A Poisson cluster model for telecommunications [1]

THOMAS MIKOSCH

UNIVERSITY OF COPENHAGEN
Laboratory of Actuarial Mathematics
Universitetsparken 5
2100 Copenhagen
Denmark

www.math.ku.dk/∼mikosch

Joint work with
Gilles Faÿ, Barbara González-Arévalo, Gennady Samorodnitsky

[1] Cherry Bud Workshop, March 27–30, 2006

1. Facts

- Since the beginning of the 1990s models have been proposed for large communication networks (Internet, local area networks,...).
- Classical queuing models for waiting and service times fail to explain typical behavior.
- There is general agreement that the process of active sources at a given time t exhibits *long range dependence.* This notion only makes sense for stationary processes.
- The integrated process (workload) is believed to be well approximated by a *self-similar* process (such as fractional Brownian motion, stable Lévy motion).
- Although the expected workload is growing roughly linearly through time (such as in classical queuing networks) there are strong deviations from linearity due to erratic behavior.

- Since work by Taqqu, Willinger, Leland,... (1993–) and others the assumption of *heavy tailed distributions* for file sizes, transmission durations, transmission rates,... has been accepted as a reasonable working hypothesis.
- There exists rather convincing evidence that file sizes, transmission durations, transmission rates,... have Pareto like distributions:

$$P(X_t > x) \approx x^{-\alpha}, \quad x \to \infty .$$

- Given the stationarity of the process of active sources, α is often found to be between 1 and 2. (infinite variance)

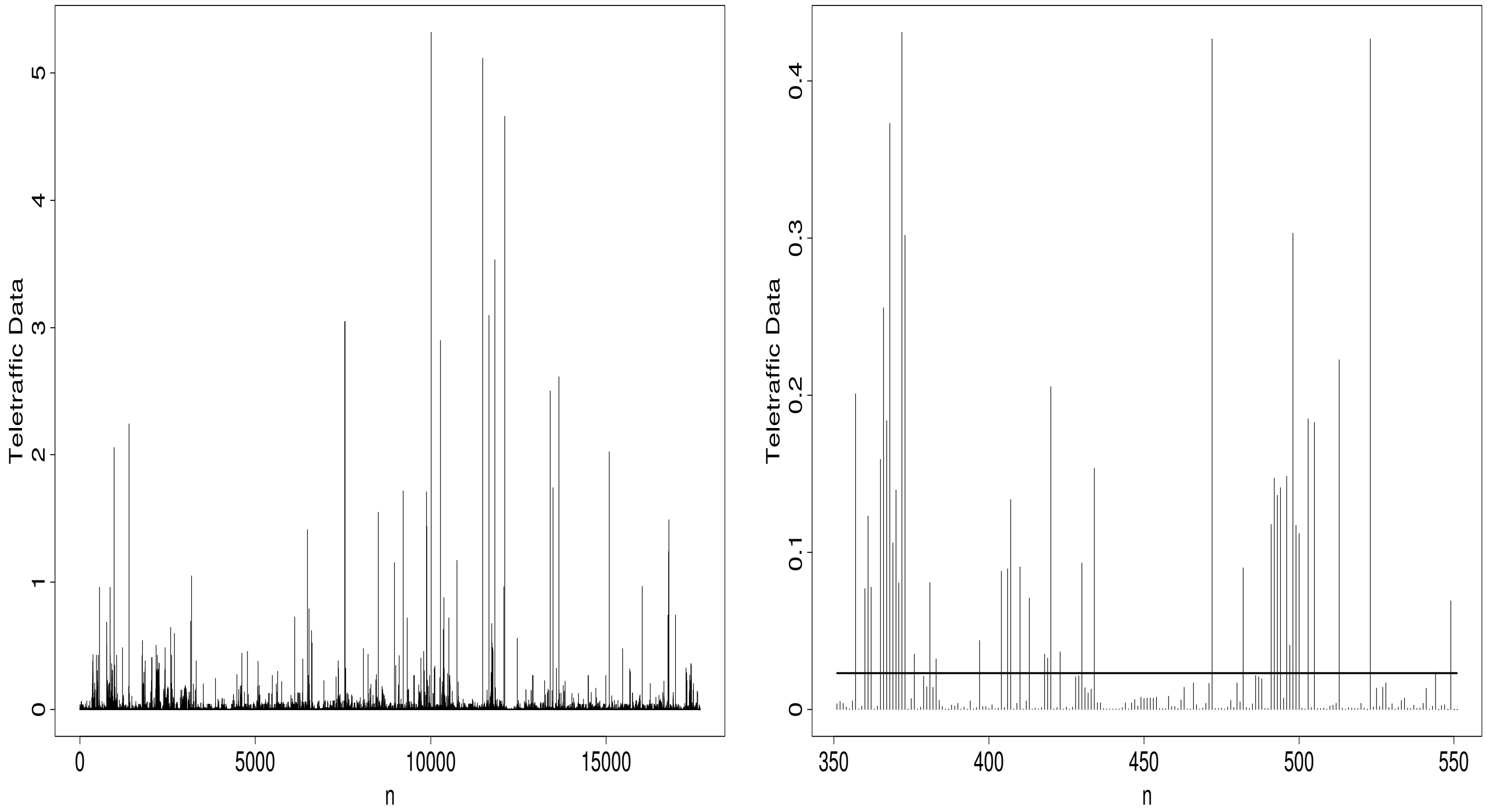

Figure 1. Time series of transmission durations.

