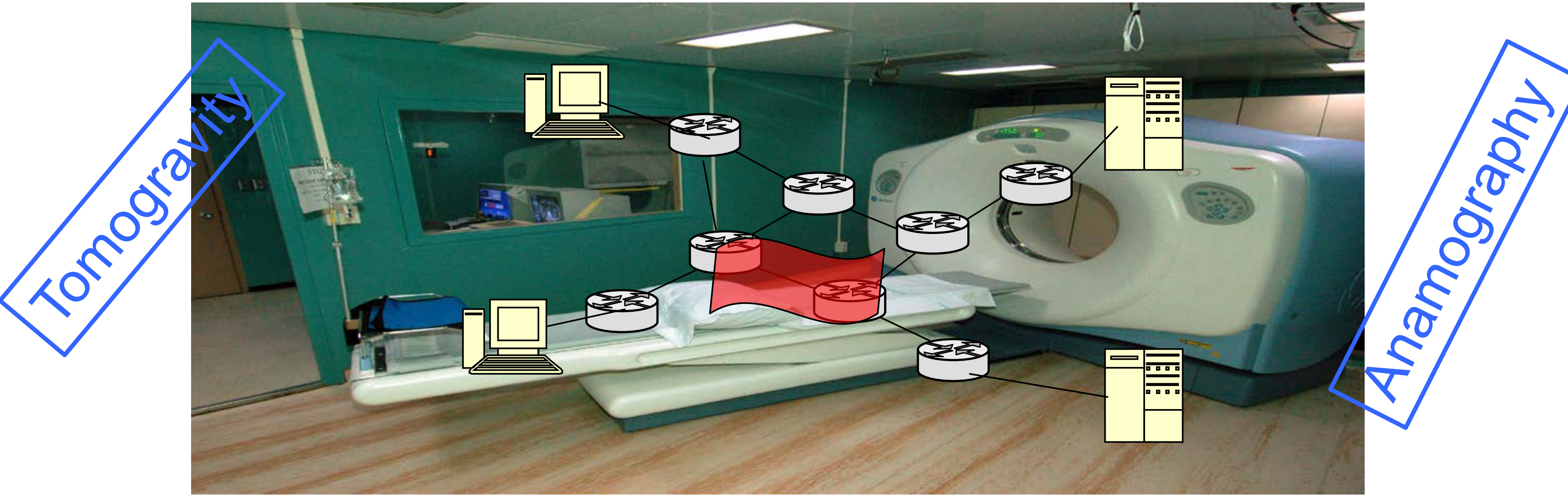
Robust Network Tomography

Srikar Tati
Institute of Network and Security Research
Pennsylvania State University

What is Network Tomography?



- Identify edge nodes and take e2e measurements
- End-to-end measurements: Delay, Log of loss rate, etc.
- Network Monitoring Applications: Diagnose bottlenecks, estimation of topology, estimation of traffic rates, all e2e measurements, etc.

Network Tomography: Introduction

- **Models network as a linear system**

  $Y=AX$

  Y- e2e measurements

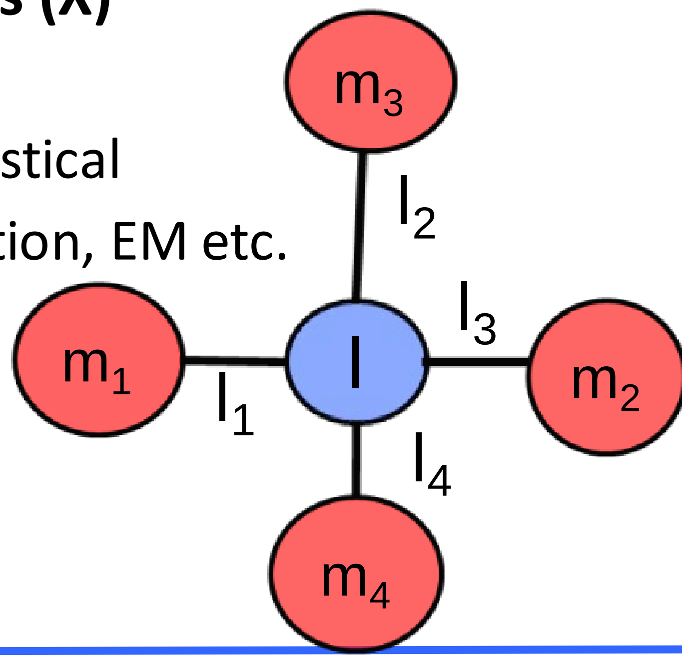
  X- link metrics

  A- path matrix

$$\begin{array}{ccccc} Y & & A & & X \\ \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{pmatrix} & = & \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} & \times & \begin{pmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{pmatrix} \end{array}$$

