

Wide-Area Traffic: The Failure of Poisson Modeling

Vern Paxson and Sally Floyd

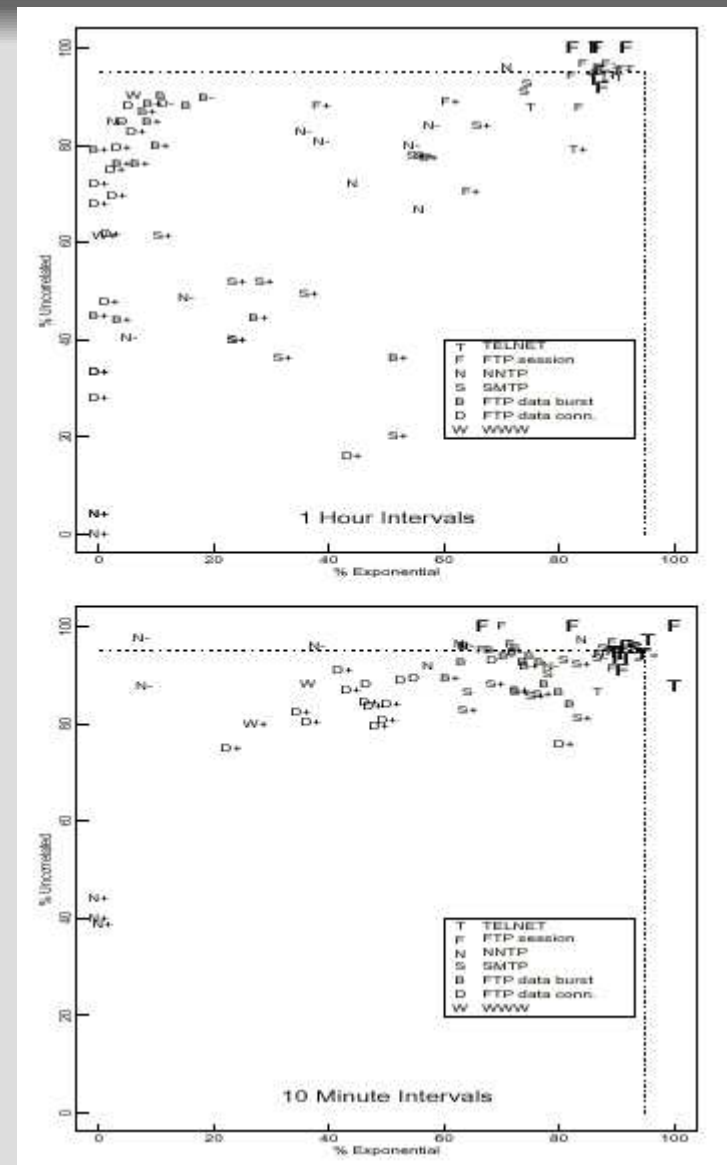
Presented by: Scott Weber
CSE 498
Lehigh University
27.February.2003

Introduction

- Poisson modeling is sometimes used to represent Internet traffic in testing
 - Not very accurate in most cases
- Self-similar modeling may be a better representation
- Analyze packet traces and TCP connection traces from variety of locations

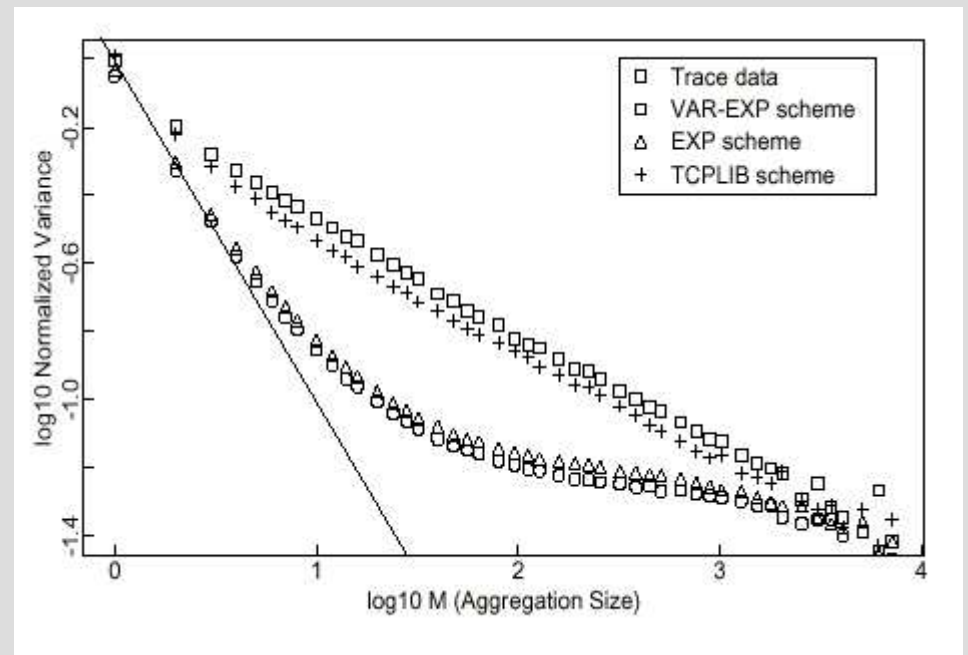
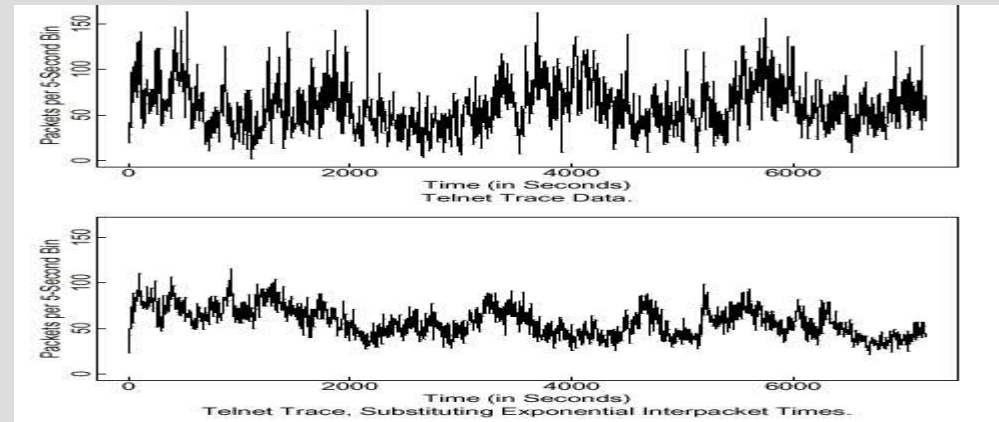
TCP Connections

