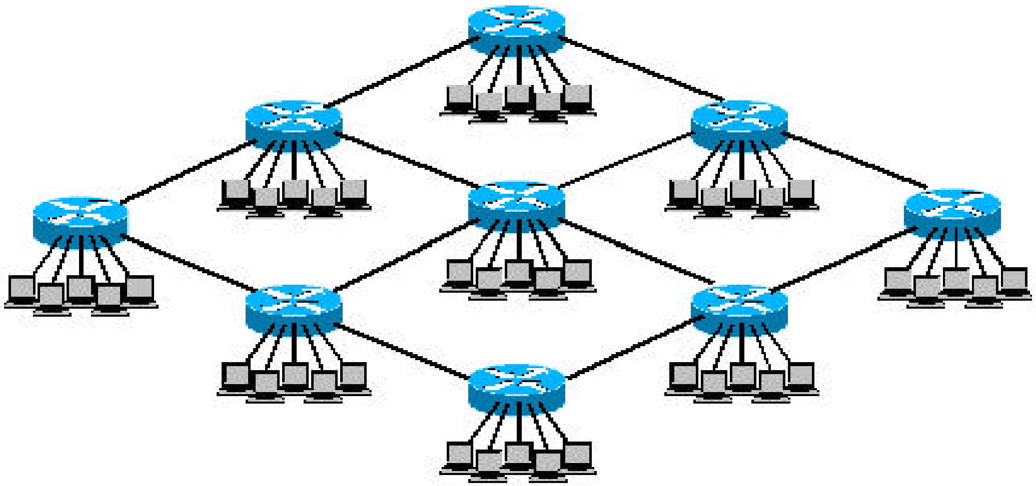
Understanding the Correlation Structure of Network Traffic



Jian Yuan, Kevin Mills

NIST/ANTD

Objective

Improving current understanding of the **correlation** structure of network traffic by identifying and studying **fundamental causalities** arising from multiple protocol layers operating in a sufficiently large network

Methodology

- Homogenizing network topology and multiple layers of protocols to achieve a sufficiently large model with well-understood parameters
- Representing in our model the essential details necessary to produce network traffic with a full-duplex “ripple effect”
- Focusing our model on describing comparative rather than absolute behavior
- Conducting a systematic search to identify, understand, and perhaps explain significant phenomena

Modeling Approach

Network Structure

- two-tiered topology: upper-tier routers and lower-tier hosts
- source hosts send at most one packet per time step, while routers forward multiple packets per time step, limited by link capacity (n_l)
- scale parameters: number of sources (n_s) and network size (L x L), e.g., L = 3 and n_s = 10 means 9 routers, and 90 pairs of connections

Traffic Sources

- alternating between wake/sleep periods with mean durations: l_{on} and l_{off}
- low (exponential ON/OFF) or high (Pareto ON/OFF with shape parameter a) user variability

Congestion-Control Algorithms

- TCP: window-based flow control, rapidly varying transmission rates, dominant in the current Internet
- TFRC: TCP friendly rate control, smoothly-changing sending rate, proposed to better serve future increase in streaming video and audio traffic

