Parameterisation of multi-map from Internet traffic traces

Girum Aklilu, Jonathan Pitts

Queen Mary, University of London

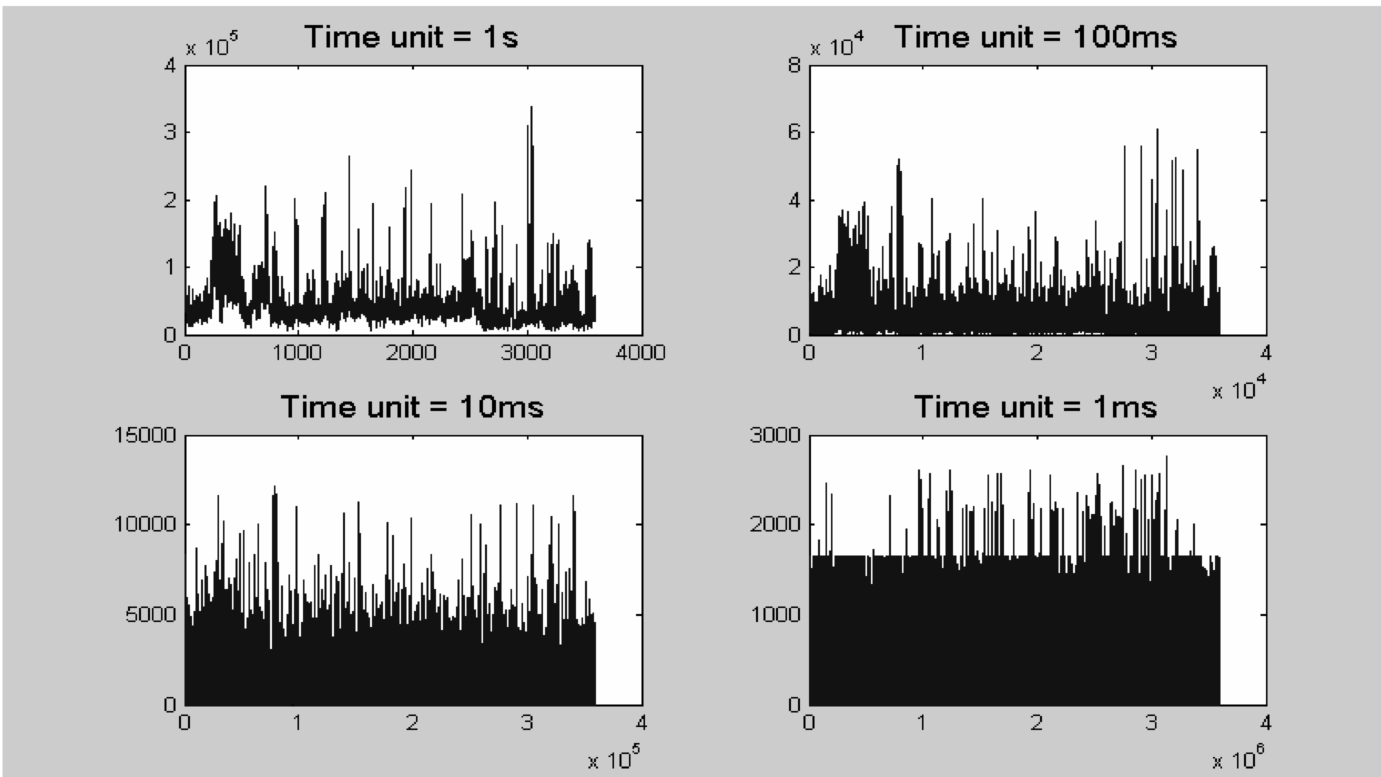
Queen Mary
University of London

Presentation Overview

- Observation of Internet traffic analysis
  - Observed scalings
- Multi-map as a traffic model
  - Structures and equations
  - Parameters of the map
- Trace based parameterisation of the multi-map
  - step by step analysis
- Comparison of real versus synthetic traffic analysis
  - R/S scaling comparison, load and variance analysis
- Future directions
  - More trace analysis
  - Model use in network analysis

