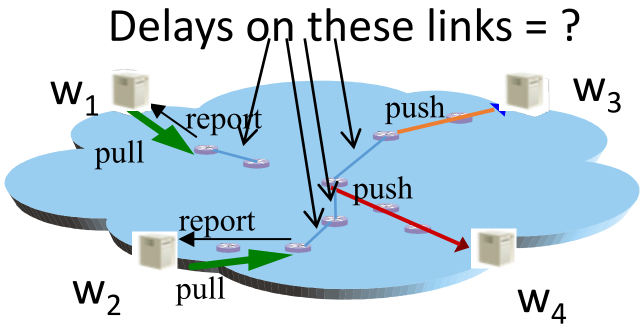
Obtaining Global Network State without Global Access via Network Tomography

Ting He

PennState College of Engineering | INSTITUTE FOR NETWORKING AND SECURITY RESEARCH

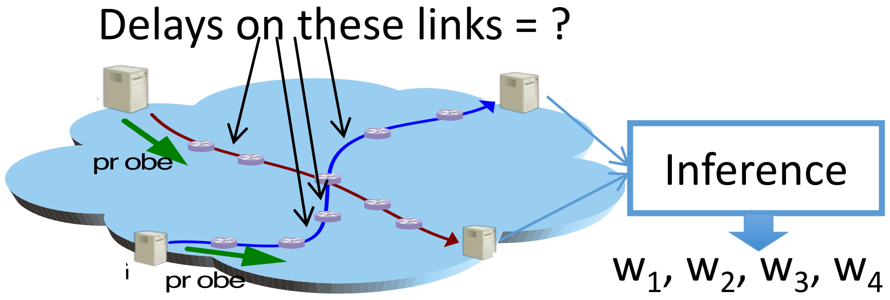
What is “Network Tomography”

- Traditional Monitoring



Require direct access to network elements

- Network Tomography



Do not require direct access (except for monitors)

- Why the name

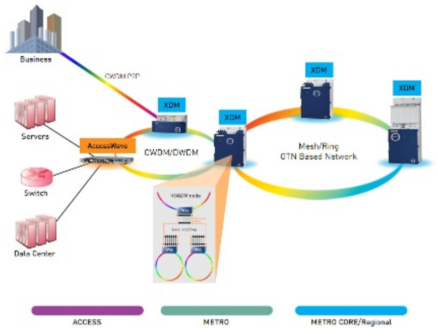
Tomography refers to imaging by sections or sectioning, through the use of any kind of penetrating wave.

Network tomography refers to inferring *internal* network state from *external* measurements

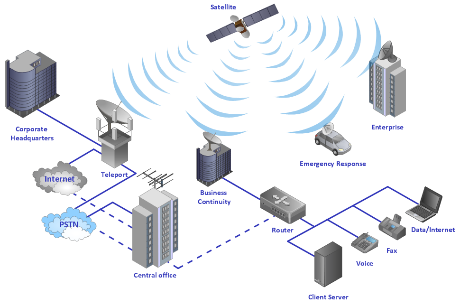
What is it good for

- Generally, anytime we want to **monitor a "closed" network**
  - Internet: Monitor an ISP from client networks, or monitor a peer ISP from an ISP

  - Optical networks: Monitor optical components from light waves

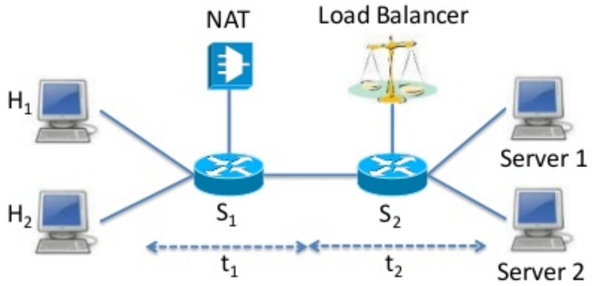


  - Hybrid networks: Monitor LTE network from the edge, or monitor the edge from LTE network for traffic engineering