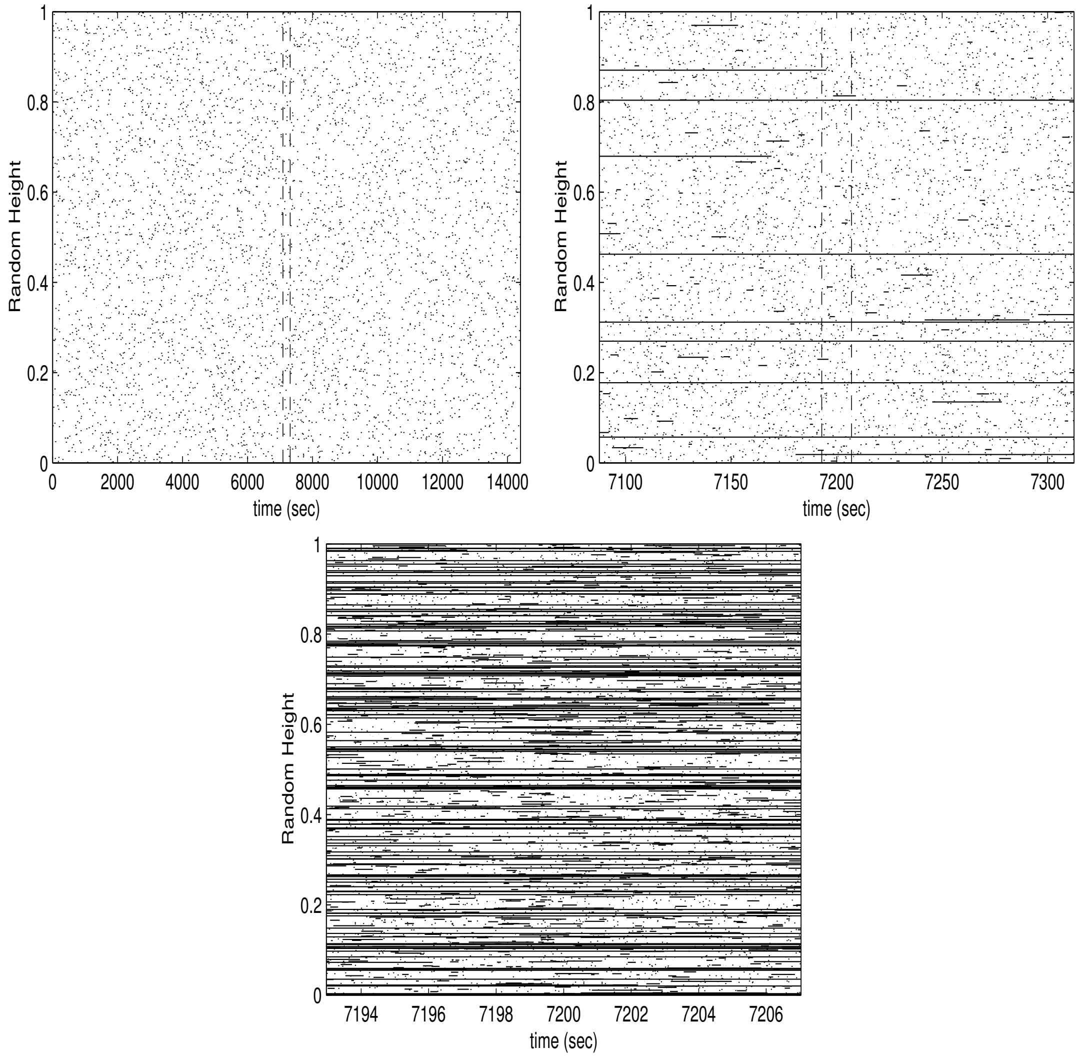

FIGURE 2. Mice and elephants plots (S. Marron).