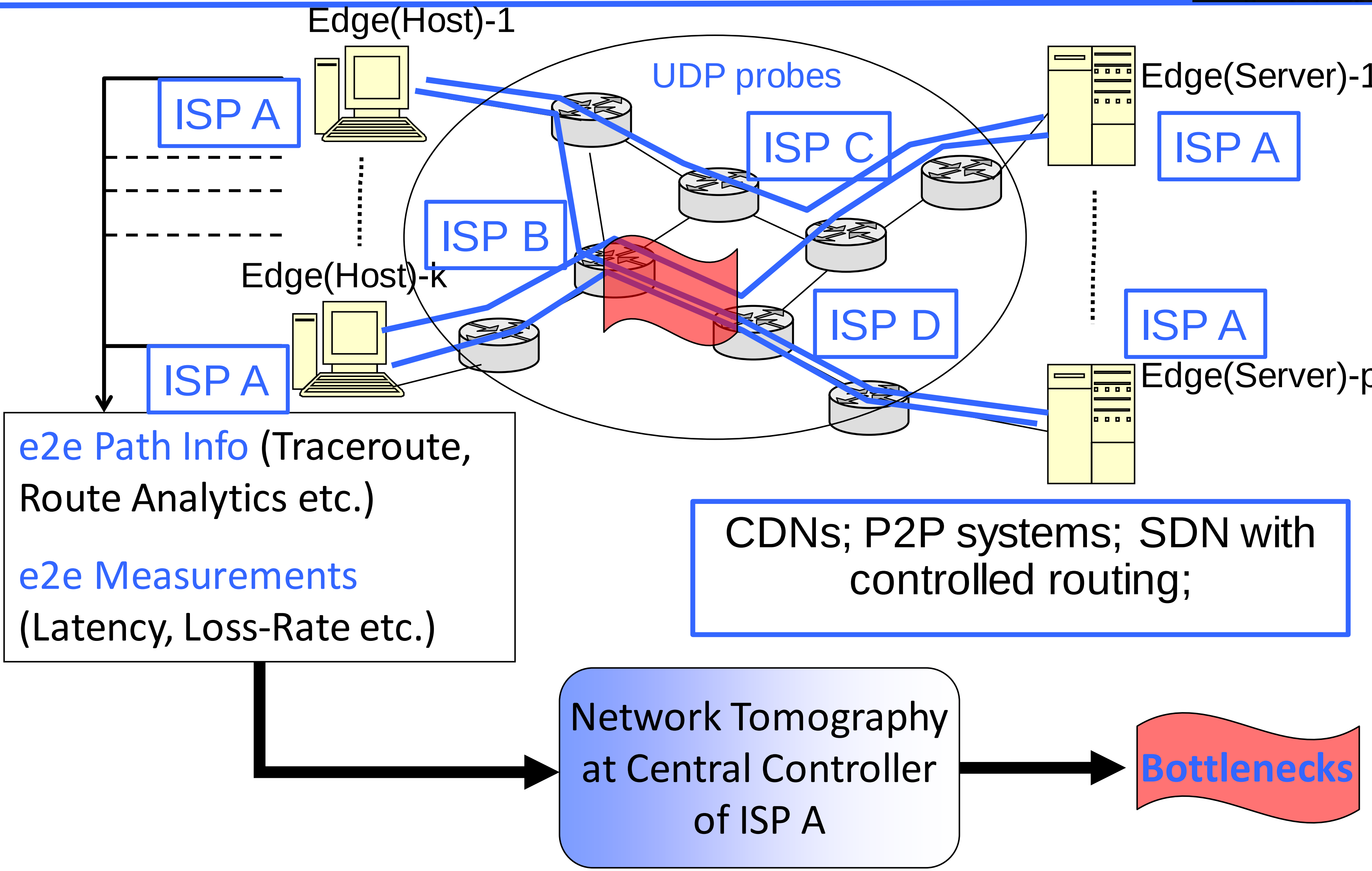
- **Estimating states of links (X)**
  - Given **Y** and **A**
  - Linear algebra and Statistical

  techniques like regularization, EM etc.



