On Propagation of Self-Similar Traffic Through an Energy-Conserving Wireless Gateway

Jie Yu, Athina P. Petropulu

Department of Electrical and Computer Engineering
Drexel University, Philadelphia, PA, 19104, USA

This work has been supported by ONR under grant N00014-03-1-0123
High-speed wireline traffic has been well recognized to be self-similar and bursty over a large range of time scales. [Lowen, Teich, 1993] [Willinger, Taqqu, 1997]

It is in stark contrast with the traditional teletraffic described by certain Markovian models. For example, the buffering becomes inefficient.

Several models have been proposed to capture these characteristics in recent works. [Willinger, Taqqu, 1997] [Yang, Petropulu, 2002] [Yu, Petropulu, 2005]

With the increasing demands for wireless internet access and the fast evolution of wireless techniques, high-speed services will soon be provided via wireless networks.

Recently, there have been some works suggesting that wireless traffic may also exhibit self-similarity. [Liang, 2002]
Motivation

Just because the demands of wireless and wireline applications are similar doesn’t imply that wireless and wireline traffic have similar characteristics.

There are some fundamental differences between wireline and wireless transmissions: 1) wireless channel is more vulnerable compared with wired lines and thus limits the channel rate; 2) the constraints on transmission power by mobiles.

We here propose an analytic buffering model for energy-conserving gateways.

We feed the model with self-similar traffic and study the statistics of the output traffic.

We identify cases where self-similarity can survive or disappear.
Self-similarity (SS)

Self-similarity and fractals are referred to the phenomenon where a certain property of an object is preserved with respect to scaling in space and/or time.

Figure 1: The first 3 stages of 2-dimensional Cantor set.

Distributional Self-similarity: the real-valued process \( X(t) \) \((t \in \mathbb{R})\), is self-similar with index \( H > 0 \), if for all \( a > 0 \) and \( t \geq 0 \),

\[
a^{-H} X(at) \overset{d}{=} X(t)
\]
Long-range dependence (LRD)

\( \{X_k\} \) is a long-range dependent process with Hurst parameter \( H \), if

\[
\lim_{l \to \infty} \frac{r(l)}{l^{2H-2}} = c
\]

where \( 1/2 < H < 1 \) and \( c \) is a positive constant.

\( r(l) \) asymptotically, when \( 1/2 < H < 1 \), behaves as \( cl^{-\beta} \) for \( 0 < \beta < 1 \) where \( \beta = 2 - 2H \), and

\[
\sum_{l=-\infty}^{\infty} r(l) = \infty
\]

While most standard time-series models assume that the covariance sequence is absolutely summable. Such processes are said to be short-range dependent (SRD), e.g. ARMA.
The relationship between LRD and SS

The aggregate process of $X_k$ of degree $m$ is a running average of non-overlapping blocks of $X_k$ with length $m$:

$$X_i^{(m)} = \frac{1}{m} \sum_{k=m(i-1)+1}^{mi} X_k.$$

The aggregate of a short-range dependent process:

$$\text{Var} X_k^{(m)} \xrightarrow{m \to \infty} c_1 m^{-1},$$

The aggregate of a long-range dependent process: ($H \in (0.5, 1)$)

$$\text{Var} X_k^{(m)} \xrightarrow{m \to \infty} cm^{(2H-2)},$$

Aggregation is equivalent to time scaling

Under time scaling, a long-range dependent process is smoothed out much slower in comparison to a short-range dependent process. Thus, the time-scaled long-range dependent process maintains similarity to the original process.
Figure 2: A sample of a general On/Off process.
Mathematically, the On/Off process can be expressed as:

\[ S(t) = \sum_{j=0}^{\infty} A_j 1_{[S_j, S_j + X_j]}(t), \quad t \geq 0, \]

- \( S_j \) is a so-called regenerative point, denoting the occurrence of the \( j \)-th On period.
- \( 1_{[s_1, s_2]}(t) \) is the indicator function, which is non-zero and equal to one only for \( t \in [s_1, s_2] \).

The Hurst parameter of the On/Off process equals:

\[ H = \frac{3 - \min(\alpha_0, \alpha_1)}{2} \]

where \( \alpha_1, \alpha_0 \) are the tail indices of the On and Off durations, respectively.
Wireline traffic:

- It is an On/Off process with Pareto distributed On-/Off-durations $X_n, Y_n \ (n \in \mathbb{N})$ and cut-off Pareto distributed ON-State rates $A_n \ (n \in \mathbb{N})$, i.e. a rate-limited EAFRP [Yu, Petropulu, 2005].

- The Pareto distribution is defined in terms of its survival function as:

$$\bar{F}(x; \alpha, K) = P(X \geq x) = \begin{cases} \left( \frac{K}{x} \right)^{\alpha}, & x \geq K, \\ 1, & x < K, \end{cases}$$

- The cut-off Pareto is defined in terms of its survival function as

$$\bar{F}_L(x; \alpha, K) = \bar{F}(x; \alpha, K)(1-u(x-L)) = \begin{cases} \left( \frac{K}{x} \right)^{\alpha}, & K \leq x \leq L, \\ 1, & x < K \\ 0, & x > L \end{cases}$$
Wireless channel:
It is modelled by a two-state Markovian model of [Kim, Krunz, 2000].
- Channel capacities alternate between $c_g$ and $c_b$.
- State durations are exponentially distributed with means $1/\beta$ and $1/\gamma$.