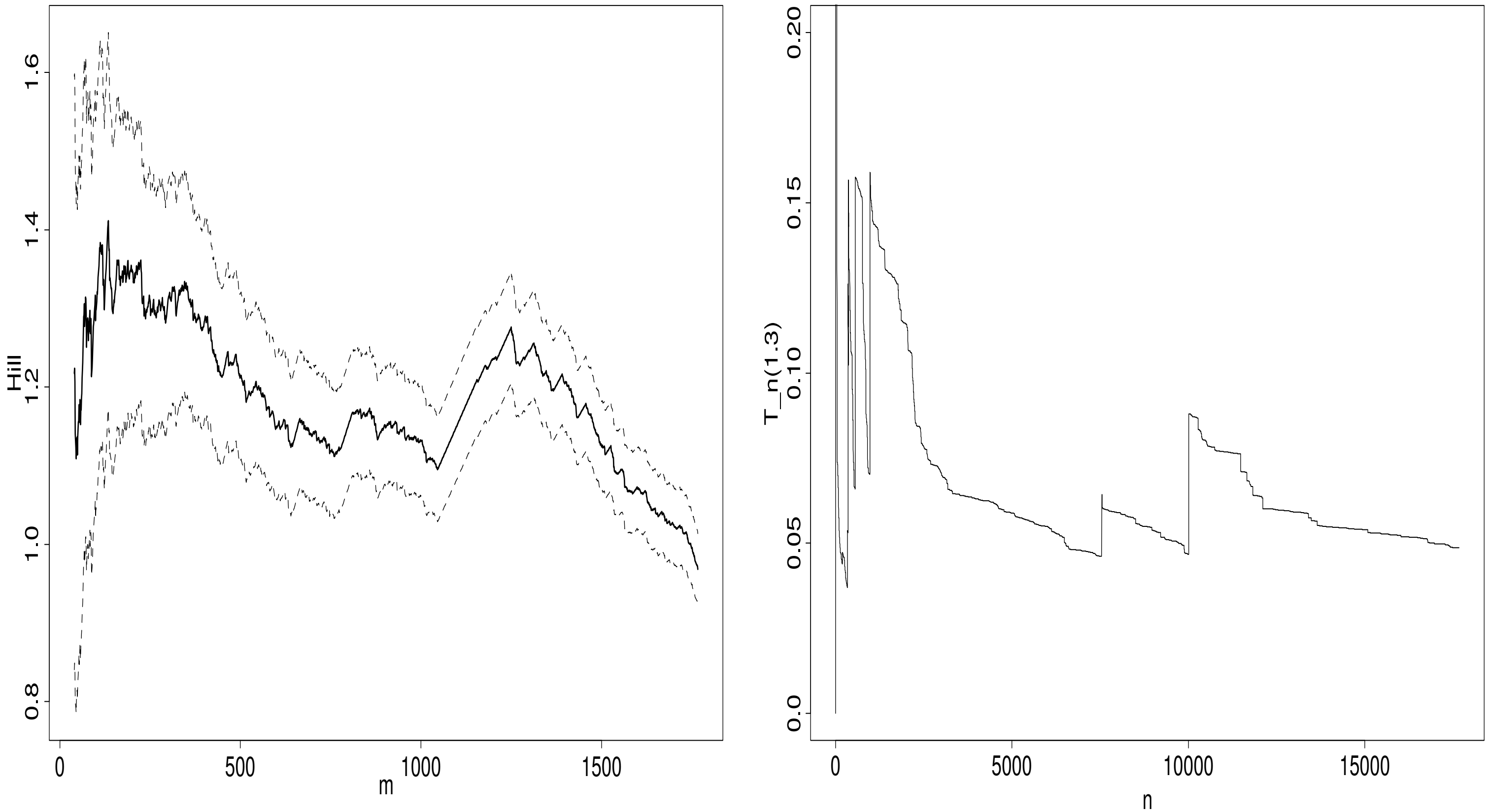

FIGURE 3. Methods for determining α.

2. Basic models

- Communication networks are too complex to be understood in detail.
- They are run by machines which are very fast (in contrast to human beings) and therefore fail a lot (in contrast to human beings who can use their brains).
- Although we do (perhaps) understand a single machine (car) and we know that the machines' joint behavior (Autobahn) is directed by a protocol (traffic lights) we do not understand their interplay (e.g. traffic jam).
- Therefore any model is nothing but a simplistic proxy to reality.
- But a "realistic" model should to some extent explain the observed facts (self-similarity of workload, long range dependence of activity process, heavy tailed distributions).

2.1. **Standard models for the process of active sources.**

2.1.1. *The ON/OFF process.*

- During a transmission, a source transmits at unit rate. Otherwise, it is silent.
- Lengths of ON and OFF periods are described by two independent iid sequences of positive random variables.
- The ON periods have heavy tailed distribution.
- See Taqqu, Willinger, Leland,... (1993-1995), Heath, Resnick, Samorodnitsky (1998), Mikosch, Resnick, Rootzén, Stegeman (2002).
- The activity of the network is understood as the superposition of a large number of iid ON/OFF sources.