Monitors	Paths
(m_1, m_2)	$q_1 = (l_1, l_3)$
(m_1, m_4)	$q_2 = (l_1, l_4)$
(m_1, m_3)	$q_3 = (l_1, l_2)$
(m_2, m_3)	$q_4 = (l_2, l_3)$
(m_2, m_4)	$q_5 = (l_3, l_4)$
(m_3, m_4)	$q_6 = (l_2, l_4)$

Network Tomography: Use Case



Characteristics of End-to-End Measurements

- **BASIS set of e2e measurements**
  - Independent paths; a subset (A_b)
  - Y=AX; A and A_b provide same solution for X
  - No. of paths in basis = RANK of path matrix

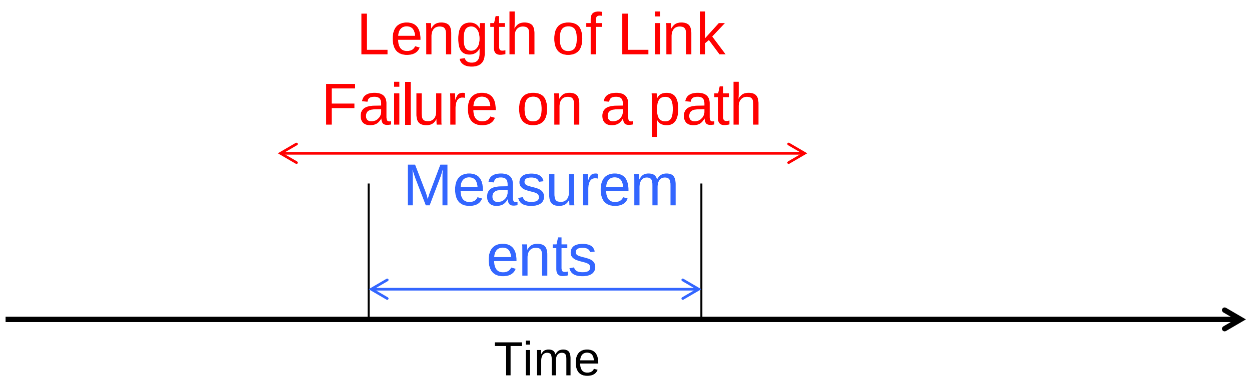
A_b

$$A = \begin{array}{c} \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{array} \begin{array}{c} \begin{array}{cccc} l_1 & l_2 & l_3 & l_4 \end{array} \\ \left(\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right) \end{array}$$

- **Basis and Rank are the performance indicators**

- **Multiple bases are possible**
  - Arbitrary Basis [Chen et al. SIGCOMM’05, Zheng et al. TOC’11]
  - Reduces overhead in collection of measurements

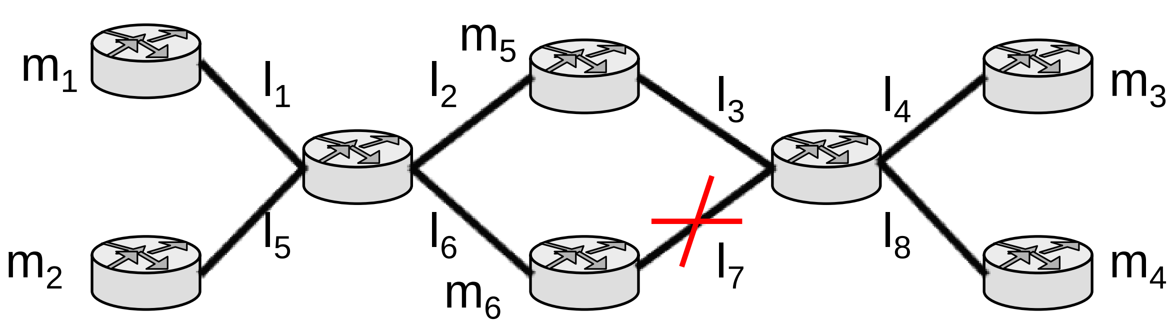
Introduction: Impact of Link Failures



- **Link failures on paths impact e2e measurements**
  - Measurement time-window (order of 10s) [Nguyen et al. IMC'07]
  - Length of IP link failures (order of 100s) [Markopoulou et al. INFOCOM'04]