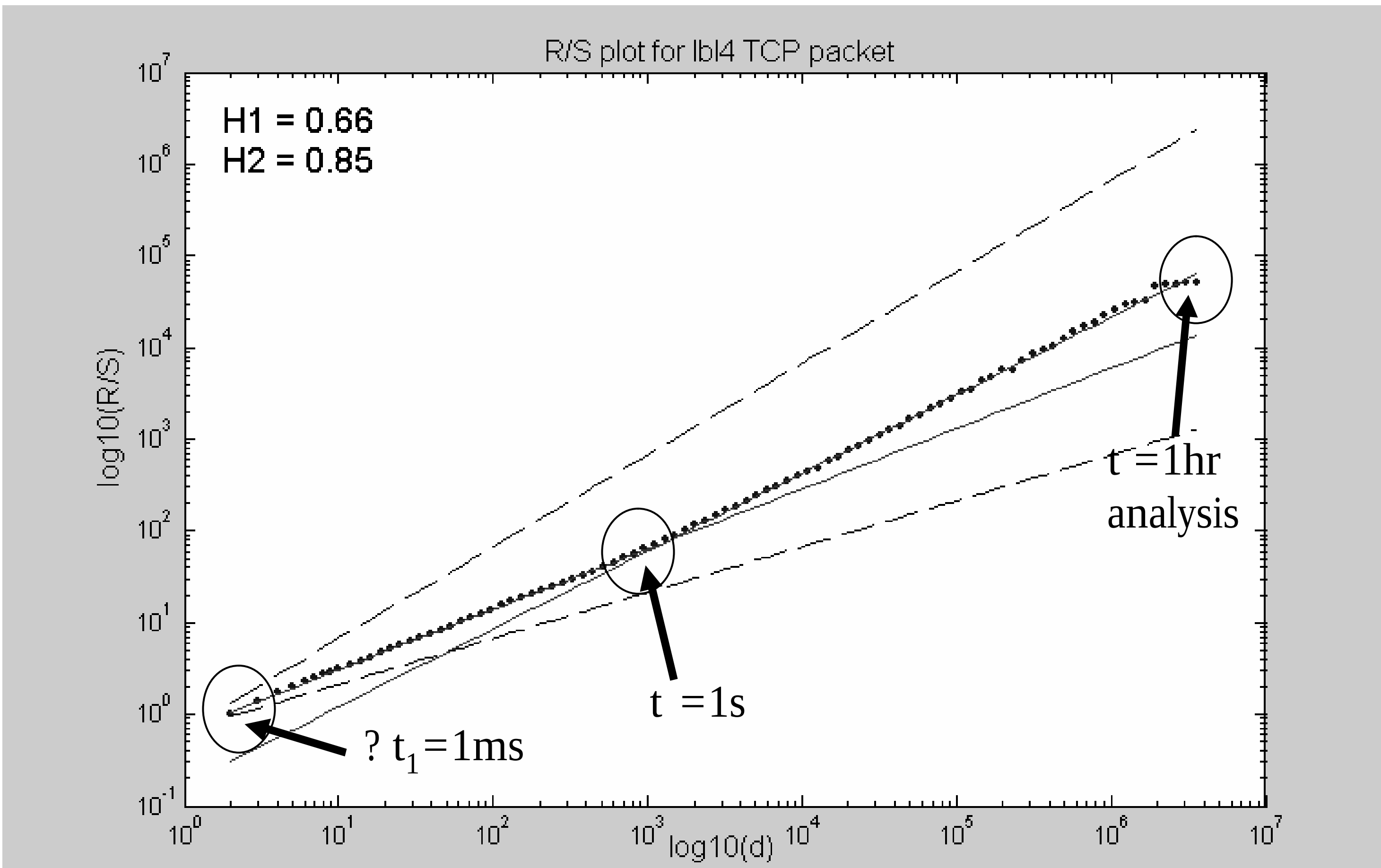
Traffic trace analysis

- TCP traffic traces from WAN at Berkeley Labs
  - An hourly trace (lbl-4)
  - 1.3 million packets in the trace