2.1.2. *The infinite source Poisson model.*

- Transmission initiations or connections of sources happen at the points of a rate λ homogeneous Poisson process
$$\cdots < \Gamma_{-1} < \Gamma_0 < 0 < \Gamma_1 < \Gamma_2 < \cdots .$$
- Transmission durations are iid random variables Y_i, independent of (Γ_i).
- During a transmission a source transmits at unit rate.
- The stationary *process of active sources* at time t
$$M_t = \textstyle\sum_{i\in\mathbb{Z}} I_{\{\Gamma_i \le t < \Gamma_i + Y_i\}}\,, \quad t \ge 0\,.$$
- Since the points (Γ_i, Y_i) constitute a PRM$(\lambda\mathrm{Leb} \times F_Y)$, a simple calculation shows
$$\gamma(h) = \mathrm{cov}(M_0, M_h) = \lambda \int_h^\infty \overline{F}_Y(t)\,dt\,.$$

- If $\overline{F}_Y(t) = L(t)t^{-\alpha}$, $\alpha > 1$, for some slowly varying L, by Karamata's theorem,
$$\gamma(h) \sim (\alpha-1)^{-1}\, h\, \overline{F}_Y(h)\,, \quad h \to \infty\,.$$
- Non-summability of γ for $\alpha \in (1,2)$ is interpreted as *long range dependence*. The *Hurst coefficient* is $H = (3-\alpha)/2 \in (0.5, 1)$.
- The *workload process*
$$A(t) = \int_0^t M_s\, ds\,, \quad t \geq 0\,,$$
has stationary increments.
- For $\alpha \in (1,2)$ scaling limits of $(A(Tt))_{t\geq 0}$ converge to spectrally positive α-stable Lévy motion. (infinite variance, independent increments)

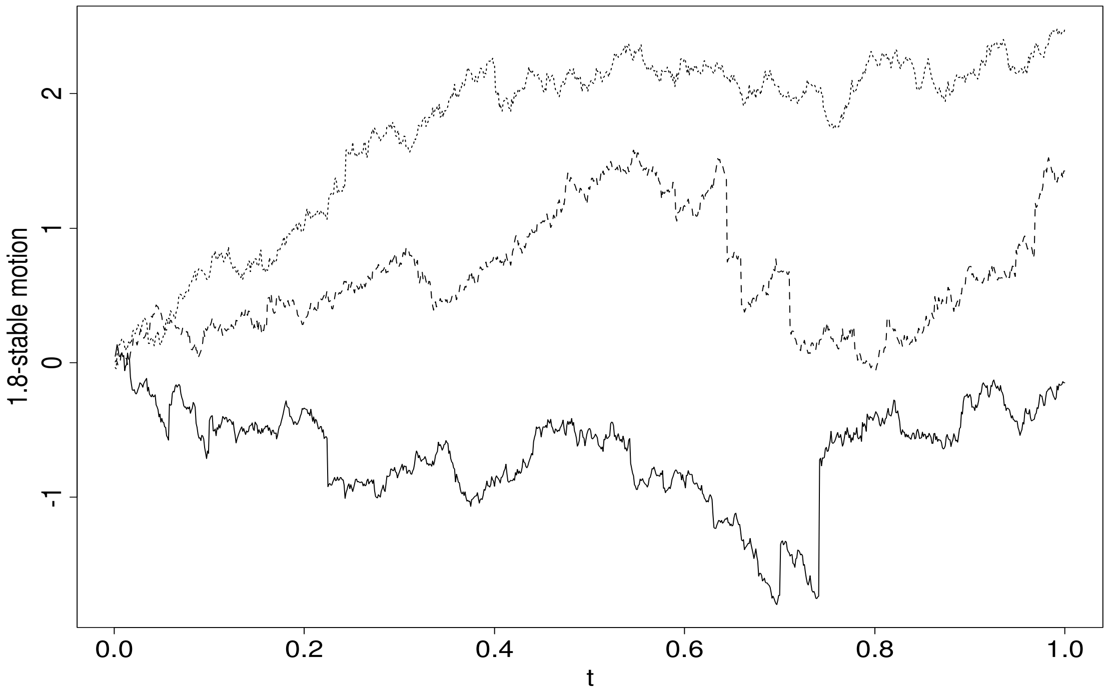

FIGURE 4. 1.8-stable sample paths.