Gateway model:
It interconnects wireline and wireless networks.
Energy-conserving Server Model

- The gateway performs desampling/packing operations on incoming traffic.
- This is a time-slotted system and during one time slot at most one packet with \( c \) information bits will be produced.
- Server model: if the amount of data in the buffer is less than \( c \), the server takes no action and waits; otherwise it sends out 1 packet.

\[
Q(n) = \begin{cases} 
S(n) + Q(n-1) - c, & 0 > \land B, \\
S(n) + Q(n-1), & 0 > \land B, \\
S(n) + Q(n-1) - c, & 0 > \land B,
\end{cases}
\]

\[
T(n) = \begin{cases} 
c, & S(n) + Q(n-1) \geq c \\
0, & S(n) + Q(n-1) < c
\end{cases}
\]

where \( < \alpha, \beta > = \max(\alpha, \beta) \) and \( \alpha \land \beta = \min(\alpha, \beta) \).
Small Buffer System \((B = c)\)

Assumptions:

(A1) The minimum rate during an On period of the incoming traffic, i.e. \(K_A\), satisfies \(K_A \ll c\). This implies moderate to low traffic load.

(A2) For a small buffer case (i.e. \(B = c\)), we ignore the previous buffer content \(Q(n - 1)\) when calculating statistics of output traffic.

(A3) The queue is stable, which implies that \(c > E[S(n)] = \frac{\mu_A \mu_1}{\mu_1 + \mu_0}\).

It holds:

\[
P(X^T > x) = P(X^T > x | A \geq c)P(A \geq c) + P(X^T > x | A < c)P(A < c)
\]

\[
P(Y^T > y) = P(Y^T > y | A \geq c)P(A \geq c) + P(Y^T > y | A < c)P(A < c)
\]

Via (A1), (A2), the dominant terms are: \(P(X^T > x | A \geq c)P(A \geq c)\)

and \(P(Y^T > y | A < c)P(A < c)\).
Small Buffer System \((B = c)\)

**Proposition 1:**
For the *small* buffer system with assumption (A1) (A2), the tail exponent of \(X^T\) is the same as that of \(X^S\). The survival function of \(Y^T\) is asymptotically power-law with tail exponent \(\min\{\alpha_1, \alpha_0\}\).

\[
P(X^T > x) \approx P(X^S > x)
\]

\[
P(Y^T > y) = \sum_{m=1}^{\infty} P\left(\sum_{i=1}^{m} X_i^S + \sum_{i=1}^{m+1} Y_i^S > y\right) \left[1 - \left(\frac{K_A}{c}\right)^{\alpha_A}\right]^m \left(\frac{K_A}{c}\right)^{2\alpha_A}
\]

More specifically,

\[
P(Y^T > y) \xrightarrow{y \to \infty} L(y) y^{-\min\{\alpha_1, \alpha_0\}}
\]

In summary, the tail indices of \(T(n)\) are related to those of \(S(n)\) as follows: \(\alpha_{X_T} = \alpha_{X_S}\) and \(\alpha_{Y_T} = \min\{\alpha_{X_S}, \alpha_{Y_S}\}\).
Large Buffer System ($B \gg c$)

- Assumption (A2) is no longer valid in the large buffer case. In this case, $T(n)$ can be in On state even when $A < C$, and thus $P(X^T > x | A < c)$ can take non-zero values.

- Via (A1), (A3), we get: $P(X^T > x) \approx P(X^T > x | A < c, Q_e < c)$ and $P(Y^T > y) \approx P(Y^T > y | A < c)$.

Proposition 2
For the large buffer system with assumption (A1)(A3), $Y^T$ has nearly the same tail index as $Y^S$, while $X^T$ is asymptotically power-low decaying with tail exponent $(\alpha_1 + 1)$, where $\alpha_1$ is the tail exponent of $X^S$.

$$P(X^T > x) \overset{x \to \infty}{\approx} C_X x^{-(\alpha_1+1)}$$

$$P(Y^T > y) \approx P(Y^S > y)$$

where $C_X = \alpha^2 C_{Q_e} K_\alpha K_1^{\alpha_1} c^{(1-\alpha_1-\alpha)}$ is a constant.
The survival functions of $X^T$ and $Y^T$ can be calculated as,

\[
P(X^T > x) = P(X^T > x \mid c = c_g)P(c = c_g)+P(X^T > x \mid c = c_b)P(c = c_b)
\]

\[
P(Y^T > y) = P(Y^T > y \mid c = c_g)P(c = c_g)+P(Y^T > y \mid c = c_b)P(c = c_b)
\]

If $B = c$ we have a small buffer system, while if $B \gg c$ we have a large buffer system; here $c$ can be $c_g$, or $c_b$.