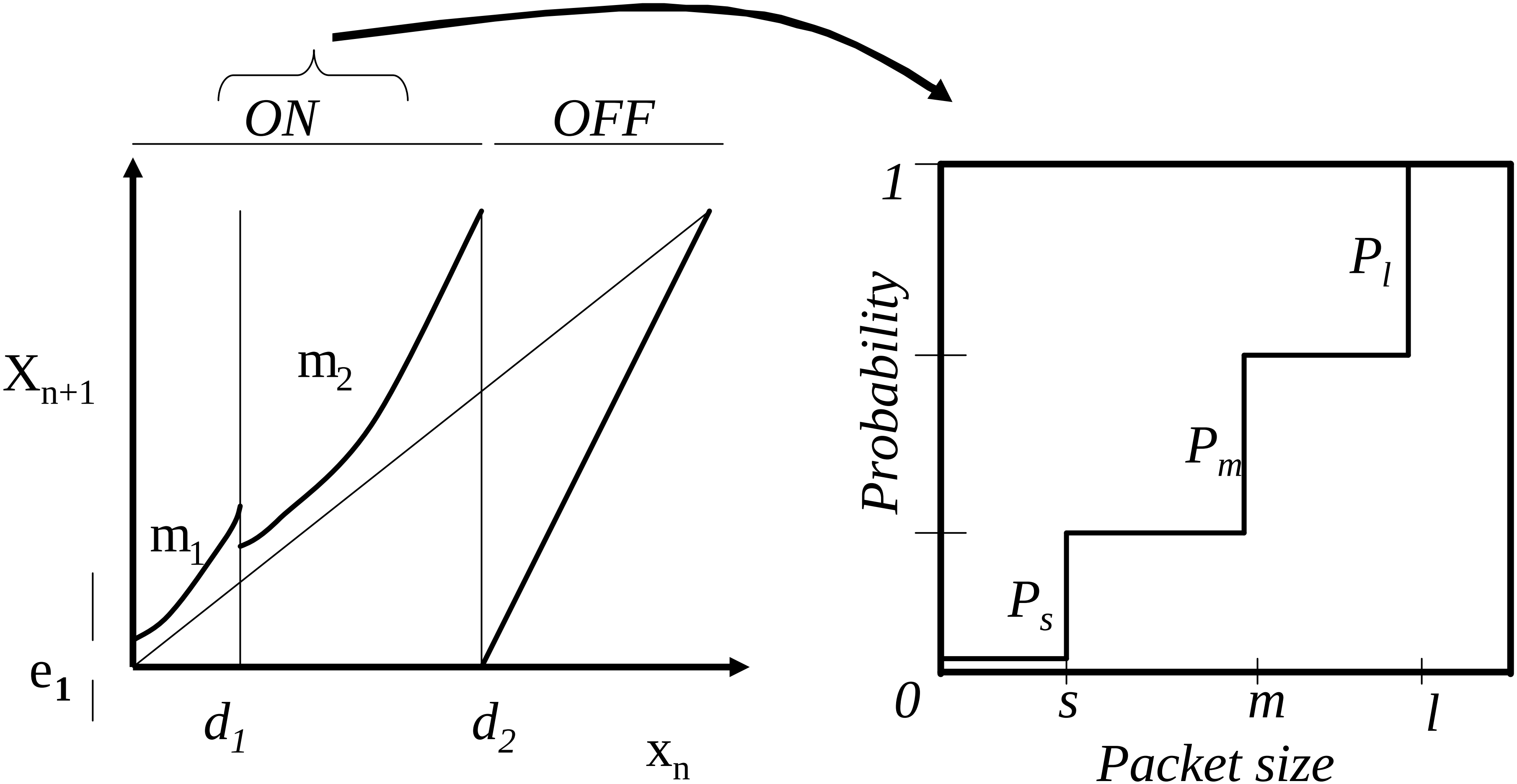


Scaling of traffic traces

- R/S analysis showed two scaling ranges
  - Two Hurst parameters with cross over point at around 1s



Structure of multi-map model



Equations of multi-map

- Hidden dynamical layer

$$x_{n+1} = F(x_n) = \begin{cases} F_1(x_n) & = e_1 + x_n + \dfrac{1-e_1-d_2}{d_2^{\,m_1}} x_n^{\,m_1} & 0 < x_n < d_1 \\ F_2(x_n) & = e_2 + x_n + \dfrac{1-e_2-d_2}{d_2^{\,m_2}} x_n^{\,m_2} & d_1 \le x_n < d_2 \\ F_3(x_n) & = x_n - d_2 \dfrac{1-x_n}{1-d_2} & d_2 \le x_n < 1 \end{cases}$$