- Letting the intensity $\lambda = \lambda_T$ grow sufficiently fast, scaling limits of $(A(Tt))_{t\geq 0}$ converge to fractional Brownian motion B_H with Hurst index $H = (3-\alpha)/2$ and covariance structure

$$\operatorname{cov}(B_H(t), B_H(s)) = 0.5(t^{2H} + s^{2H} - |t-s|^{2H})\,.$$

- See Mikosch, Resnick, Rootzén, Stegeman (2002).
- Fractional Brownian motion B_H with $H \in (0.5, 1)$ inherits long range dependence for the increment process $B_H(h) - B_H(h-1)$.
- Similar results exist for *superpositions of ON/OFF processes* given the number $M = M_T$ of superimposed processes grows sufficiently fast with T.
- *High frequency of arrivals* is generated either by increasing the intensity λ_T of the Poisson process *or* the number M_T of ON/OFF sources.

- If λ_T or M_T increase too slowly α-stable Lévy motion appears in the limit.

3. The Poisson cluster process

- At the points Γ_i of a rate λ homogeneous Poisson process on $\mathbb{R}$ the first packet of the ith flow (ith activity) arrives.
- The ith flow of packets consists of K_i packets which arrive at times

$$Y_{ik} = \Gamma_i + S_{ik} = \Gamma_i + \sum_{j=1}^{k} X_{ij}\,, \quad 0 \le k \le K_i\,.$$

- $(X_{ik})_{i,k}$ are iid, (K_i) are iid; (X_{ik}), (K_i), (Γ_i) are independent.
- The counting process

$$N(B) = \#\{(i,k) : i \in \mathbb{Z}\,, 0 \le k \le K_i : Y_{ik} \in B\}$$

is stationary.
- Let $0 \le T_1 \le T_2 \le \cdots$ be an enumeration of the points of N.

- For statistical analyses one cannot distinguish between the arrivals $Y_{i0} = \Gamma_i$ and Y_{jk}, $k \geq 1$.
- *Notice:* The points $(\Gamma_i, K_i, (X_{ik})_k)$ constitute a PRM$(\lambda \,\text{Leb} \times F_K \times F_X^\infty)$, N^*, in $\mathbb{R} \times \mathbb{N}_0 \times \mathbb{R}^\infty$.
- and

$$N(a,b] = \int_{\mathbb{R}\times\mathbb{N}_0\times\mathbb{R}^\infty} \textstyle\sum_{j=0}^k I\{\gamma + \sum_{i=0}^j x_i \in (a,b]\} dN^*(\gamma, k, (x_i))\,.$$

3.1. How can we get long range dependence for the increments $N(h, h+1]$?

- If $\mathrm{var}(K) < \infty$

$$\int_1^\infty \gamma_N(h)\,dh$$

$$= \lambda E \sum_{k=1}^{K} (K-k+1) \int_0^1 (x \wedge (2-x))\overline{F}_{S_k}(x)\,dx < \infty\,,$$

for the generic renewal process $S_k = X_1 + \cdots + X_k$.

- Long range dependence is impossible unless $\mathrm{var}(K) = \infty$ *whatever the distribution of* X.
- This is in agreement with teletraffic measurements: S_K *is large due to a large number* K.
- A *weighted renewal argument* Alsmeyer (1992) yields

$$\gamma_N(h) \sim \lambda\,(EX)^{\alpha-2}(\alpha-1)^{-1}h\,P(K > h)\,,$$

if $P(K > x) = x^{-\alpha}L(x)$ for some $\alpha \in (1,2)$.

3.2. **Where do the heavy tails of S_K come from?**

- S_K can be large due to large K or large X.
- $P(X > x) = x^{-\alpha}L(x)$, $EK < \infty$ and $P(K > x) = o(P(X > x))$. Then

$$P(S_K > x) \sim EK\, P(X > x)\,.$$

- $P(K > x) = x^{-\beta}L(x)$ for some $\beta \geq 0$, $EX < \infty$ and $P(X > x) = o(P(K > x))$. Then

$$P(S_K > x) \sim (EX)^{\beta} P(K > x)\,.$$

- The assumptions are close to necessity.

3.3. **Asymptotic results.**

- $N(t) = N[0,t]$ satisfies the strong law of large numbers $N(t)/t \stackrel{\text{a.s.}}{\to} \lambda\,(EK+1)$, see figure.
- Scaling limits are either Brownian motion (if $\mathrm{var}(K) < \infty$) or α-stable Lévy motion (if $P(K > x)$ is regularly varying with index $-\alpha \in (-2,-1)$ and $EX < \infty$).
- This is *disappointing* but similar to the workload in the ON/OFF and infinite source Poisson cases.
- One starts with long range dependent increments (if $\mathrm{var}(K) = \infty$) and loses them in the limit: Lévy motion has independent increments.
- One even loses the notion of long range dependence in the narrow sense: the limit has infinite variance.