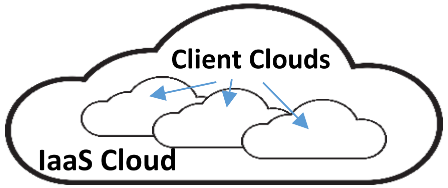


What is it good for cont’d

- Generally, anytime we want to **monitor a “closed” network**
  - ➢ Network with middle boxes: Detect middle box-related performance issues



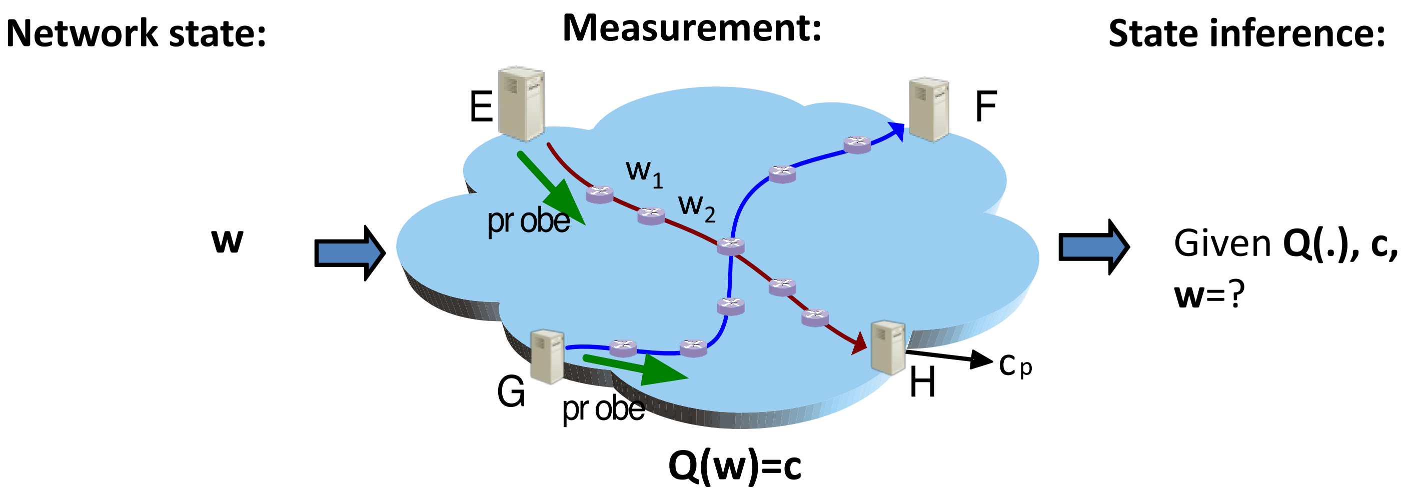
  - ➢ IaaS Clouds: Verify service level agreement (SLA) with cloud provider



  - ➢ Coalition networks: Optimize service placement based on coalition network state