  - where parameters $x_n \in (0,1) \quad m_i \in (1,2) \quad d_i \in (0,1)$

- Visible dynamical layer

$$y(x_n) = \begin{cases} 1 & 0 < x_n < d_2, \\ 0 & d_2 \le x_n < 1, \end{cases}$$

ON (packet)

OFF (no packet)

Parameterisation of the multi-map

- Packet size distribution analysis
  - Packet sizes - *s*, *m*, *l*, & corresponding - P_s, P_m, P_l

- Time scale and load analysis
  - d_2, ? t_2

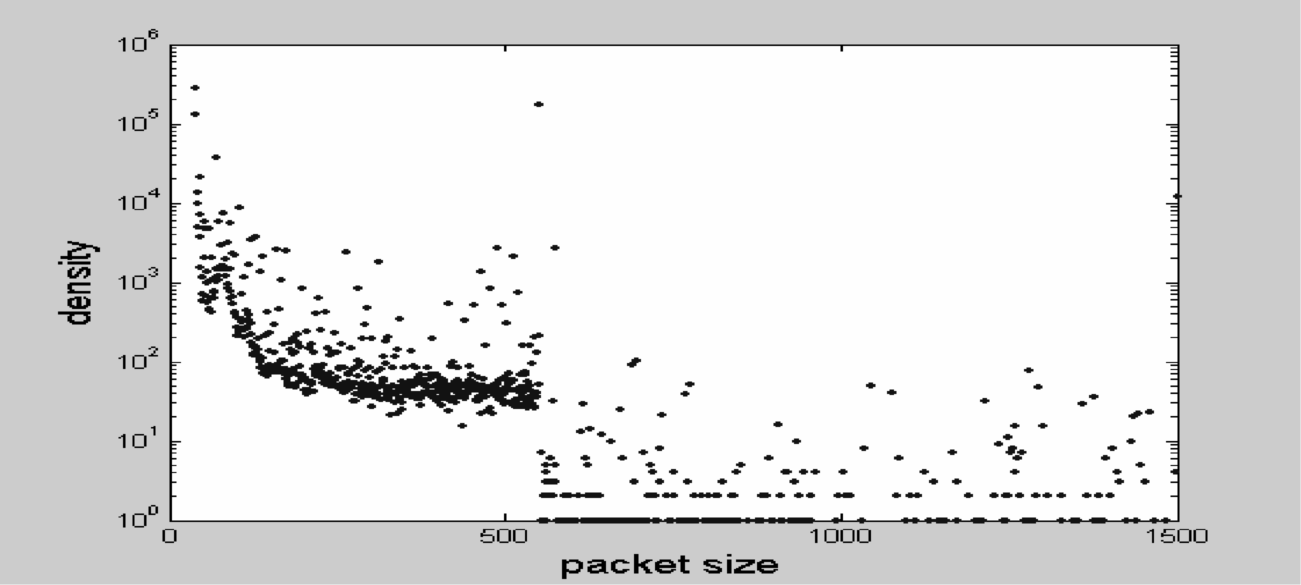
- Multi scaling analysis
  - m_1, m_2, *e*, d_1