Impact of Link Failures: Example

- Path Matrix
  - Rank=8 and 8 links;
- Basis $b_1 = (q_1,q_2,q_4,q_{11},q_{15},q_5,q_6,q_7)$
- Basis $b_2 = (q_5,q_6,q_7,q_8,q_9,q_{10},q_{11},q_{12})$
- **Link l_7 fails**
  - b_1: $(q_1,q_2,q_4,q_{11},q_{15},q_5,q_6,q_7)$
  - b_1: Rank = 3
  - b_2: $(q_5,q_6,q_7,q_8,q_9,q_{10},q_{11},q_{12})$
  - b_2: Rank = 7



Pair of edge nodes	Paths
(m_1, m_3)	$q_1 = (l_1,l_6, l_7,l_4)$
(m_2, m_4)	$q_2 = (l_5,l_6, l_7,l_8)$
(m_2, m_3)	$q_3 = (l_5,l_6, l_7,l_4)$
(m_1, m_4)	$q_4 = (l_1,l_6, l_7,l_8)$
(m_1, m_2)	$q_5 = (l_1, l_5)$
(m_3, m_4)	$q_6 = (l_4, l_8)$
(m_1, m_5)	$q_7 = (l_1, l_2)$
(m_2, m_5)	$q_8 = (l_5, l_2)$
(m_5, m_3)	$q_9 = (l_3, l_4)$
(m_5, m_4)	$q_{10} = (l_3, l_8)$
(m_2, m_6)	$q_{11} = (l_3, l_7)$
(m_1, m_6)	$q_{12} = (l_1, l_6)$
(m_2, m_6)	$q_{13} = (l_5, l_6)$
(m_6, m_3)	$q_{14} = (l_7, l_4)$
(m_5, m_6)	$q_{15} = (l_7, l_8)$

Robustness metric: Expected Rank

- **Link failures**
  - Link failures are independent
  - Failure probabilities are known for each link: Bernoulli R.V
  - Failure scenarios where single or multiple links fail; Probability distribution:

$$\mathbf{P}(v) = \prod_{i=1}^{|E|} (p_i v[i]) + (1 - p_i)(1 - v[i])$$

vector v : status of links in a failure scenario, p_i : failure probability of link l_i

- **Robustness of measurements**
  - Expected Rank of a set of paths **R** over all failure scenarios (**$2^{|E|}$**):

$$ER(\mathbf{R}) = \sum_{v \in \{0,1\}^{|E|}} r(\mathbf{R}_v)\mathbf{P}(v)$$

r(): rank function, $\mathbf{R}_v$: set of paths available under failure vector v

Robust Network Tomography: Problem

INSR

Problem: How to pick a general set of paths that provides robust and inexpensive measurements for network tomography?

- **Budget-constrained Optimization Problem**

Given set of paths R_M, **ER** : $2^{R_M} \rightarrow R^+$, the probing cost, **PC** : $2^{R_M} \rightarrow R^+$, and budget B, find $R^* \subseteq R_M$ such that:

$$\mathbf{R^* = \text{argmax } ER(R)}$$

$$\mathbf{R^* \subseteq R_M, PC(R) \leq B}$$

ER: Expected Rank

PC: Probing cost is equal to sum of costs of individual paths

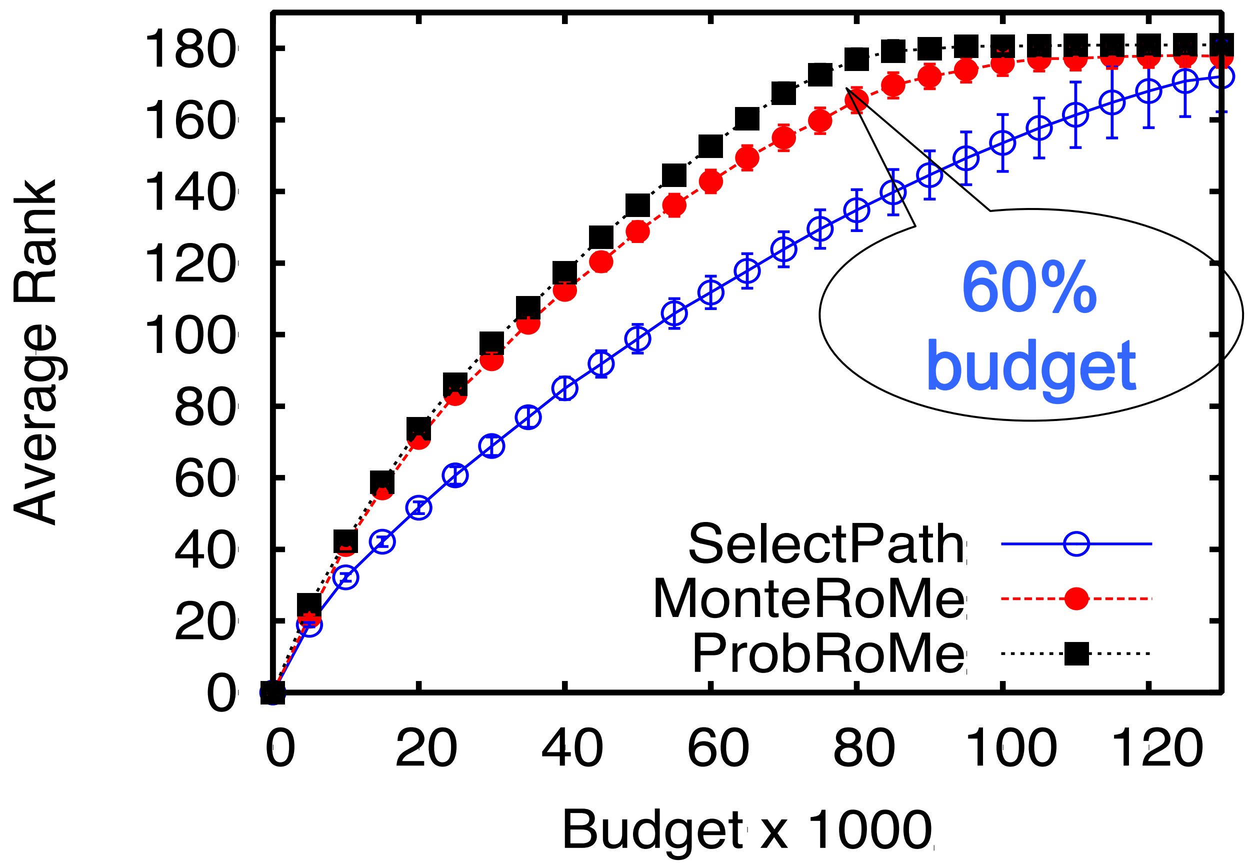
Robust Network Tomography: Solution

- **Theorem.** The budget-constraint optimization problem is NP-Hard
- **Solution:** With known statistical knowledge of link failures and unknown statistical knowledge of link failures (Reinforcement Learning)

- **RoMe** (**Ro**bust **Me**asurements)
  - Greedy based approach
  - Polynomial complexity with probabilistic approximation of ER
  - Solution has tight approx. bound
- **Basis/Linear independence constraint (state-of-the-art)**
  - RoMe gives an optimal solution (Theory of matroids)

