A Renaissance in Network Measurement

- Not just monitoring
  - Traffic patterns: link, path, node, applications, management
  - Quality of Service: delay, loss, reliability
  - Protocol dynamics: TCP, VoIP, ..
  - Network infrastructure: routing, security, DNS, bottlenecks, latency..

- Knowing, understanding, improving network performance

- Data centric view of networking
  - Must arise from real problems, or observations on data
  - Abstractions based on data
  - Results validated by data
  - In fact: the scientific method in networking
  - Not just getting numbers, *Discovery*
THE RISE OF MEASUREMENT

- Papers between 1966-87 (P.F. Pawlita, ITC-12, Italy)
  - Queueing theory: several thousand
  - Traffic measurement: around 50

- Now have dedicated conferences:
  - Passive and Active Measurement Conference (PAM) 2000 …
  - ACM Internet Measurement Conference (IMC) 2001 …

IMC 2010 : MELBOURNE!

- Conference: ~ Nov 1-3
- Deadline: ~ May 1
- Venue: BMW Edge at Federation Square
IMC 2010 : MELBOURNE!

- Conference: ~ Nov 1-3
- Deadline: ~ May 1
- Venue: BMW Edge at Federation Square

ACTIVE VERSUS PASSIVE MEASUREMENT

- Typical Passive Aims
  - “At-a-point” or “Network Core”
  - Link utilisation, Link traffic patterns, Server workloads
  - Long term monitoring:
    - Dimensioning, Capacity Planning, Source modelling
  - Engineering view: Network performance

- Typical Active Aims
  - “End-to-End” or “Network edge”
  - End-to-End Loss, Delay, Connectivity, “Discovery” ….
  - Long/short term monitoring: Network health; Route state
  - Internet view: Application performance and robustness
TOMO - GRAPHY

Tomos section

Graphia writing
EXAMPLES OF TOMOGRAPHY

- Atom probe tomography (APT)
- Computed tomography (CT) (formerly CAT)
- Cryo-electron tomography (Cryo-ET)
- Electrical impedance tomography (EIT)
- Magnetic resonance tomography (MRT)
- Optical coherence tomography (OCT)
- Positron emission tomography (PET)
- Quantum tomography
- Single photon emission computed tomography (SPECT)
- Seismic tomography
- X-ray tomography

COMPUTED TOMOGRAPHY
**Network Tomography**

Began with Vardi [1996]

"Network Tomography: estimating source-destination traffic intensities from link data"

**Classes of Inversion Problems**

- End-to-end measurements $\rightarrow$ internal metrics
- Internal measurements $\rightarrow$ path metrics

**The Metrics**

- Link Traffic (volume, variance) (Traffic Matrix estimation)
- Link Loss (average, temporal)
- Link Delay (variance, distribution)
- Link Topology
- Path Properties (network "kriging")
- Joint problems (use loss or delay to infer topology)

---

**The Early Literature (Incomplete!)**

- Traffic Matrix Tomography
  - AT&T (Zhang, Roughan, Donoho et al.)
  - Sprint ATL (Nucci, Taft et al.)

- Loss/Delay/Topology Tomography
  - AT&T (Duffield, Horowitz, Lo Presti, Towsley et al.)
  - Rice (Coates, Nowak et al.)

- **Evolution:**
  - loss, delay $\rightarrow$ topology
  - Exact MLE $\rightarrow$ EM MLE $\rightarrow$ Heuristics
  - Multicast $\rightarrow$ Unicast (striping)
**Active Probing versus Network Tomography**

- **Active Probing**
  - Typically over a *single path*
  - Use tandem FIFO queue model
  - Exploit discrete packet effects in *semi-heuristic queueing analysis*
  - Typical metrics: link capacities, available bandwidth

- **Network Tomography**
  - Typically “network wide”: multiple destinations and/or sources
  - Simple black box node/link models, *strong assumptions*
  - Classical inference with twists
  - Typical metrics: per link/path loss/delay/throughput

---

**Broader Vision of Network Inference**

- **Active Probing**
  - solving formal problems in inverse queueing
  - seeking optimal probing methods and methodologies

- **Network Tomography**
  - black box models -> queueing compatible models
  - trees -> general network topologies

- **Pathways to Impact**
  - unimplementable research analysis -> tools for consumer watchdogs
  - specialist software -> smart phone apps

- **Networks/Measurement Disconnect**
  - measurement friendly networks (building tomography into the NBN)
  - underlying timing infrastructure (RADclock project)
TWO PROBLEMS IN LOSS TOMOGRAPHY OVER TREES