Routing

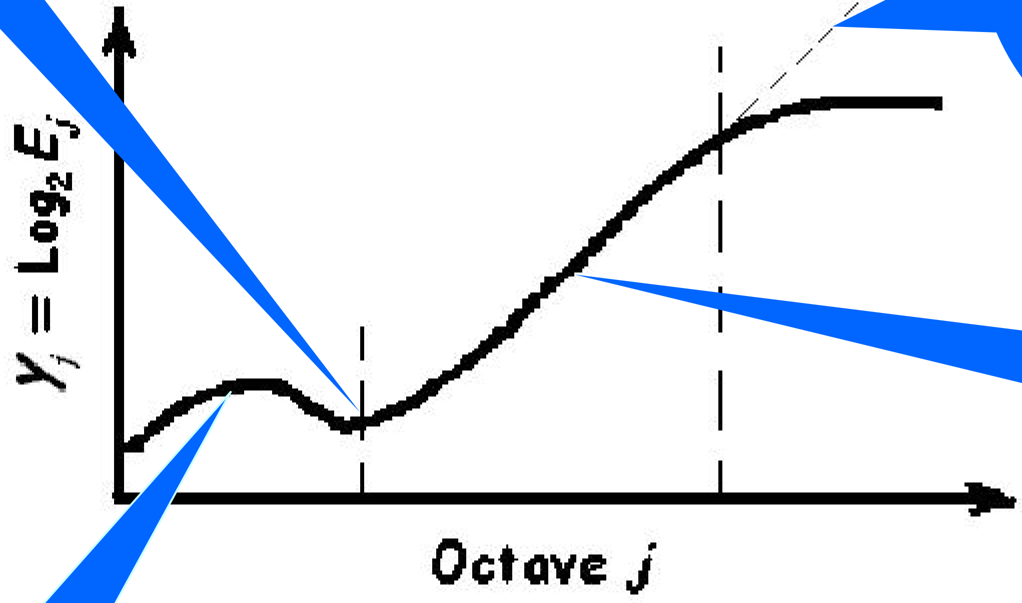
- constant, shortest path selected for each source-destination pair

Analysis Approach: Wavelet-based

$$E_j = \frac{1}{N_j} \sum_k \left| d_x(j,k) \right|^2$$

This point is larger than round-trip time

Large-time scaling phenomena are now believed to arise from the application layer property of heavy-tailed file sizes*



TCP can produce a linear part within limited time scales*

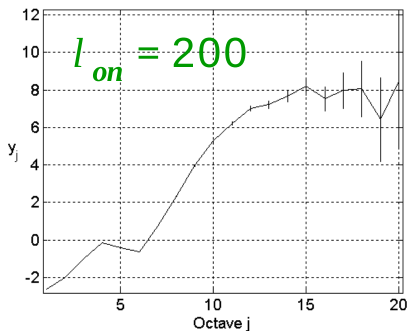
Small timescale irregular behavior arising from TCP

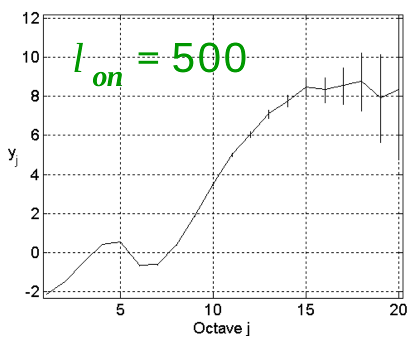
Energy Scale Plot

* Some current understanding!!!

Effect of Application Level

It was believed that ON/OFF sources with heavy-tailed ON and/or OFF periods lead to LRD in the aggregated process that is determined by the heavy-tailedness of ON or OFF periods.



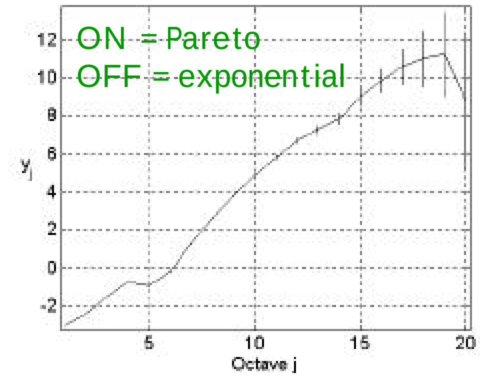


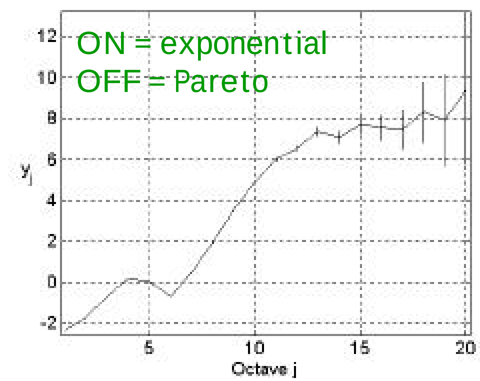
Results from our model cast doubt on current understanding that long-range dependence arises from either heavy-tailed ON or OFF traffic sources