Robust Network Tomography: Results

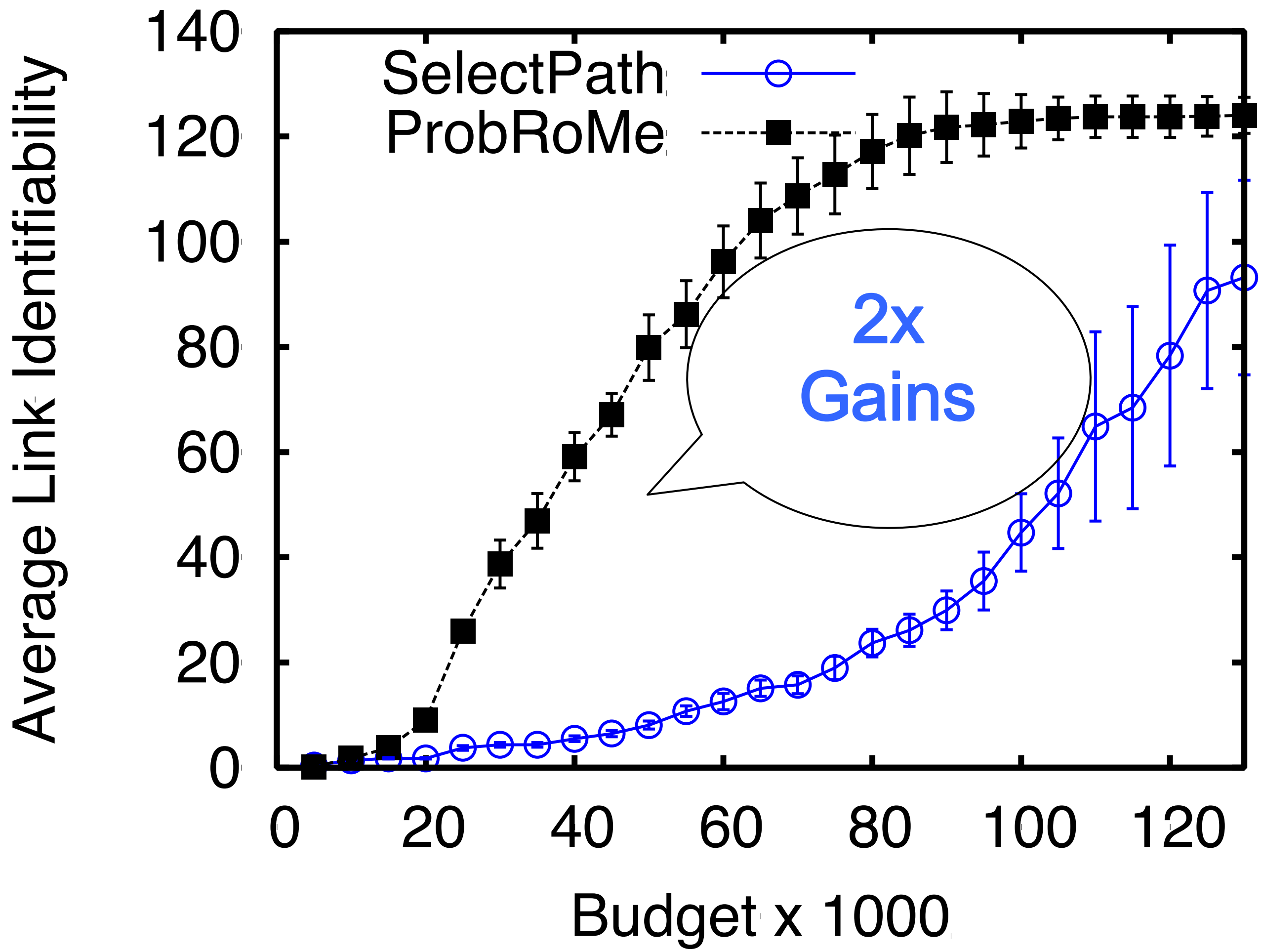
Realistic Topology with 161 nodes and 330 links; Candidate Paths=1600; Realistic link failure model



Robust Network Tomography: Results

Link Identifiability: No. of links for which you can find the unique solution.

It is less than or equal to rank



Conclusion and Future Work

- **Impact of failures on network monitoring technique**
  - Communication between edge nodes get affected while collecting measurements
  - Selection of robust measurements in network tomography
  - RoMe: An efficient solution
  - RoMe has 2x gains for Link Identifiability metric

- **Future Work**
  - Maximizing link identifiability directly; a new metric like *Expected Rank*
  - Probe Sampling and Probe Selection: Joint optimization problem for Loss Tomography application

Thank you!

Questions?

INSR

BACKUP SLIDES

Solution: Statistical Knowledge

- **Expected Rank (ER) is submodular and montonically increasing function**
  - Rank function is a classic submodular function
  - Budgeted submodular maximization problem [S. Kuller et al., A. Krause et al.]