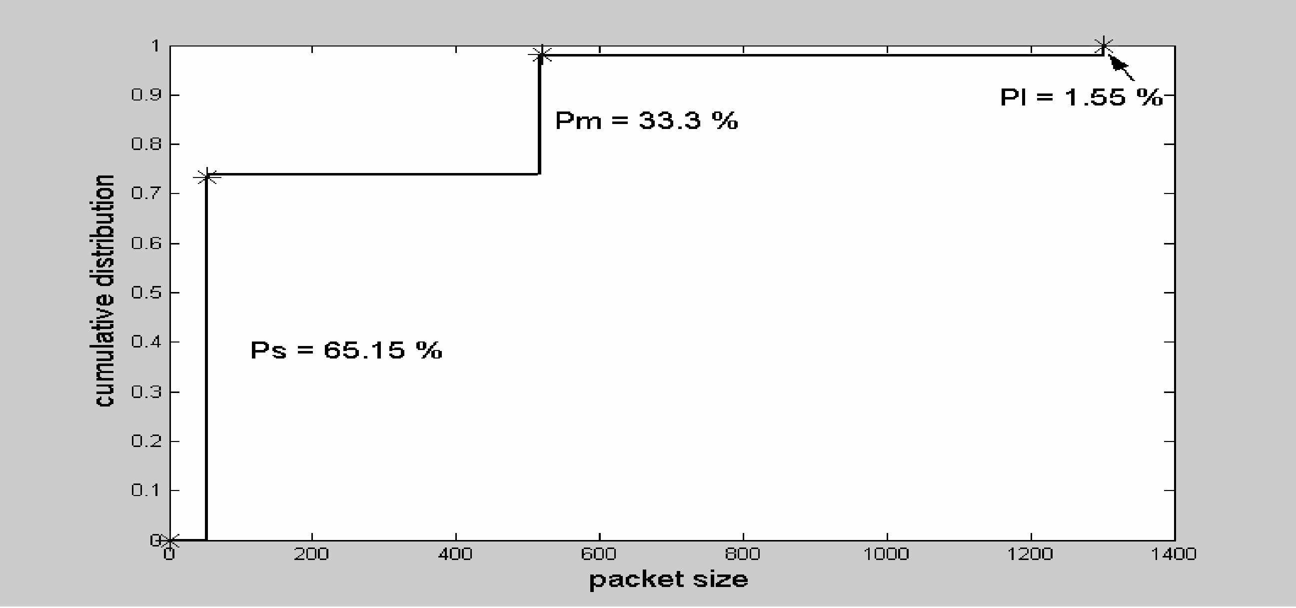
Distribution analysis

- Packet size analysis based on Imix Internet packet mixture [*Journal of Internet test*]
  - Result of 342 million packet analysis at NLANR
  - Accurate correlation when compared to realistic Internet traffic
  - Tri-modal distribution analysis from traces
  - Output - three mean packet sizes and their distribution
    - Small, medium and large sizes and respective probabilities

"Journal of Internet test Methodologies", Edition 2, Agilent technologies

Distribution analysis (contd.)



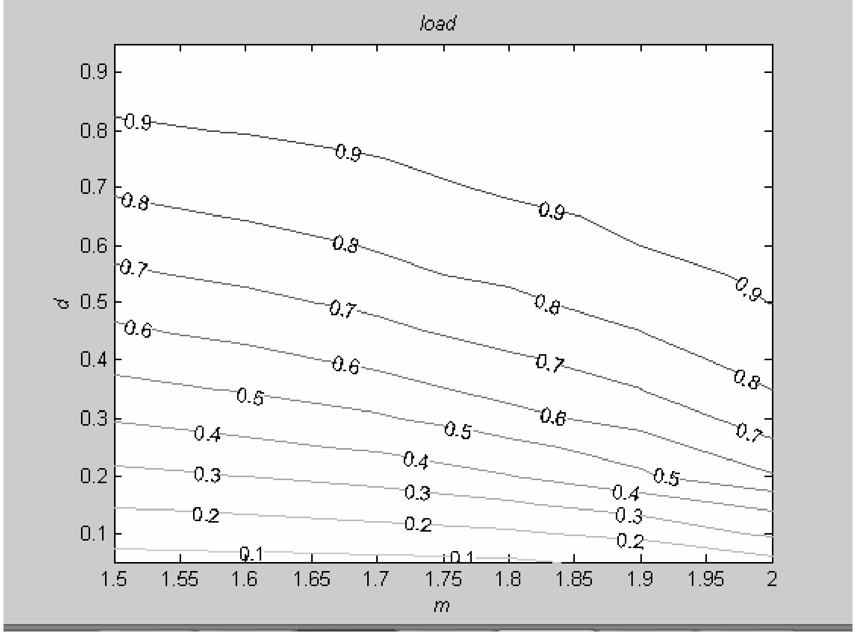


Time scale and load analysis

- Variance and load calculated for each trace with regards to ? t_1
  - smallest time scale unit the trace analysed at
- Respective analysis of iteration time
  - ? $t_2 = P_l / C$
    - where P_l is max packet size and C is link rate
  - Hence one iteration corresponds to ? t_2
- Scaling comparison achieved with equal time units