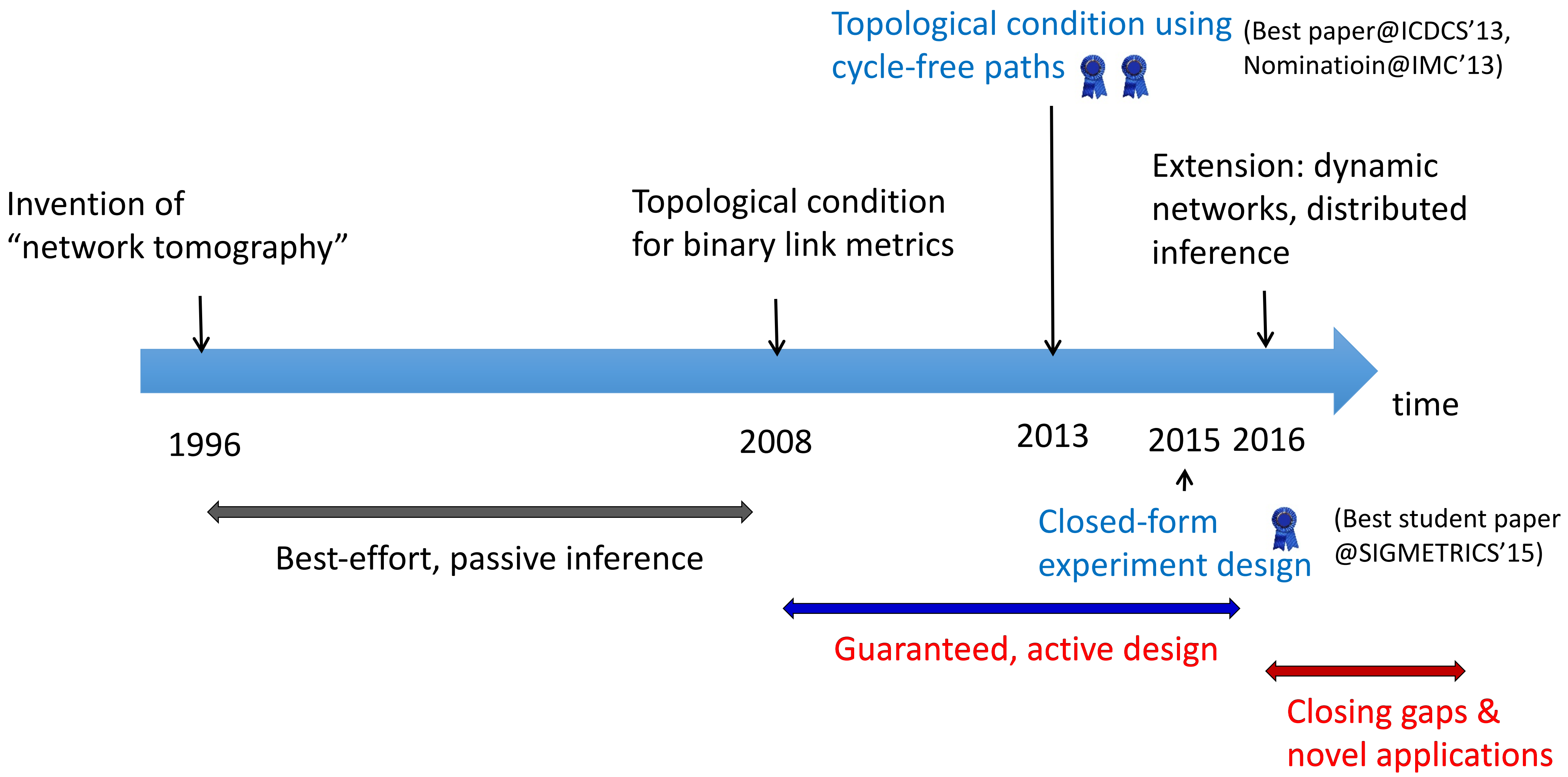
How does it work

- **Idea:** *Inverting* the mapping from internal network state to external measurements.



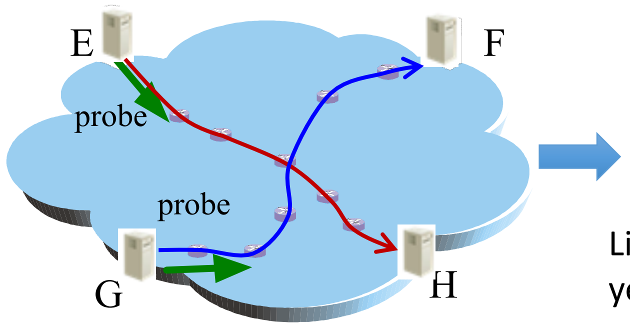
- Examples:
  - Infer link delay/jitter: **Q(w)** = $w_1+\ldots+w_n$
  - Infer link delivery rate (loss): **Q(w)** = $w_1*\ldots*w_n$
  - Infer link available bandwidth: **Q(w)** = $\min(w_1,\ldots,w_n)$