Statistical Knowledge: RoME

- **Greedy Algorithm**

– Iteratively adds paths to R_{out} from candidate paths R_M based on greedy heuristic

$$w_q = \frac{ER(R_{out} \cup \{q\}) - ER(R_{out})}{PC(\{q\})}$$

- **Major Limitation**

– ER function is called $O(|R_M|^2)$ times; $|R_M|$: total candidate paths

– Computation of ER has exponential complexity

$$O(ER(R)) = 2^{|E|}O(r(R)) = 2^{|E|}O(|E| \times r^2)$$

|E|: total links; r():rank func.;
r: rank of path matrix

$$ER(\mathbf{R}) = \sum_{v \in \{0,1\}^{|E|}} r(\mathbf{R}_v)\mathbf{P}(v)$$

Statistical Knowledge: Computation of ER

- **MonteRoMe: Monte Carlo method**
  - Generate failure scenario samples according to their probabilities
  - Not very accurate with few samples
- **ProbRoMe: Probabilistic Approximation**

$$E(Z_{\mathbf{R}}) \leq \sum_{w \in \mathbf{R}_{ind}} E(X_w) + \sum_{q \in \mathbf{R}_{dep}} E(D_q)$$

  - General set **R**: linearly independent ($\mathbf{R}_{ind}$) and dependent paths ($\mathbf{R}_{dep}$)
  - $\mathbf{R}_{ind}$: If available, rank=1, o.w 0; $\mathbf{R}_{dep}$: Not straight forward
  - **Equality:** One linearly dependent path ($q \in \mathbf{R}_{dep}$)

$Z_{\mathbf{R}}$	Random variable for rank of set **R**
X_w =1 or 0	Path **w** $\in \mathbf{R}_{ind}$ is available
D_q =1 or 0	**q** is up and at least one path in $\mathbf{R}^q_{ind}$ is failed

$$O(ER(\mathrm{R})) = \cancel{2^{|E|}}O(r(\mathrm{R})) = \cancel{2^{|E|}}O(|E| \times r^2)$$

Computational Complexity

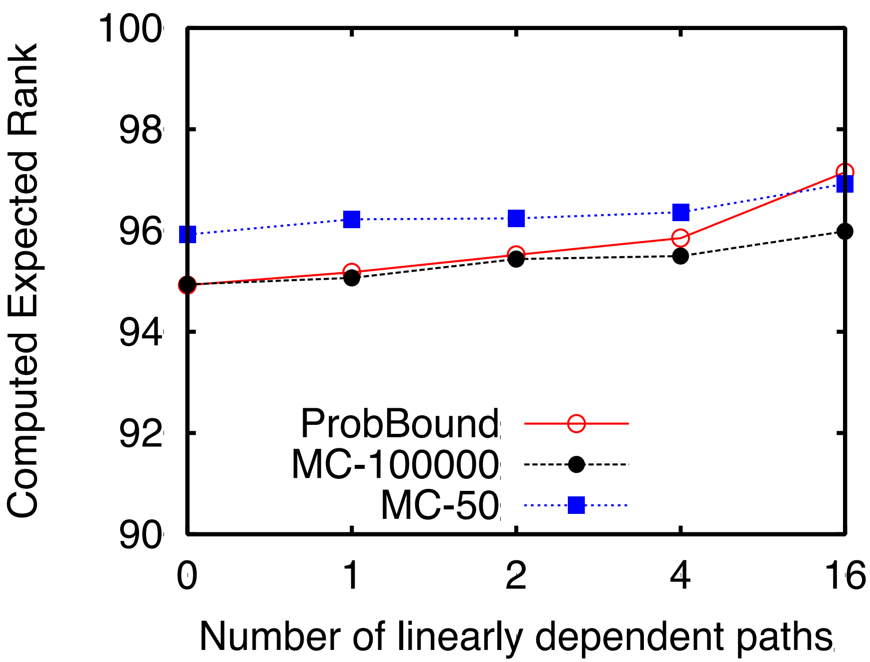
Statistical Knowledge: Computation of ER

- **Approximation Bound of RoMe [A. Krause et al.]**

$$(1-\frac{1}{\sqrt{e}})*OPT-\frac{2B}{C_{\min}}\times\varepsilon$$

ε: Error due to prob. approximation, *OPT*: Optimal value, *B*: Budget; C_{min}: Min. probing cost

