Self-Similarity in Ethernet [Leland94a] and the Web [Crovella97a]

CSci551: Computer Networks
SP2006 Thursday Section

John Heidemann

Key Ideas

- [Leland94a]
  - self-similar network traffic
    - similar at different timescales
    - definition: has infinite variance
    - consequenes: variance decays less than exponentially; …
  - looks at Ethernet
  - described math needed to check for self-sim
  - guesses at maybe why tfc is self-sim
- [Crovella97a]
  - web traffic and TCP is self-similar and has heavy tailed on-times
    - heavy tailed: distribution has infinite variance
  - shows *why* internet tfc is self-sim

Traffic Modeling

- Paxson showed microscopic traffic effects
  - ex: reordering, "little stuff", …
- What about macroscopic traffic behavior?
  - what do traffic aggregates look like?
  - conventional wisdom:
    - traffic is generated by Poisson sources
      - or at least, that's a good approximation
    - traffic will "smooth out" at large timescales

Traffic Analysis: A Bit of History

- Phone network is homogeneous and static
  - Call arrivals at trunk groups independent
  - Interarrival times exponentially distributed
  - Call durations are exponentially distributed
- why? *people* are on both ends
- so voice traffic is
  - relatively predictable
  - very amenable to mathematical analysis
    - queueing theory (EE549)
    - with care, some of these techniques can be applied to networking as well
    - examples from class: Markov modeling in Shakih paper
  - Q: will this change with more fax & data?

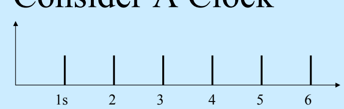
Basic Random Processes

- How to model the process arrival
  - ex. users to a computer room, or packets to a router, or
- Define an *arrival probability*
  - $a(t)$:= P[object arrives at time t] := pdf (probabilty density function)
  - $A(t)$:= P[object arrives at $t_0 < t$] := cdf (cumulative distribution function)
  - arrival rate := λ
    - $\lambda^{-1} = E[a(t)]$

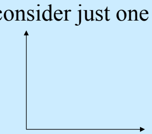
$$A(t) = \int_0^t a(t)dt$$

Consider A Clock

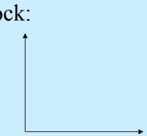
clock ticking once/second



now consider just one tick of the clock:



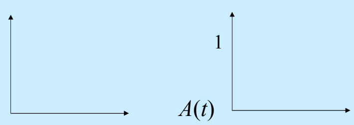
$a(t)$
P[ticks at time t]



$A(t)$
P[ticks at or before time t]

Consider A Probabilistic Clock

• assume the clock ticks uniformly randomly with ticks in the range of 0 to 2s
• note that the *rate* the clock ticks (λ) is the same as a regular clock, but exactly when the clock ticks is random

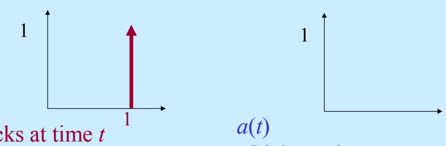


a(*t*)
P[ticks at time *t*]

A(*t*)
P[ticks at or before time *t*]

Consider A Clock and Time

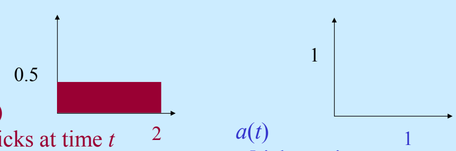
but for a clock that ticks regularly, a(t) *changes* with time (I.e., depending on how long ago the clock ticked)



a(*t*)
P[ticks at time *t*
| *clock just ticked*]

a(*t*)
P[ticks at time *t*
| *clock ticked ½s ago*]

A Probabilistic Clock and Time

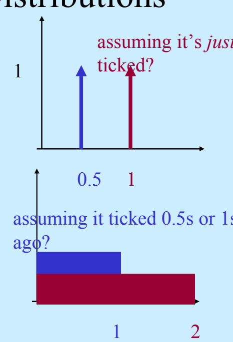


a(*t*)
P[ticks at time *t*
| *clock just ticked*]

a(*t*)
P[ticks at time *t*
| *clock ticked 1s ago*]

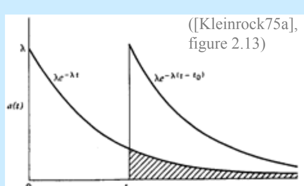
Memoryless Distributions

- in both of these examples, the distribution *changes* based on prior state
  - it has *memory*
- Is there an *a*(*t*) that doesn't change as a function of when the clock last ticked?
  - i.e., is *memoryless*



Poisson Processes

- Poisson traffic
  - *a*(*t*) = $\lambda e^{-\lambda t}$ probability next arrival is at *t*
  - λ is arrival rate (steady over time)
  - is *memoryless*
    - a(t) is the same, even after t_0

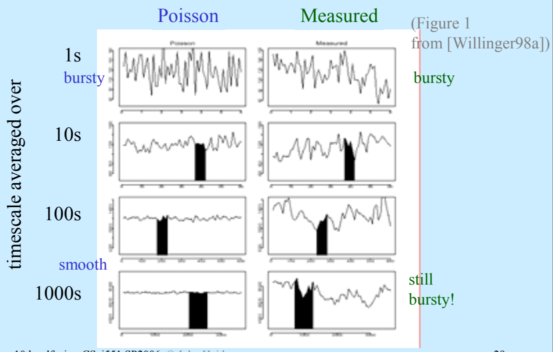


([Kleinrock75a], figure 2.13)