What has been done

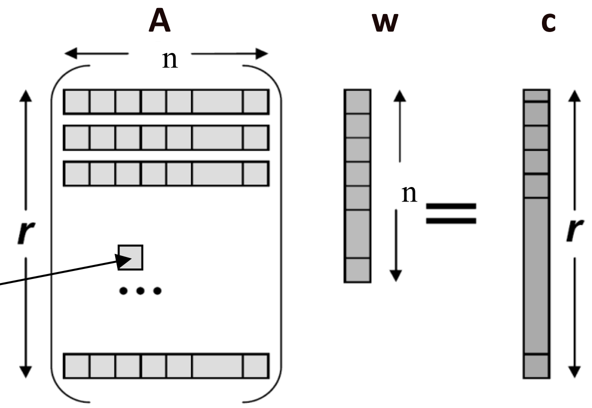


Topological Condition

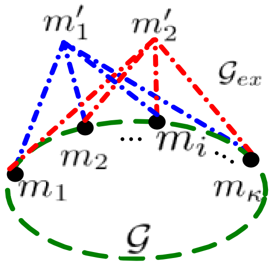
Assume: *Constant* additive link metrics (e.g., mean delay/jitter, log delivery ratio).



Link i on path j?
yes – 1, no – 0



Given measurement matrix **A** and path metrics **c**, link metrics **w** =?



Algebraic condition:

$\mathcal{G}$ is identifiable iff rank(A) = #links

- Difficult to verify
- No insight

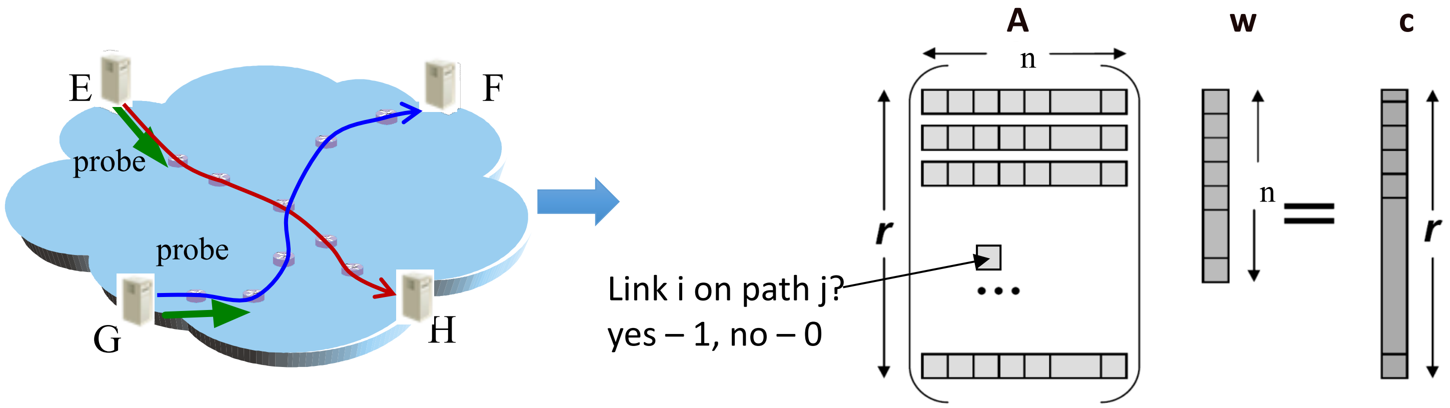
Topological condition:

$\mathcal{G}$ is identifiable iff $\mathcal{G}_{ex}$ is 3-vertex-connected

- Easy to verify
- Guide monitor placement & path construction

Experiment Design

Assume: *Stochastic* additive link metrics (e.g., loss, delay, jitter).



w ~ f(**w**; **θ**)

How to allocate probes among paths to minimize error in estimating **θ**?

