi_i/\varphi_0)^2\) which becomes the lower limit of the plateau width due to the averaging over load. When the lower limit is much smaller than the upper, \((\varphi_0 / \gamma)^2,\) the averaging is contributed only by the former resulting in the PL tail
\[
P_T(\Phi) \sim \frac{a \gamma}{\varphi_0} \frac{1}{\Phi} - (1 + \beta), \quad \varphi_0 \ll \Phi \ll \frac{\varphi_0^2}{\gamma}, \tag{8}
\]
followed by an exponentially small decay.
Discussion. – The intermittency is the reason of a strong adverse effect of the fluctuations on network feedback. We illustrate this for the feedback mechanism provided by an idealized transmission control protocol (TCP) which handles most of data transfer. It works in cycles, each consisting of sending a group of $W$ packets from the source and receiving acknowledgments of their delivery at the destination [32]. The transmission rate equals $W/t_0$ where the cycle duration $t_0$ is normally the round trip time. If a loss is detected during the cycle, the protocol halves the transmission rate of service could be tens of seconds.

To estimate the protocol operation time scale note that the memory buffer size $c_i$ of any link is related to its capacity (maximal sending rate) $r_i$ by the engineering “rule of thumb” [32], $c_i = t_0r_i$, ensuring any full buffer to empty during the same time $t_0$. At $t_i r_i \lesssim 1$ at the congestion threshold, we find $t_i / t_0 \gtrsim 1/c_i$. We show this region of protocol operations as the hatched area in fig. 2.

This region is spread over many orders of magnitude in $\gamma^{-1}$ which is not directly measurable but is bounded from above, $\gamma^{-1} \lesssim c$ (the perfect design limit, as explained after eq. (7)), and from below, $\gamma^{-1} \gg aW$, as determined by the condition $\langle \Phi \rangle_c \sim a \gamma \ll 1/W$ (otherwise the network would be permanently congested). As $c \sim 10^6$ for the typical buffer size (in packets) [32], and $a \sim 10$ for the average number of links in an end-to-end path across the Internet [6,17], we see that $aW$ is close to $\sqrt{c}$, so
that almost the entire hatched area corresponds to the mesoscopic mode of intermittent losses.

On the macroscopic time scale (at the bottom of the hatched area in fig. 2) the probability $\langle \Phi^2 \rangle_c$ of detecting two loss indicators in two consecutive cycles is equal to $\langle \Phi^2 \rangle_c$. However, the ratio $\langle \Phi^2 \rangle_c / \langle \Phi \rangle_c^2$ increases with $\gamma^{-1}$ reaching $c^1-\beta/a$ at $\gamma^{-1} \sim c$, where it varies from $10^2$ for $\beta = \frac{1}{2}$ to $10^3$ for $\beta \rightarrow 0$. Such a ratio for detecting three loss indicators in three consecutive cycles is even more striking, varying from $10^6$ to $10^9$. Conversely, the same multiple-loss indicators may correspond to very different average losses $a \gamma$. And it is the time-independent average loss which does matter as the intermittent fluctuations would die out with time as operations move from the dangerous dark-shaded (red online) area to the safe light-shaded (green online) one.

**Conclusion.** — We stress that the above discussion of the idealised TCP feedback is just an example of a possible network overreaction to the intermittent fluctuations. The latter would arise in any network provided that i) link (or node) operations can be described by a finite-capacity model which will suffer from local congestion fluctuations at the threshold of capacity and on a mesoscopic time scale, due to inevitable input noise; ii) the network has a PL load distribution which greatly enhances the fluctuations on highly-loaded network elements.

The identification of a new, mesoscopic time scale for data loss in heterogeneous networks and establishing the existence of strong fluctuations of losses exacerbated by the heterogeneous nature of the network constitute the main results of this letter. The fluctuations become most dangerous, leading to a possible operational instability of the network, when they are misinterpreted by a network feedback mechanism as described above. Such mechanism is network-specific and whether it triggers fluctuation-induced congestion requires ad hoc considerations; however, the danger of overreaction would be imminent if the reaction time were comparable to the mesoscopic time.