TCP

l_{off} = 2000 and a = 1.2

$L = 3$, $n_s = 10$, and $n_l = 5$



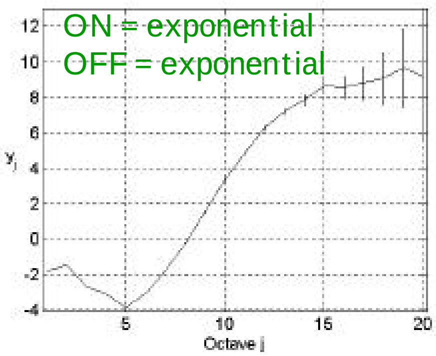


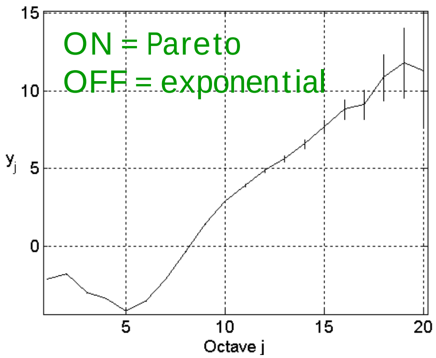
As l_{on} increases the right part of the energy-scale plot (low-frequency energy) tends to increase and the linear part tends to be reinforced

While long-range dependence appears with heavy-tailed ON periods, we do not find long-range dependence with heavy-tailed OFF periods

Effect of Transport Level

It was believed that with TCP, or flow-controlled UDP, long-range dependence of aggregate traffic is insensitive to details in the protocol stack or network configuration as long as connection durations or object sizes being transported are heavy-tailed.





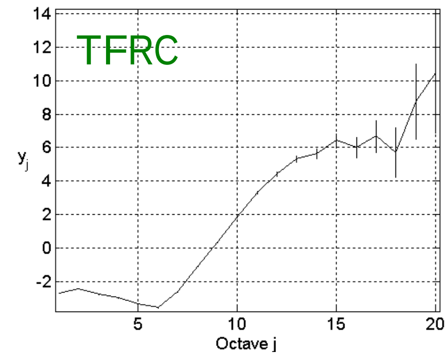
Energy scale plots for TFRC with low user variability (top) and high user variability (bottom)

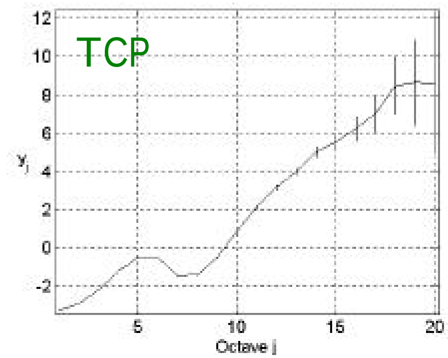
Results from our model cast doubt on current understanding that long-range dependence arises solely from heavy-tailed traffic sources