- Simple testing can show whether Poisson might fit
 - Telnet (T) and FTP-Session (F) seem pretty good fit, especially on smaller time scale



Telnet Sessions

- Poisson / exponential model not useful for packet arrival during a telnet session
- Empirically derived TCPLIB model is a much better estimator



Complete Telnet Sessions

- A complete model of telnet traffic is developed
 - Poisson connection arrivals
 - Log-normal connection sizes
 - TCPLIB packet interarrivals
- Based on empirical data, seems to be a good model

FTP Data Connections

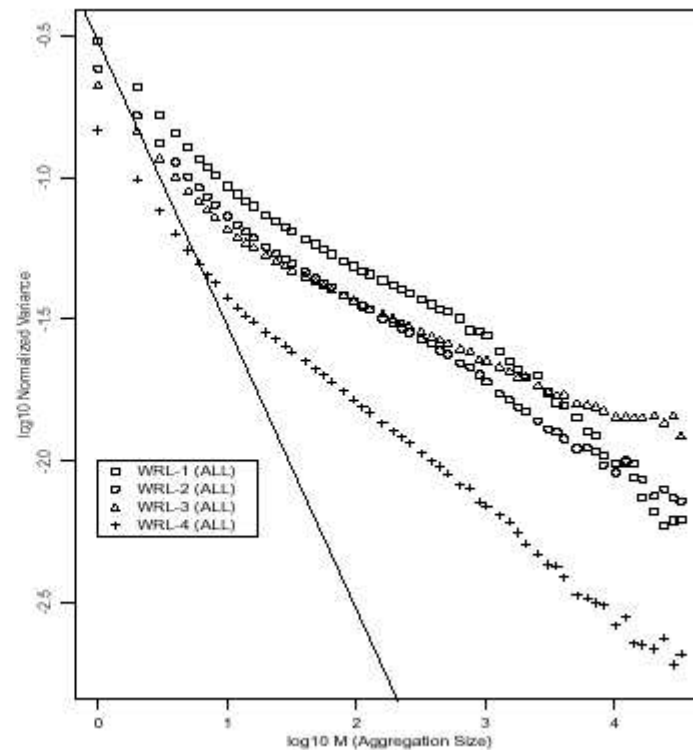
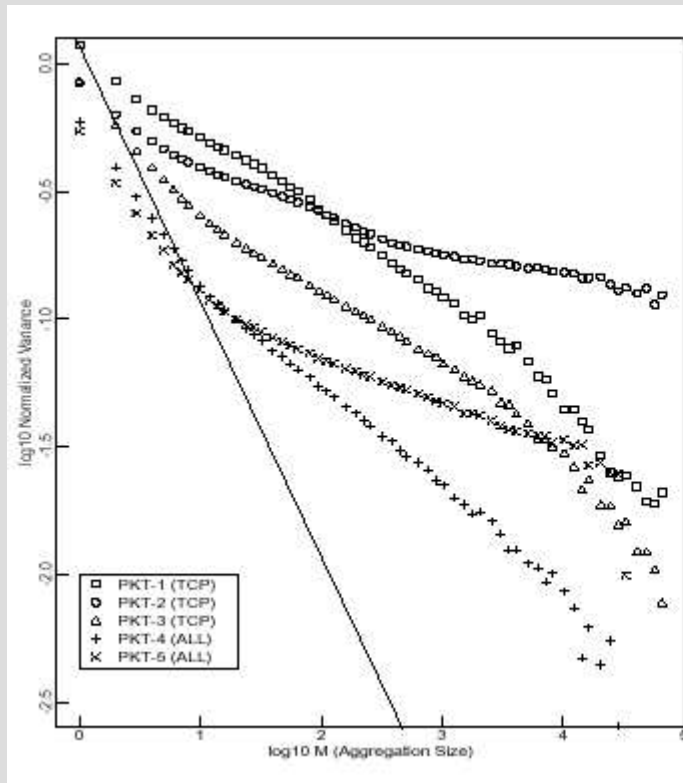
- Connections tend to come in bursts
 - "mget"
 - "ls" followed by "get"
- Small number of connections account for most of data transferred
 - Corresponds to general distribution of file sizes
- Modeling the small number of connections is most important
 - Pareto distribution fits the heavy tail

Self-similarity

- Possible causes
 - Multiplexing of sporadic sources
 - Queueing model
- Easy to pick out visually
 - straight line with $(0 > \text{slope} > -1)$ on a variance-time plot
- Hopefully usable to model general traffic

Preliminary Results

- General traffic does not fit self-similarity well
 - probably due to large-traffic bursts



Conclusions

- Understand the validity of your model before you use it