Advantages of Poission Modeling

- analytically tractable
  - can solve hard but relevant problems
- applicable to telephone traffic
  - (may require somewhat more complexity, but basic details are here)
- smoothes out when you combine many independent users
  - allows easier planning
- applicable to many computer problems
  - where did we see this before? xxx

Traffic at Different Time-scales



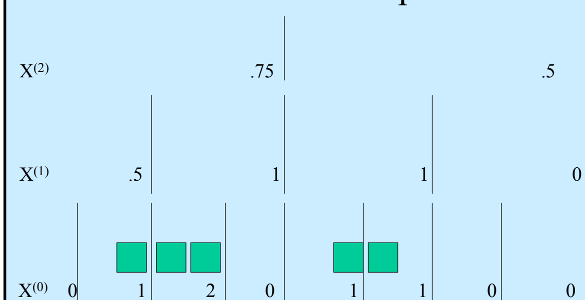
Measured Data

- shows burstiness at many different scales
  - No natural burst length
  - … unlike Poisson
- but what does this mean?
  - hard to model with Poisson
    - but will need many parameters
    - not just arrival rate λ
    - models with just simple λ may not match real net traffic
  - would prefer better model

Defining Time-scale

- Given a random variable X
  - say, pkts/s or bytes/s (or pkts/1ms)
  - let X_t be measurement at time *t,* spaced δ apart
- Define *time-scale m* recursively:
  - $X_t^{(m+1)} := X_t^{(m)} X_{t+\delta}^{(m)}$
  - basically, add things up into larger intervals
- Poisson gets *smoother* at larger *m*, but Internet traffic stays *bursty*
- Question: how can we model bursty traffic?

Time Scale Examples

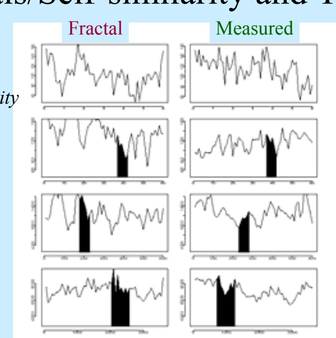


A Hint: Fractals

- Def: A class of objects with surprising scaling properties
- Example: Length of coastline depends on level of detail
- No "natural" length for these objects

Fractals/Self-similarity and Traffic

self-similarity
i.e., looks like itself (bursty) at different time-scales



fractal tfc looks like measured tfc

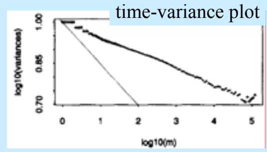
but… how can we *characterize* burstiness

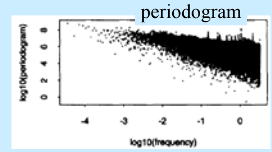
(and *why* does it happen?)

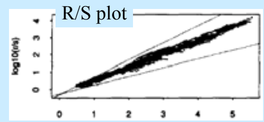
Characteristics of Self-similarity

- variance decays slowly as *m* increases
  - self-sim: $Var(X(m)) \sim a\, m^{-\beta}$, $0 < \beta < 1$
  - Poisson: $Var(X(m)) \sim a\, m^{-1}$
  - i.e., self-sim *remains bursty* (high variance)
- autocorrelation [r(*k*)] decays hyperbolically rather than exponentially
  - self-sim: $r^{(m)}(k) \sim k^{-\beta}$, where r(*k*) is autocorrelation spaced by *k*
  - Poisson: $r^{(m)}(k) \sim 0$ as $m \rightarrow \infty$
  - i.e., self-sim has *long-range dependence*
- spectral density is power-law near origin

Observing Self-Similarity





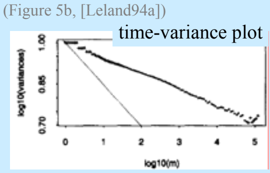


(Figure 5, [Leland94a])

general approach:
- evaluate data over multiple timescales
- measure trend, infer *H*, the *Hurst Parameter*
- somewhat subjective, so validate in multiple ways

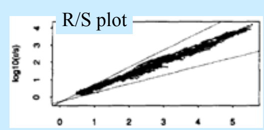
Time-Variance Plot

(Figure 5b, [Leland94a])



- variance decays slowly as *m* increases
  - self-sim: $Var(X(m)) \sim a\, m^{-\beta}$, $0 < \beta < 1$
  - i.e., self-sim *remains bursty* (high variance)
- time-variance plot captures decay of variance
  - time (in *m*, scale) vs. variance

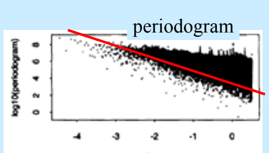
R/S Plot



(Figure 5a, [Leland94a])

- autocorrelation [r(*k*)] decays hyperbolically rather than exponentially
  - self-sim: $r^{(m)}(k) \sim k^{-\beta}$, where r(*k*) is autocorrelation spaced by *k*
  - i.e., self-sim has *long-range dependence*
- R/S Plot captures autocorrelation via *H* (the *Hurst parameter*)
  - H is the slope on the R/S plot
  - $H = 1 - \beta/2$

Spectral Density



- periodogram / Whittle estimator measures spectral density
- can get statistical bounds (confidence intervals) on H
- largely superceeded by wavelet analysis today
  - will see examples next time

(Figure 5, [Leland94a])

got to here 23-Mar-06