TFRC

l_{on} =200, l_{off} = 2000, and a = 1.2

L = 3, n_s = 10, and n_l = 5





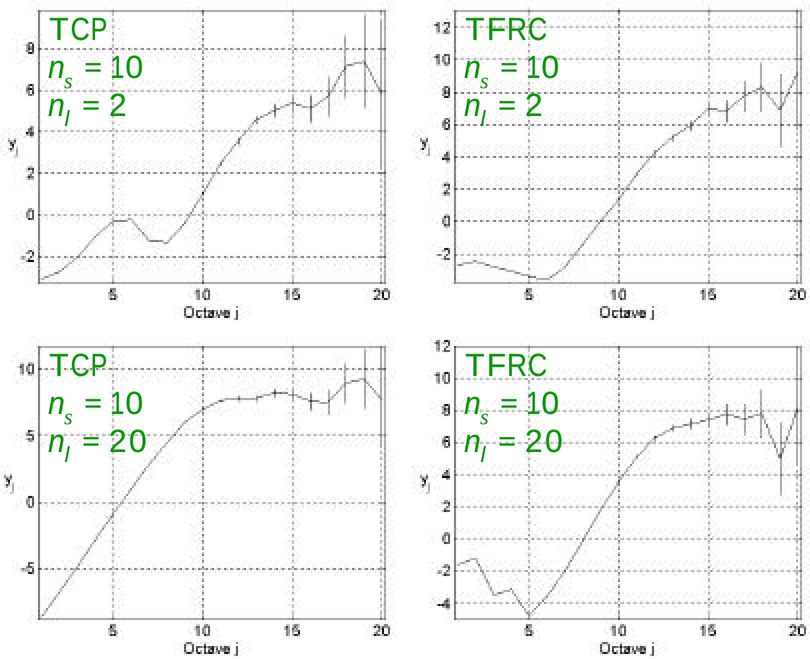
Effect of shrinking link capacity on correlation structure for TFRC (top) and TCP (bottom) with high user variability

Effect of Relative Bandwidth

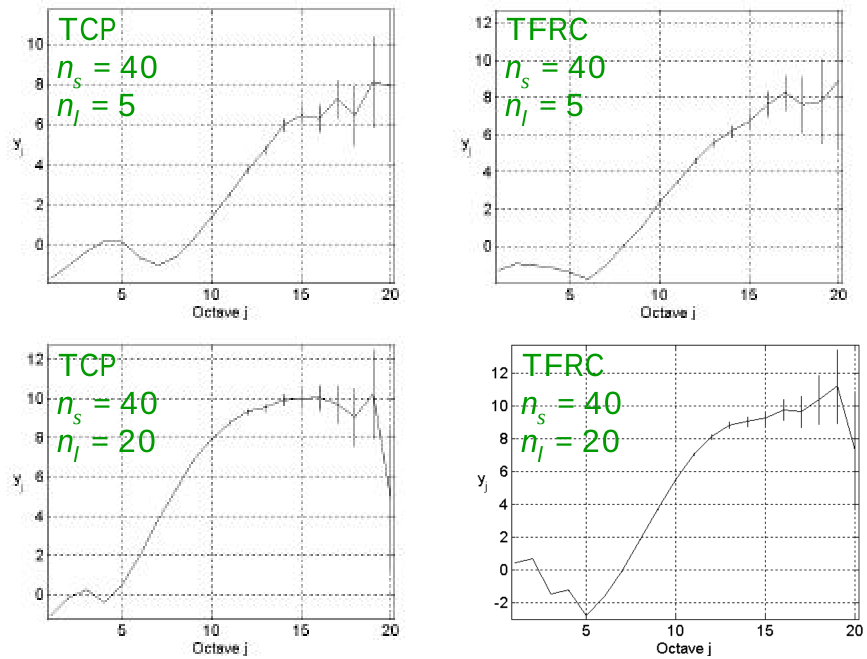
How does network congestion interact with the offered traffic? > congestion > correlation
Will shrinking or expanding link capacity alter the correlation structure? Yes.

Two Parameters: Source Count (n_s) and Link Capacity (n_l)

Exponential ON/OFF L = 3, l_{on} =200, and l_{off} = 2000



Energy scale plots for TCP (left) and TFRC (right) as link capacity increases from 2 (top) to 20 (bottom) with 10 traffic sources