***

This work was supported by the EPSRC Grant EP/E049095/1 and by the Leverhulme grant RPG-380.

REFERENCES

[1] Watts D., Proc. Natl. Acad. Sci. U.S.A., 99 (2002) 5766.
[2] Motter A. E. and Lai Y.-C., Phys. Rev. E, 66 (2002) 065102.
[3] Motter A. E., Phys. Rev. Lett., 93 (2004) 098701.
[4] Moreno Y., Pastor-Satorras R., Vázquez A. and Vespignani A., Europhys. Lett. 62 (2003) 292.
[5] Wang B. and Kim B. J., EPL, 78 (2007) 48001.
[6] Barrat A., Barthélemy M. and Vespignani A., Dynamic Processes on Complex Networks (Cambridge University Press, Cambridge) 2008.
[7] Ashtab D. J., Jarrett T. C. and Johnson N. F., Phys. Rev. Lett., 94 (2005) 058701.
[8] Buldyrev S. V., Parshani R., Paul G., Stanley H. E. and Havlin S., Nature, 464 (2010) 1025.
[9] Vespignani A., Nature, 464 (2010) 984.
[10] Arenas A., Díaz-Guilera A. and Guimerà R., Phys. Rev. Lett., 86 (2001) 3196.
[11] Freeman L., Sociometry, 40 (1977) 35.
[12] Newman M. E. J., Phys. Rev. E, 64 (2001) 016132.
[13] Girvan M. and Newman M., Proc. Natl. Acad. Sci. U.S.A., 99 (2002) 7821.
[14] Faloutsos M., Faloutsos P. and Faloutsos C., Comput. Commun. Rev., 29 (1999) 251.
[15] Barabási A.-L. and Albert R., Science, 286 (1999) 509.
[16] Pastor-Satorras R., Vázquez A. and Vespignani A., Phys. Rev. Lett., 87 (2001) 258701.
[17] Albert R. and Barabási A.-L., Rev. Mod. Phys., 74 (2002) 47.
[18] Pastor-Satorras R. and Vespignani A., Evolution and Structure of the Internet (Cambridge University Press, Cambridge) 2004.
[19] Goh K., Kahng B. and Kim D., Phys. Rev. Lett., 87 (2001) 278701.
[20] Goh K., Oh E., Jeong H., Kahng B. and Kim D., Proc. Natl. Acad. Sci. U.S.A., 99 (2002) 12583.
[21] Barahelény M., Phys. Rev. Lett., 91 (2003) 189803.
[22] Goh K.-I., Ghim C.-M., Kahng B. and Kim D., Phys. Rev. Lett., 91 (2003) 189804.
[23] Boccaletti S., Latora V., Moreno Y., Chavez M. and Wang D. U., Phys. Rep., 424 (2006) 175.
[24] Dorogovtsev S. N., Goltsev A. V. and Mendes J. F. F., Rev. Mod. Phys., 80 (2008) 1275.
[25] Dimitropoulos X. A., Krioukov D. V. and Riley G. F., Passive and Active Network Measurement, edited by Dovalolis C., Lect. Notes Comput. Sci., Vol. 3431 (Springer, Berlin) 2005, p. 177.
[26] Vázquez A., Pastor-Satorras R. and Vespignani A., Phys. Rev. E, 65 (2002) 066130.
[27] Stepanenko A. S., Constantinou C. C., Avantis T. T. and Baughan K., Electron. Lett., 38 (2002) 350.
[28] Cao J., Cleveland W. S., Lin D. and Sun D. X., Lect. Notes Stat., 171 (2003) 83.
[29] de Menezes M. A. and Barabási A.-L., Phys. Rev. Lett., 92 (2004) 028701.
[30] http://www.caida.org/research/traffic-analysis, packet size distribution comparison between Internet links in 1998 and 2008.
[31] Stepanenko A. S., Constantinou C. C., Yurkevich I. V. and Lerner I. V., Phys. Rev. E, 77 (2008) 046115.
[32] Kurose J. F. and Ross K. W., Computer Networking: A Top-Down Approach (Addison-Wesley, Boston) 2010.