Time scale and load analysis – parameter d_2

- Proportion of time increments that have a packet and hence the load, relates to parameter d_2
- Equivalence to single intermittency map in its structure
  - Based on look up table of parameters *d*, *m* and *load*



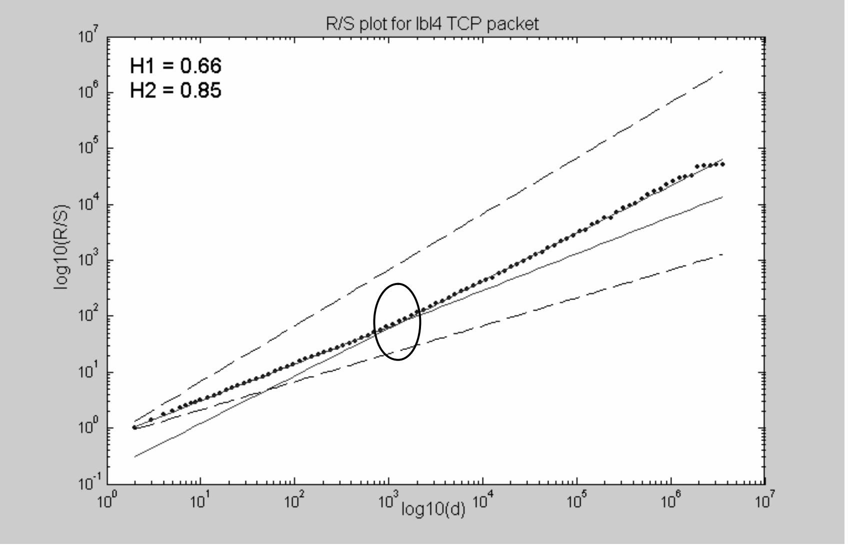
  - Output – parameter d_2 for specific load and m_2

Time scale and load analysis – constraints on parameter d_2

- Constraints on d_2 [Samuel] leads to changes in parameters
  - If d_2 lies outside regions 0.1 to 0.9
    - *H* value is lower than expected
    - Hence keep d_2 within the region
    - Results in changes in proportion of time increments that have a packet
    - Leads to change in ? t_2 – iteration time unit

Multi scaling analysis – *H* test

- Hurst parameter using R/S analysis
  - Advantage of showing distinctive scaling regions



  - Output – two H values with cross over point
    - $H_1, H_2, t = n_{t*} ? t_1$

Multi scaling analysis – Parameters m_1 and m_2

- The m parameters are directly linked to the Hurst parameters [Mondragon]
- Equation of m from H parameters

$$H = \left(\frac{3m - 4}{2m - 2} \right)$$

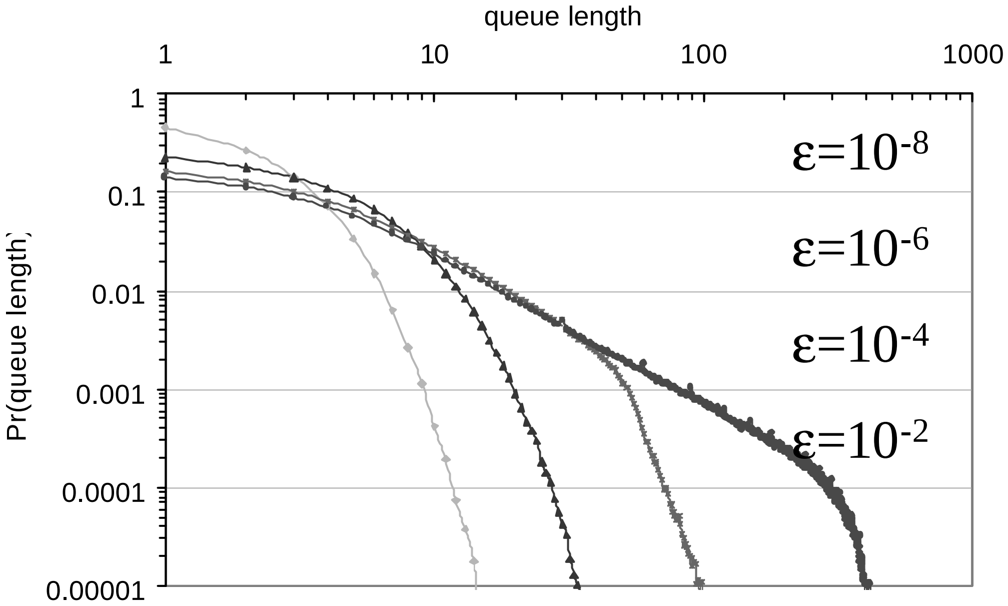
$$\Rightarrow \qquad m = \left(\frac{4 - 2H}{3 - 2H} \right)$$