• Removing the temporal independence assumption
  Arya, Duffield, Veitch, 2007

• Exploiting Sparsity
  Arya, Veitch, 2009 (ongoing)

LOSS TOMOGRAPHY USING MULTICAST PROBING
**The Loss Model**

Stochastic loss process on link $k$ acts deterministically on probes arriving to $f(k)$

Node and Link Processes

- $\{X_k(i) : i \in Z\}$: loss process on link $k$
- $\{Z_k(i) : i \in Z\}$: probe 'bookkeeping' process for node $k$

\[
\begin{array}{c}
\text{f(k)} \quad X_{f(k)} \quad 1 \quad 0 \quad 0 \quad 1 \quad \ldots \\
\text{Z_k} \quad 1 \quad 1 \quad 0 \quad 0 \quad \ldots \\
\text{X_k} \quad 1 \quad 0 \quad 0 \quad 0 \quad \ldots \\
\end{array}
\]

- Link loss process
- Bookkeeping process

\[X_k(i) = Z_k(i)X_{f(k)}(i)\]

\[
\begin{align*}
\Pr[X_k(i) = b | X_{f(k)}(i) = 1] &= \Pr[Z_k(i) = b], & b \in \{0, 1\} \\
\Pr[X_k(i) = 0 | X_{f(k)}(i) = 0] &= 1
\end{align*}
\]

**Adding Probability: Loss Dependencies**

Independent Probes

- Spatial:
  - loss processes on different links independent
- Temporal:
  - losses within each link independent

Model reduces to a single parameter per link, the

\[\text{passage or transmission probabilities } \{\alpha_k\}\]
FROM LINK PASSAGE TO PATH PASSAGE PROBABILITIES

Path probabilities: only ancestors matter

\[ X_k(i) = Z_k(i)X_f(k)(i) = \prod_{j \in \alpha(k)} Z_j(i) \]

Let \[ A_k = \Pr[X_k(i) = 1] \]

\[ A_k = \alpha_kA_f(k) = \prod_{j \in \alpha(k)} \alpha_j \]

Sufficient to estimate path probabilities

ACCESSING INTERNAL PATHS

Aim: estimate \( A_k(i) \)

\[ Y_k(i) = X_b(i) \lor X_c(i) \]

Let \[ \gamma_k(i) = \Pr[Y_k(i) = 1] \]

\[ \beta_b(i) = \Pr[X_b(i) = 1 | X_k(i) = 1] = \frac{A_b(i)}{A_k(i)} \]

\[ \gamma_k(i) = A_k(i)\{1 - (1 - \beta_a(i))(1 - \beta_b(i))\} \]

Obtain a quadratic in \( A_k(i) \)

Original MINC loss estimator for binary tree

[Cáceres, Duffield, Horowitz, Towsley 1999]
**Temporal Independence: How Far to Relax?**

**Before**
- Spatial: link loss processes independent
- Temporal: link loss processes Bernoulli
- Parameters: link passage probabilities $\{\alpha_k\}$

**After**
- Spatial: link loss processes independent
- **Temporal:** link loss processes stationary, ergodic
- **Parameters:** joint link passage probabilities $\{\alpha_k(I)\}$ over index sets $I = \{i_1, i_2, \ldots, i_s\}$
  - Full characterisation/identification possible!
TARGET LOSS CHARACTERISTIC

- Loss run-length distribution (density, mean)

\[ \text{Loss-run length} \]

• Importance
  • Impacts delay sensitive applications like VoIP (FEC tuning)
  • Characterizes bottleneck links

ACCESSING TEMPORAL PARAMETERS: GENERAL PROPERTIES

- Sufficiency of joint link passage probabilities
  
  \[ \Pr[Z_k(i) = 0, Z_k(i + 1) = 1] = \Pr[Z_k(i) = 1] - \Pr[Z_k(i) = 1, Z_k(i + 1) = 1] \]

• Mean Loss-run length

\[ E[L_k] = \frac{\Pr[Z_k(i) = 0]}{\Pr[Z_k(i) = 0, Z_k(i + 1) = 1]} = \frac{1 - \Pr[Z_k(i) = 1]}{\Pr[Z_k(i) = 1] - \Pr[Z_k(i) = 1, Z_k(i + 1) = 1]} \]

• Loss-run distribution

\[ \Pr[L_k \geq j] = \frac{\Pr[Z_k(i) = 1, \ldots, Z_k(i + j - 1) = 1] - \Pr[Z_k(i) = 1, \ldots, Z_k(i + j) = 1]}{\Pr[Z_k(i) = 1] - \Pr[Z_k(i) = 1, Z_k(i + 1) = 1]} \]
**Joint Passage Probabilities**