- **Empirical Analysis**

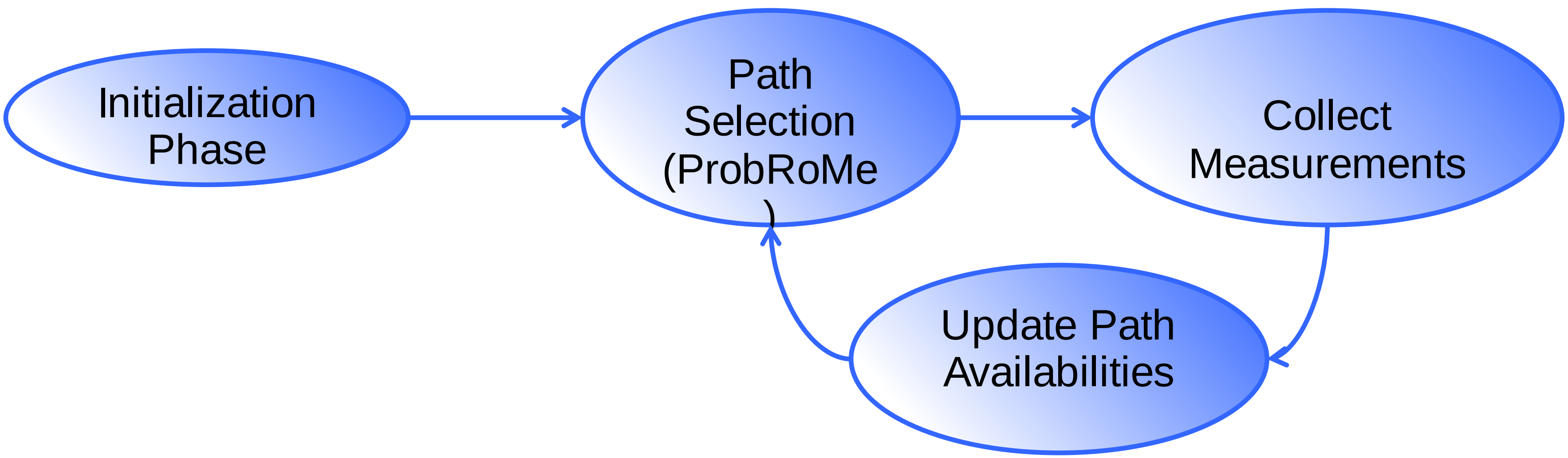


- 99 linearly independent paths+variable dep. paths
- Tight upper bound for **ER(R)** with few dependent paths

- Suits RoMe
- Maximize Rank: Picks few linearly dependent paths in initial iterations

Solution: Unknown Statistical Knowledge

- **Reinforcement learning approach**
  - Different epochs
  - While collecting measurements, observe path status and learn path availabilities (θ)
  - Path selection in each epoch: Use expected path availabilities ($\hat{\theta}$) with ProbRoMe



Solution: Unknown Statistical Knowledge

- **Performance Analysis**
  - Reward and Regret after each epoch
  - Linear reward is common; Submodular reward: Rank
  - Regret at epoch *n* is cumulative difference between optimal action and current action till epoch *n*
  - Upper bound on Regret at an epoch *n*

$$O\left(\frac{\Delta}{\delta^2} N L^3 \log n\right)$$

$$\Delta := \max_{\mathsf{R} \in \mathcal{A}} \left(ER(\mathsf{R}^*; \theta) - ER(\mathsf{R}; \theta) \right)$$

$$L = \max_{\mathsf{R} \in \mathcal{A}} |\mathsf{R}| \qquad \delta := EA(\mathsf{R}^*; \theta) - \max_{\mathsf{R} \in \mathcal{S}} EA(\mathsf{R}; \theta)$$

Robust Network Tomography: Evaluation

- **Simulation Setup**
  - Realistic topologies from Rocketfuel
  - Randomly select edge nodes (monitors)
  - Realistic link failure model [2]
- **Budget-constrained: ProbRoMe, MonteRoMe (50 samples) and SelectPath (modified) [1]**
- **Performance Evaluation**
  - Sample failure scenarios through random generation (500)
  - Evaluate Rank and Link Identifiability (direct application)

[1]Y. Chen, D. Bindel, H. Song, and R. H. Katz, "An algebraic approach to practical and scalable overlay network monitoring," *ACM SIGCOMM Comp. Com. Rev.*, vol. 34, no. 4, pp. 55–66, 2004.

[2] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot, "Characterization of failures in an IP backbone," *IEEE INFOCOM*, vol. 4, pp. 2307–2317, 2004.

Unknown Statistical Knowledge: Results

Medium Topology ; Candidate Paths=400