- Output m_1 and m_2

"A model of packet traffic using random wall model", Mondragon, 14th ITC

Multi scaling analysis – Parameter *e*

- *e* parameter allows effective control of LRD
  - cut-off point in case of single intermittency map
  - Queue analysis result



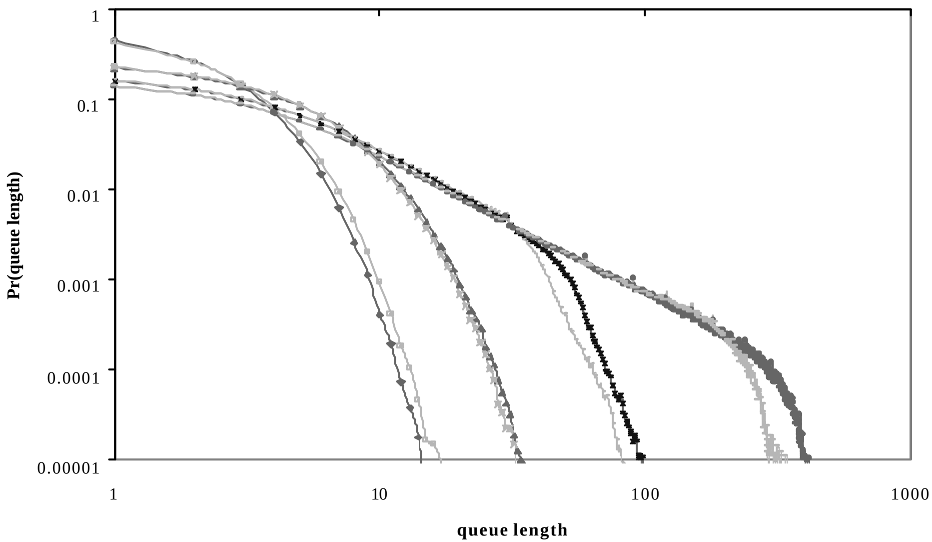
- Output *e* versus cross-over point

Multi scaling analysis – Parameter d_1

- Parameter d_1

$$d_1 = d_2 \left(\frac{e}{1 - d_2} \right)^{\frac{1}{m_2}}$$

  - Controls the cross over point
  - Equivalence achieved with single intermittency map



Parameter value results –lbl4

- Packet size distribution analysis
  - Packet sizes – *s = 45 bytes*, *m= 520 bytes*, *l = 1451 bytes*,
  - $P_s = 0.6515$, $P_m = 0.333$, $P_l = 0.0155$