Experiment design $\boldsymbol{\phi} := (\phi_y)_{y=1}^{|P|}$ (ϕ_y: prob. probing path y)

Ground truth $\boldsymbol{\theta}$ → Likelihood $f(x,\, y;\, \boldsymbol{\theta},\, \boldsymbol{\phi})$ → Measurements $\boldsymbol{x}$ → Estimate $\hat{\boldsymbol{\theta}}$

Optimal design **ϕ** that maximizes the “information” about **θ**

Experiment Design cont'd

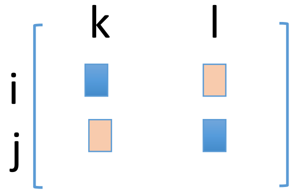
D-optimal

$$\max_{\boldsymbol{\phi}} \det I(\boldsymbol{\theta};\ \boldsymbol{\phi})$$

A-optimal

$$\min_{\boldsymbol{\phi}} Tr\left(I^{-1}(\boldsymbol{\theta};\ \boldsymbol{\phi})\right)$$

$$I_{i,k}^{(y)}(\boldsymbol{\theta})I_{j,l}^{(y)}(\boldsymbol{\theta}) = I_{i,l}^{(y)}(\boldsymbol{\theta})I_{j,k}^{(y)}(\boldsymbol{\theta})$$



$$\det I(\boldsymbol{\theta};\ \boldsymbol{\phi}) = B(\boldsymbol{\theta}) \prod_{y=1}^{|L|} \phi_y$$

$$Tr\left(I^{-1}(\boldsymbol{\theta};\ \boldsymbol{\phi})\right) = \sum_{y=1}^{|L|} \frac{1}{\phi_y} A_y(\boldsymbol{\theta})$$

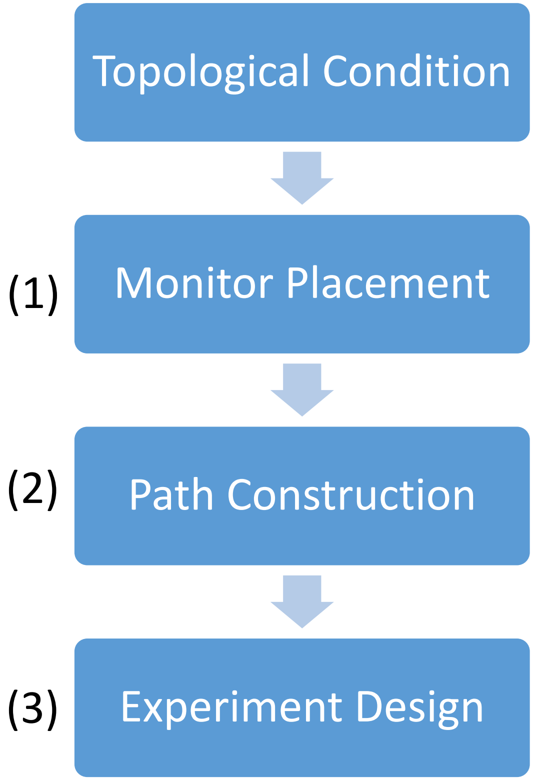
$$\phi_y = \frac{1}{|L|} \quad \text{for } y = 1, 2, \ldots, |L|$$