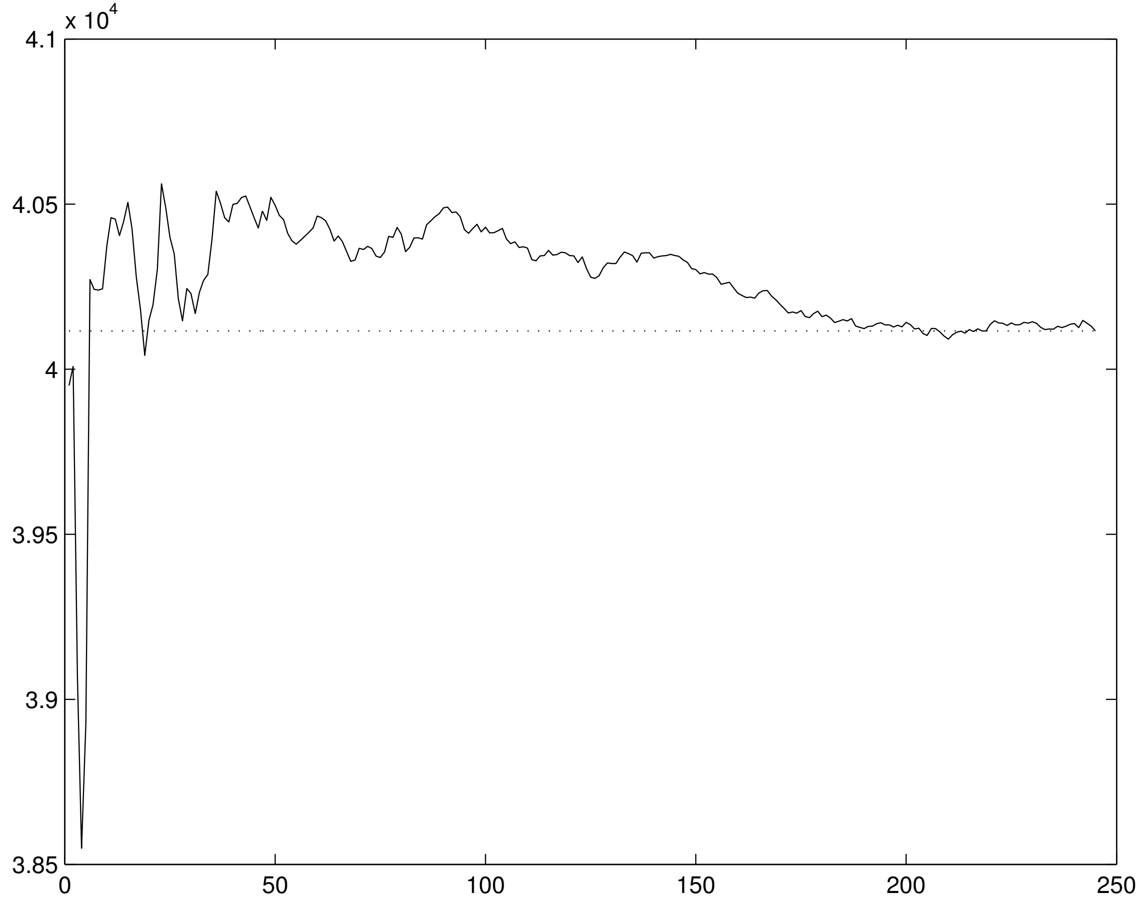

FIGURE 5. $N(t)/t$ for 10^7 packet arrivals (245 seconds) at UNC.