- Joint **link** passage probability
  \[ \alpha_k(I) = \text{Pr}[Z_k(I) = 1] \]
  e.g. \( \alpha_k(\{1, 2\}) = \alpha_k(\{i, i+1\}) \)

- Joint **path** passage probability
  \[ A_k(I) = \text{Pr}[X_k(I) = 1] = A_{f(k)}(I)\alpha_k(I) \]

**Estimation: Joint Path Passage Probability**

\[ Y_k(I) = \bigvee_{\gamma_r \in R_k} X_r(I) \]
\[ \gamma_k(I) = \text{Pr}[Y_k(I) = 1] \]
\[ \beta_k^1(I, B) = \text{Pr}[\gamma_k^1(I) = B | X(k)(I) = 1] \]
\[ \gamma_k^1(I) = A_k(I) \beta_k^1(I, 1) \]
\[ \gamma_k^2(I) = A_k(I) \beta_k^2(I, 1) \]
\[ \gamma_k(I) = A_k(I) \left\{ 1 - (1 - \beta_k^1(I, 1)) \cdot (1 - \beta_k^2(I, 1)) \right\} + \sum_{B_1 \neq 1, B_2 \neq 1 \atop B_1 \lor B_2 = 1} \beta_k^1(I, B) \beta_k^2(I, B) \]
\[ \Rightarrow \tilde{A}_k(I) = g(\gamma_k(I), \tilde{\gamma}_k^1(I'), \tilde{\gamma}_k^2(I')) , \quad I' \subseteq 1 \]
Estimation: Joint Path Passage Probability

- Estimation of $A_k(I)$ in general trees
  - Requires solving polynomials with degree equal to the degree of node $k$
  - Numerical computations for trees with large degree
  - Recursion over smaller index sets

- Simpler variants
  - Subtree-partitioning
    - Requires solutions to only linear or quadratic equations
    - No loss of samples
    - Also simplifies existing MINC estimators
  - Avoid recursion over index sets by considering only subsets of receiver events which imply $X_k(I) = 1$

Simulation Experiments

- Setup
  - Loss process
    - Discrete-time Markov chains
    - On-off processes: pass-runs geometric, loss-runs truncated Zipf
  - Estimation
    - Passage probability $\alpha_k(\{1\}) = \alpha_k[1]$
    - Joint passage probability for a pair of consecutive probes $\alpha_k(\{1,2\}) = \alpha_k[11]$
    - Mean loss-run length: $\mu_k = E[L_k] = \frac{1 - \alpha_k[1]}{\alpha_k[1] - \alpha_k[11]}$
  - Relative error: $\left| \hat{\theta} - \theta \right| / \theta$
EXPERIMENTS

• Estimation for shared link in case of two-receiver binary tree

\[ \alpha[1] \quad \alpha[11] \]

Relative Error vs. Mean loss-run length

Markov chain

On-off process

EXPERIMENTS

• Estimation for shared link in case of two-receiver binary tree, continued..

\[ \mu \]

Relative Error vs. Mean loss-run length

\[ \alpha[1] = 5\%, \mu = 3.5 \]

Markov chain

On-off process
**EXPERIMENTS**

- Estimation of $\mu$ for larger trees

Trees taken from router-level map of AT&T network produced by Rocketfuel (2253 links, 731 nodes)

Random shortest path multicast trees with 32 receivers. Degree of internal nodes from 2 to 6, maximum height 6

**VARIANCE**

- Estimation for shared link in case of tertiary tree

GM: general temporal
GS: subtree version
SM: simplified OR
SS: subtree version
SE: subtree & AND
CONCLUSIONS

• Estimators for temporal loss parameters, in addition to loss rates
  • Estimation of any joint probability possible for a pattern of probes

• Class of estimators to reduce computational burden
  • Subtree-partition: simplifies existing MINC estimators

• Future work
  • Asymptotic variance
  • MLE for special cases (Markov chains)
  • Hypothesis tests
  • Experiments with real traffic

TWO PROBLEMS IN LOSS TOMOGRAPHY OVER TREES

• Removing the temporal independence assumption
  Arya, Duffield, Veitch, 2007

• Exploiting Sparsity
  Arya, Veitch, 2009 (ongoing)
CONCLUSIONS

- Network Tomography has many guises and flavours
- Many interesting problems remain!

Thank you