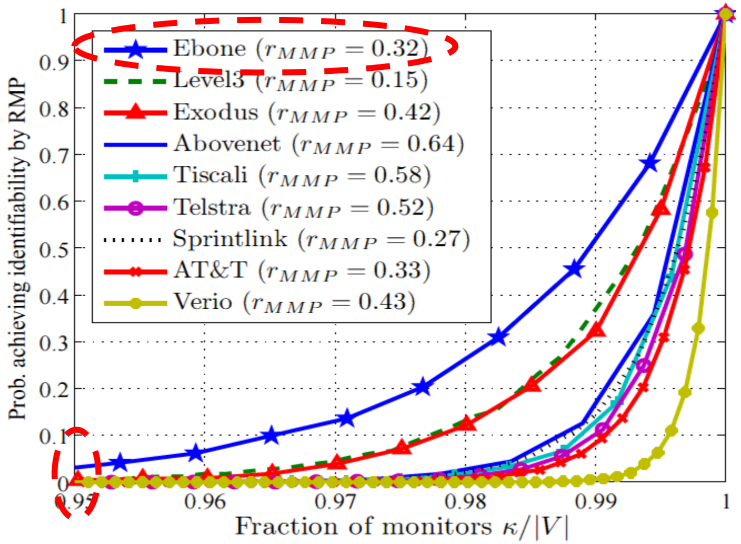
$$\phi_y = \frac{\sqrt{A_y(\boldsymbol{\theta})}}{\sum_{i=1}^{|L|} \sqrt{A_i(\boldsymbol{\theta})}} \quad \text{for } y = 1, 2, \ldots, |L|$$

- D-optimal design = uniform probing
- A-optimal design ≠ uniform probing

* Condition satisfied for loss tomography and jitter tomography

How much is the improvement

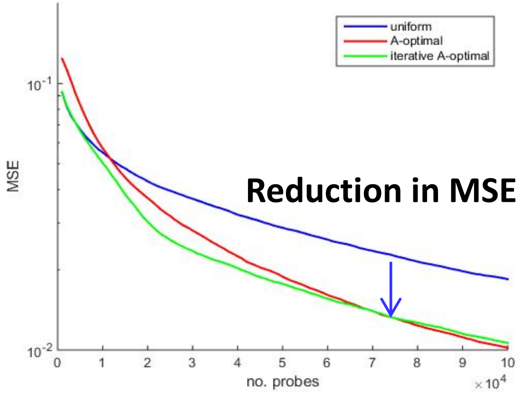


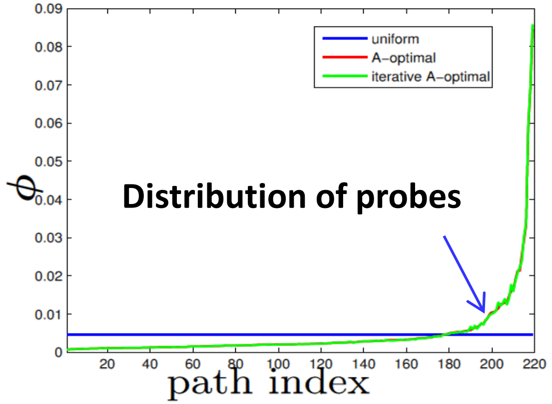


(1) Use 36-85% fewer monitors

ISP	n	m	κ	r_{succ}	Υ	t_{STPC} (s)	t_{RWPC} (s)
Abovenet	294	182	117	80.00%	99.61%	10.12	58.20
EBONE	381	172	55	75.00%	99.69%	13.65	139.37
Tiscali	404	240	138	70.00%	99.67%	28.07	171.58
Exodus	434	201	85	67.00%	99.76%	21.13	226.15
Telstra	758	318	164	24.00%	99.76%	80.38	2999.96
AT&T	2078	631	208	NA	NA	685.46	131.1 hrs
Sprintlink	2268	604	163	NA	NA	608.18	46.8 hrs
Verio	2821	960	408	NA	NA	697.86	170.3 hrs

(2) Speed up path construction by 6-879x





(3) Reduce error by 40%

What is in the future

- **Advancing the science**
  - Non-additive, non-binary metrics
  - Constraints/heterogeneous costs in monitor placement
  - Load balancing in experiment design
  - Dynamic topology
  - Distributed inference
- **Discovering novel applications**
  - Tomography-based inference in social networks, neural networks, integrated circuits,…
- **Closing the loop**
  - Self-optimized * (self-optimized path selection, self-managed services,…)

Obtaining Global Network State without Global Access via Network Tomography

Ting He
https://sites.google.com/site/thtinghe/

THANK YOU

Backup

What has been done