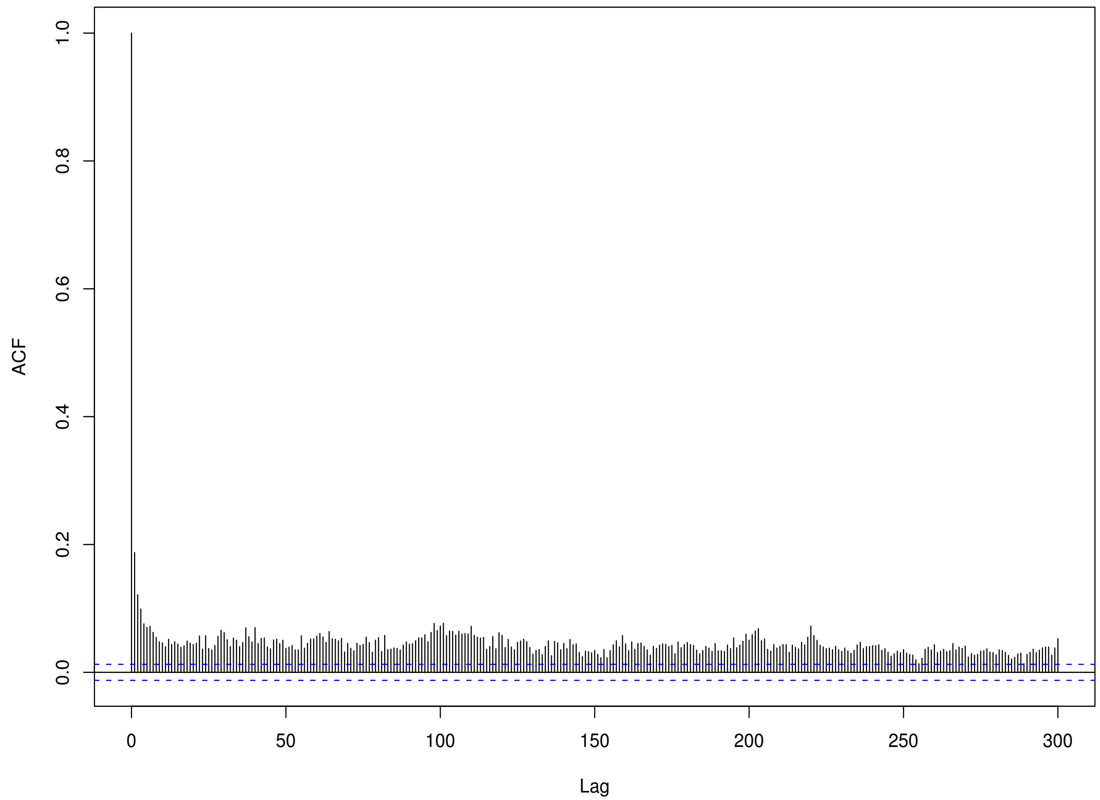

FIGURE 6. Sample ACF of the increments of the UNC packet arrival data.

3.4. **How can one overcome this problem?** Mikosch and Samorodnitsky (2006).

- *Partial answer:* For a truncated cluster point process $N_0[0,t] \wedge (K+1)$ with regularly varying $P(K > x)$ with index $-\alpha \in (-2,-1)$, the expected arrivals $ET_n^{(0)}$ have to grow faster than n^{α}.
- Or one has to increase the intensity $\lambda_T \to \infty$ at some rate.

3.5. **Can one estimate the parameters of the model?**

- Faÿ, Roueff, Soulier (2005) have developed *wavelet estimation techniques* (local Whittle estimation of α). See also Hohn, Veitch, Abry (2003) for some empirical studies.

- For a bounded real-valued function ψ with support $[0,1]$, the *wavelet coefficients* are defined as

$$d_{jk} = 2^{-j/2} \int_0^\infty \psi(2^j s - k)\, N(s)\, ds\,, \quad j \in \mathbb{Z}\,, k \in \mathbb{Z}\,.$$

  The set of available wavelet coefficients:

$$\Delta = \{(j,k) : 0 < J_0 < j \le J_1\,, 0 \le k \le 2^{J-j} - 1\}$$

  The reduced local Whittle contrast function

$$W(\alpha') = \log \sum_{(j,k)\in\Delta} \frac{d_{jk}^2}{2^{(2-\alpha')j}} + \text{const}\,.(2-\alpha')\,.$$

- In the infinite source model Faÿ, Roueff, Soulier (2005) show consistency of the local Whittle estimate under additional conditions on J_0, J_1, J and derive rates of convergence.
- In the Poisson cluster model the method works well, see figure.