The survival function of On-state durations:

\[
P(X^T > x) = \left(\frac{K_1}{x}\right)^{\alpha_1}P(c = c_g)+C_2x^{-(\alpha_1+1)}\frac{\beta}{\beta+\gamma} \xrightarrow{x \to \infty} \left[K_1^{\alpha_1} \frac{\gamma}{\beta+\gamma}\right]x^{-\alpha_1}
\]

where $1/\beta$, $1/\gamma$ are mean durations of good and bad states.

Thus, $\alpha_{X^T} = \alpha_1$. 
The survival function of Off-state durations:
\[ P(Y^T > y) \sim y^{-\min\{\alpha_1, \alpha_0\}} P(c = c_g) + y^{-\alpha_0} P(c = c_b) \xrightarrow{y \to \infty} y^{-\min\{\alpha_1, \alpha_0\}} \]

Thus, \( \alpha_Y^T = \min\{\alpha_1, \alpha_0\} \).

The Hurst parameter of output traffic is calculated by
\[ H^T = \frac{3 - \min\{\alpha_X^T, \alpha_Y^T\}}{2} \]

The output traffic is self-similar if \( H \in (0.5, 1) \) or \( \min\{\alpha_X^T, \alpha_Y^T\} < 2 \).

- If the incoming traffic holds that \( \alpha_1, \alpha_0 \in (1, 2) \), the output traffic is still self-similar, since \( \alpha_X^T = \alpha_1 \in (1, 2) \) and \( \alpha_Y^T = \min\{\alpha_1, \alpha_0\} \in (1, 2) \).
- If Off periods of the incoming traffic has finite variance (\( \alpha_0 = 2 \)) and the gateway is always under large buffer mode (i.e., \( B \gg c_g, c_b \)), the output traffic loses the self-similarity, since \( \alpha_X^T = \alpha_1 + 1 > 2 \) and \( \alpha_Y^T = \alpha_0 = 2 \).
Simulation Setting

- We generate the incoming traffic \( S(t) \) as an On/Off process:
  - Pareto distributed On-/Off-durations, \( \bar{F}(x; \alpha, K) \) with \( \alpha_1 = 1.6, K_1 = 1 \) and \( \alpha_0 = 1.4, K_0 = 1 \) respectively;
  - cut-off Pareto distributed On-state rates, \( \bar{F}_L(x; \alpha, K) \) with \( \alpha = 1.19, K = 48, L = 10^{4.64} \).
- The parameters have been estimated by data we collected in LAN of ECE Dept, Drexel in 2003 [Yu, Petropulu, Feb 2005]

- For wireless channel, time slot is chosen as 0.001s
  - rates alternate between two states: \( c_g = 5000 \) (bits/time slot) and \( c_b = 500 \) (bits/time slot)
  - the means of states durations are \( 1/\beta = 0.1sec \) and \( 1/\gamma = 0.0333sec \)

- The channel information rates are 1200 bit/time unit in *good* states and 290 bit/time unit in *bad* states.

- The buffer size is assumed to be \( B = 1200 \) bits.
Simulation Results–1

Small buffering system satisfies \( B = c \) \((B = 1200, c = c_g = 1200)\)

![Graph showing log-log complementary distributions](image)

\[ H^T = \frac{3 - \min(\alpha_{XT}, \alpha_{YT})}{2} = 0.8 = H^S \]

Figure 4: The log-log complementary distributions.
Large buffering system satisfies $B \gg c$ ($B = 1200$, $c = c_g = 290$)

Figure 5: The log-log complementary distributions.

$$H^T = \frac{3 - \min(\alpha_{XT}, \alpha_{YT})}{2} = 0.8 = H^S$$
Wireless gateway: alternating buffering system satisfies

\[ B = c_g \gg c_b \ (B = 1200, \ c_g = 1200, \ c_b = 290) \]

\[
H_T = \frac{3 - \min(\alpha_{X_T}, \alpha_{Y_T})}{2} = 0.8 = H^S
\]

Figure 6: The log-log complementary distributions.
Variance-time Plot

\[
\frac{\text{Var}(X^{(m)})}{\text{Var}(X^{(1)})} \sim m^{-\beta} = m^{2-2H}, \quad X^{(m)}_i = \sum_{i=(k-1)m+1}^{km} X_k
\]

- **Case 1 (left):** The self-similarity is survived, if \( \alpha_1, \alpha_0 \in (1, 2) \).
- **Case 2 (right):** The self-similarity is lost, if \( \alpha_1, \in (1, 2), \alpha_0 = 2 \) and the wireless gateway is always under large buffer model (\( B = 1200, c_g = 300, c_b = 290 \)).
Conclusion

- The proposed model can help us understand and study the effect of the gateway that feeds wireline traffic into the wireless network.

- The analysis presented here suggests that under certain conditions the self-similarity can be preserved through the gateway, while self-similarity will disappear in the case that if the On durations of input traffic are heavy-tailed while Off have finite variance and the gateway operated always under the large buffer model.