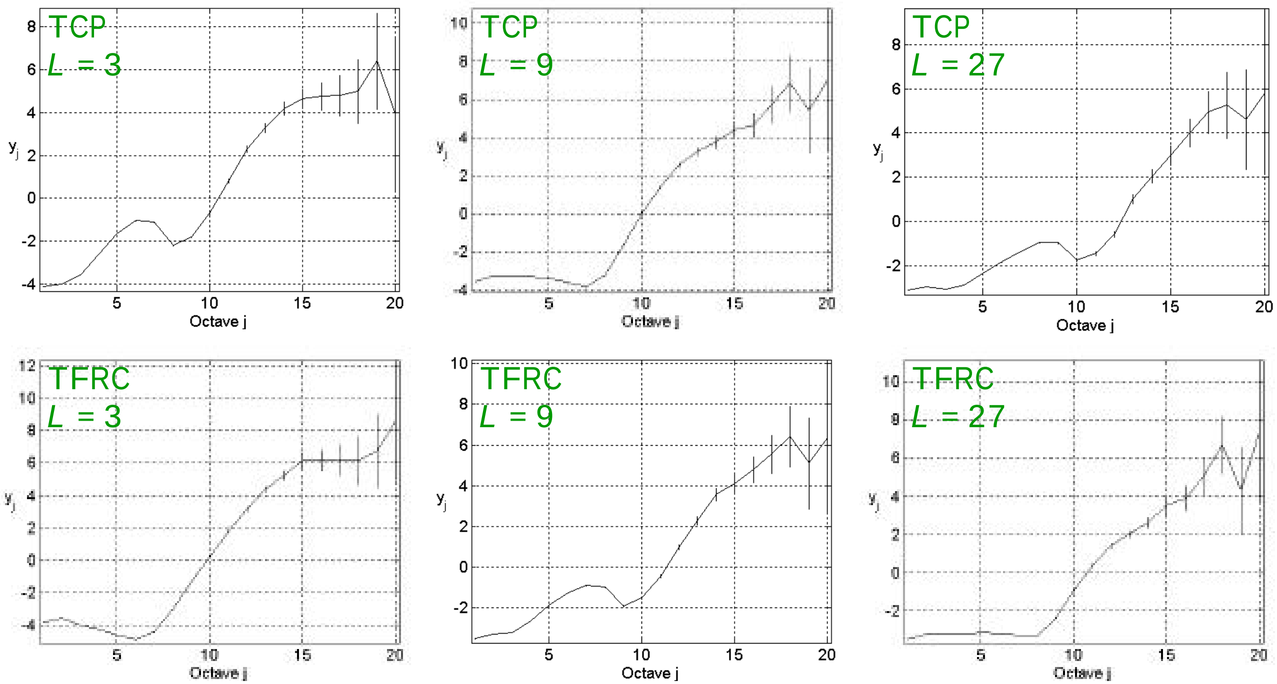


Energy scale plots for TCP (left) and TFRC (right) as link capacity increases from 5 (top) to 20 (bottom) with 40 traffic sources

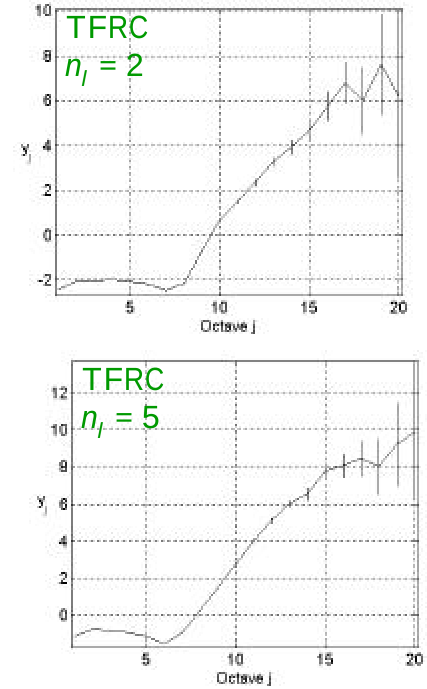
Effect of Network Size

Is correlation structure insensitive to network size? No.
Will long-range dependence continue to exist in the future Internet? Maybe.

Parameter: Network Size (L)
Exponential ON/OFF l_{on}=200, and l_{off} = 2000



Energy scale plots for TCP (top) and TFRC (bottom) with link capacity = 1 and with 5 traffic sources as network size increases from L = 3 (left) to L = 9 (middle) to L =27 (right)



Energy scale plots for TFRC at network size L = 15 and with 10 traffic sources as link capacity increases from 2 (top) to 5 (bottom)