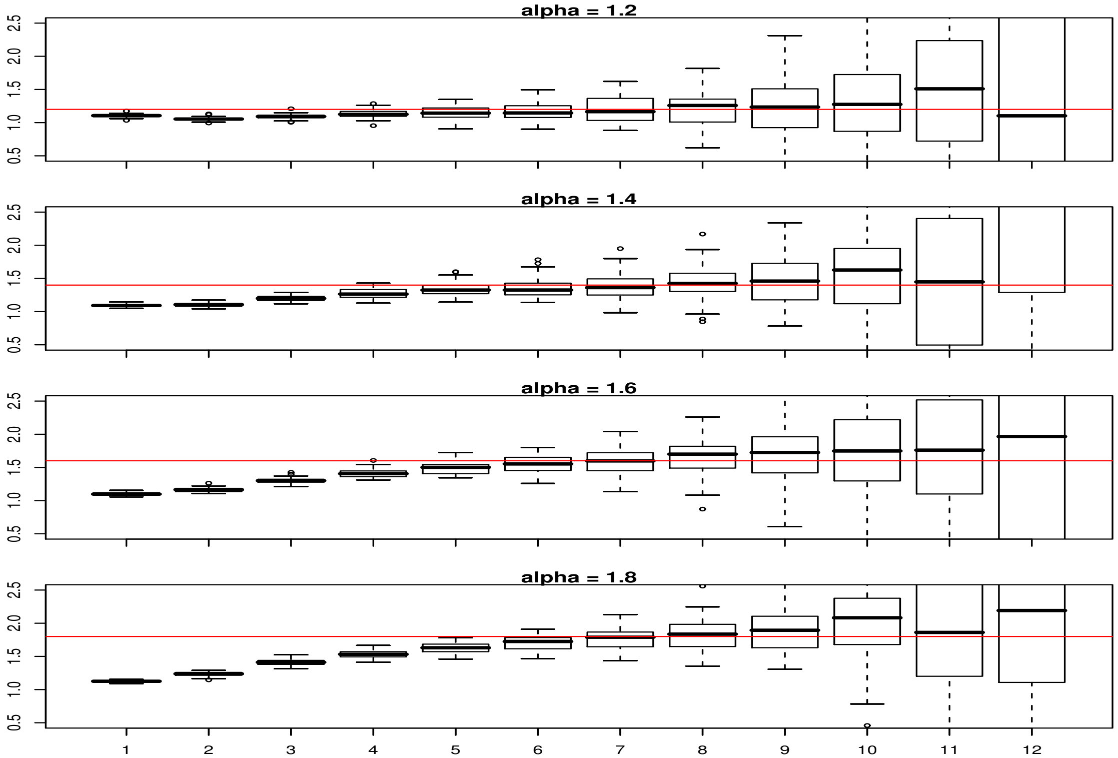

FIGURE 7. Estimation of α from simulated processes N.

3.6. **The distribution of the interarrival times under the Palm measure.**

- Under the Palm measure, $T_0 = 0$ a.s. and the interarrival times of the non-decreasing enumeration $0 \le T_1 \le T_2 \le \cdots$ of the non-negative points of N constitute a stationary ergodic process.
- For the distribution F_0 of the interarrival times under the Palm measure, (Palm-Khintchine)

$$\begin{aligned} P(T_1 > t) &= \lambda\,(EK+1)\int_1^\infty \overline{F}_0(x)\,dx \\ &= \exp\{-\lambda(t + EK\int_0^t \overline{F}_X(x)\,dx)\}\,. \end{aligned}$$

- After differentiation,

$$\overline{F}_0(t) = \frac{1 + EK\,\overline{F}_X(t)}{EK+1}\exp\{-\lambda\,(t + EK\int_0^t \overline{F}_X(x)\,dx)\}\,.$$

- Notice: $\overline{F}_0$ is nearly exponential whatever the distribution of X and K.

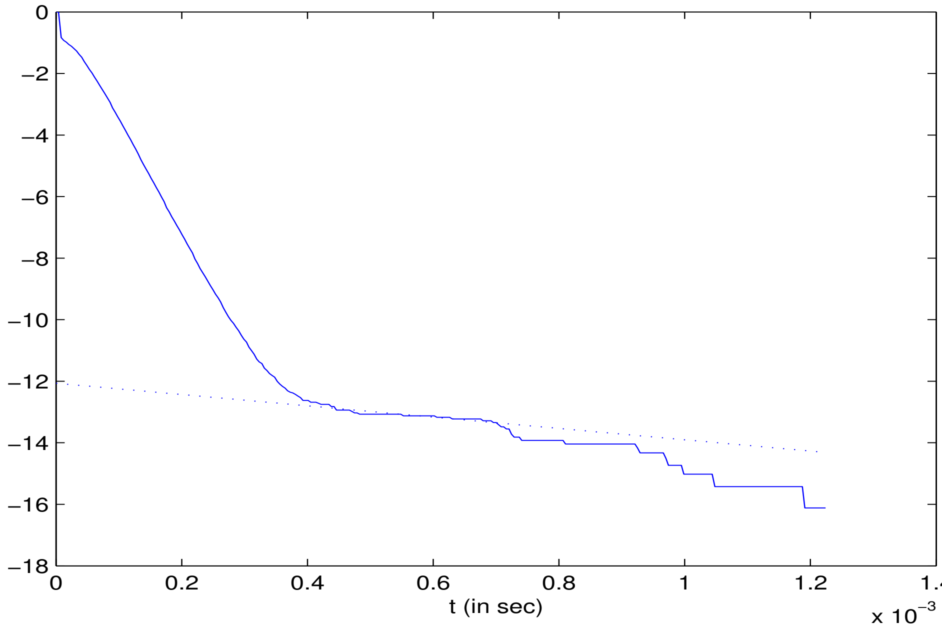

FIGURE 8. Estimation of λ from UNC data by regression from $\log \overline{F}_n$.

4. Some conclusions

- Models for teletraffic are (too) simple. One would wish to incorporate effects of the protocol or the interaction between different sources.
- At the moment no better models are availabe.
- The phenomenon of heavy-tailed distributions for file sizes, transmission durations, transmission rates, etc., is a well accepted fact and should be part of the model.
- Heavy tails give a plausible explanation of the long-range dependence of the process of active sources.
- The statistics of teletraffic data depend on the models available.
- The statistics inside these models is non-trivial and needs further efforts.
- It would be interesting to investigate whether suitable time series models for telecommunications can be developed.