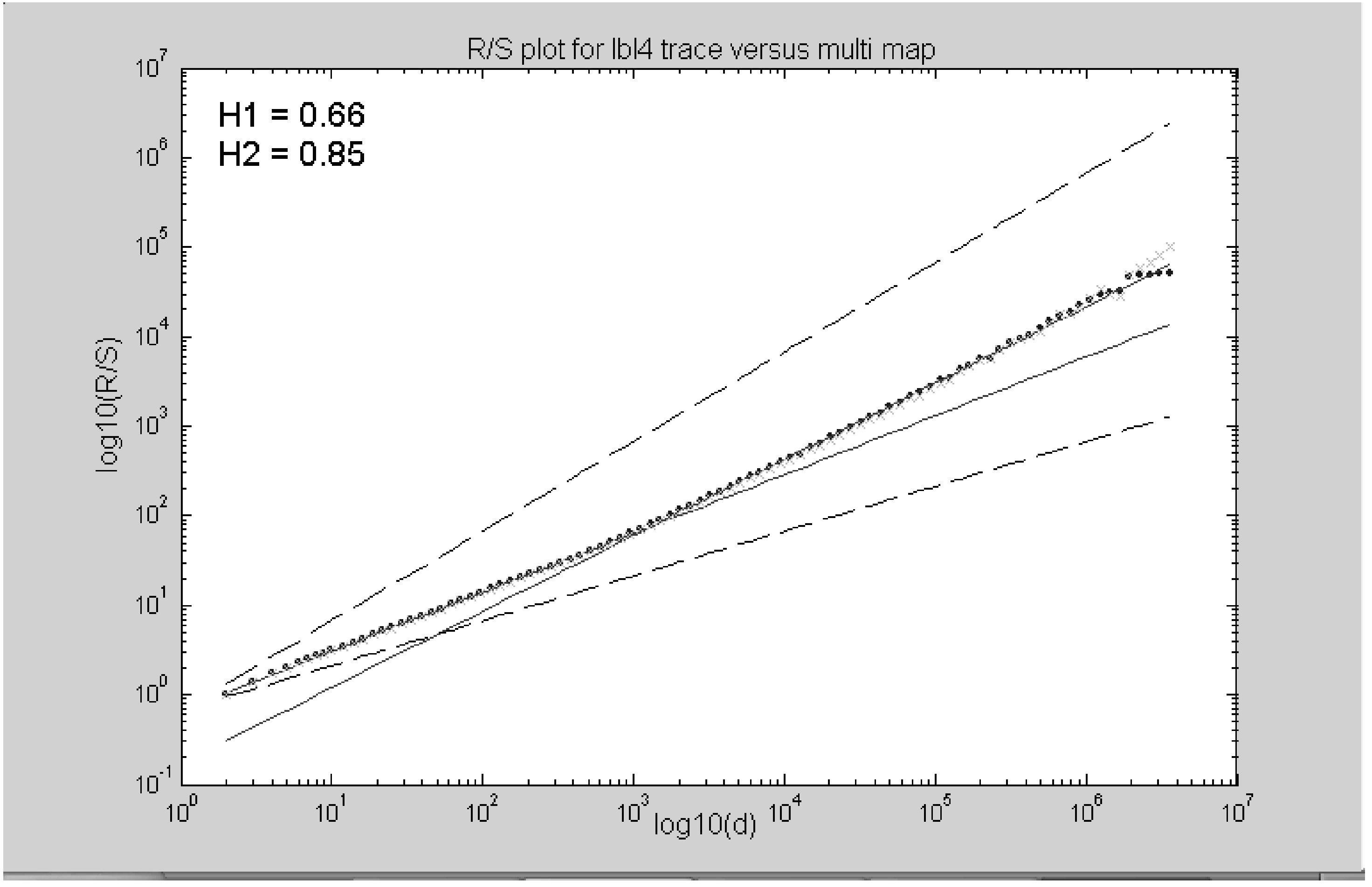
- Time scale and load analysis
  - $d2 = 0.1$, $? t_2 = 500\mu s$

- Multi scaling analysis
  - $m_1 = 1.8$, $m_2 = 1.6$, $e = 1*10^{-6}$, $d_1 = 1.9*10^{-5}$

Comparison of real versus synthetic traffic

- Mean load
  - Trace — Multi-map
    - 330 bit/s — 293 bits/s
    - 0.2397 pps — 0.232 pps
- Variance
  - Trace — Multi-map
    - 167.8 bit/s — 137 bit/s
    - 0.527 pps — 0.5 pps
- H parameter
  - Trace — Multi-map
    - 0.66, 0.85 — 0.65, 0.82

Comparison of real versus synthetic traffic from multi map



Future Directions

- More trace analysis
  - York trace
- Use of traces in network analysis
  - Advantages of multi-map
    - Parsimonious, effective parameterisation, fast
  - Multi service complex traffic scenarios
    - Traffic generation for different classes of traffic
    - Performance analysis