

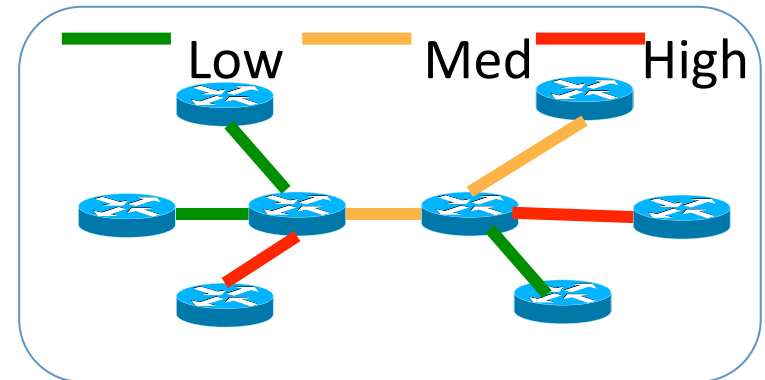
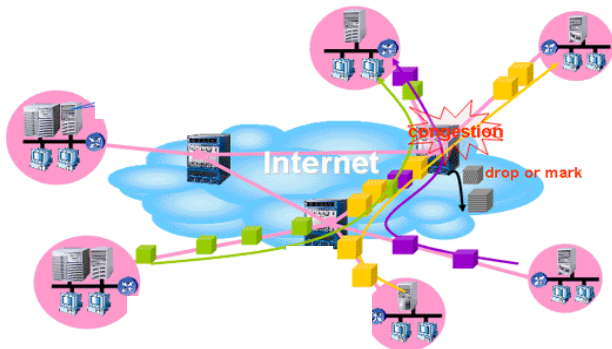
Using Video-Based Measurements to Generate a Real-Time Network Traffic Map

Yi Sun, Junchen Jiang, Vyas Sekar,
Hui Zhang, Fuyuan Lin, Nanshu Wang



**Carnegie
Mellon
University**

Real world has Google Maps, why don't we!



Not a new idea, but has been elusive

- Coverage
 - Need millions of vantage points
- Overhead
 - B/w measurement incurs non-trivial cost
- Real-time views
 - Continuous updates

Opportunity of Internet Video!

- Coverage
 - Need millions of vantage points



→ Video traffic 30-50%, 30M Netflix streaming subscribers

- Overhead
 - B/w measurement incurs non-trivial cost



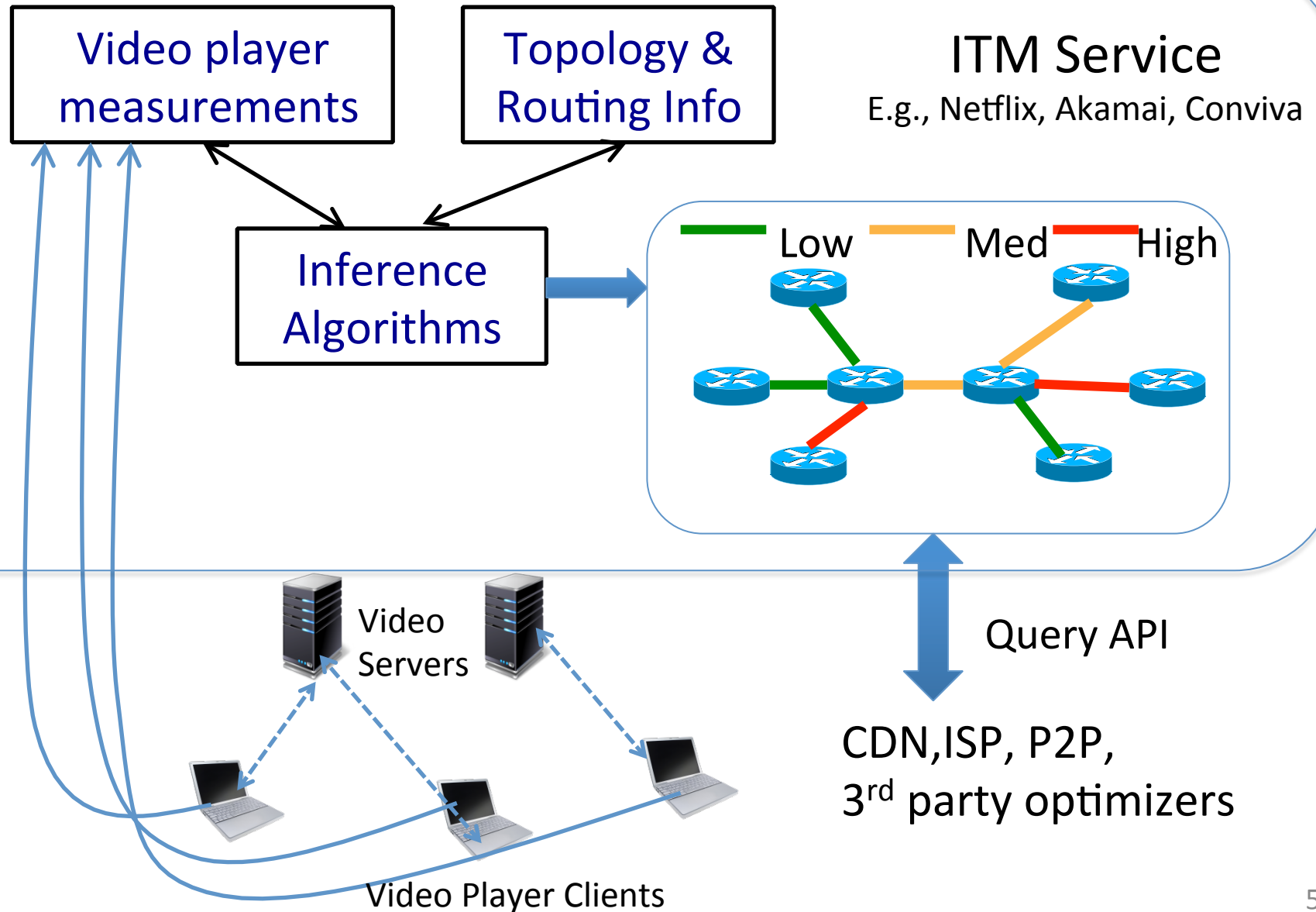
→ Passive throughput measurements

- Real-time views
 - Continuous updates



→ Akamai, Netflix, Conviva, PPLive etc already do this

Internet Traffic Map Service



Concrete Problem Definition

For each epoch, src, dst:

Throughput

Bytes

For each epoch, src, dst:

Path (PoP-level)

e.g., iPlane

Video player
measurements

Topology &
Routing Info

Inference
Algorithms

P1: InferCapacity
Link \rightarrow Capacity

P2: InferUtilization
Link, Epoch \rightarrow Utilization

Concrete Problem Definition

For each epoch, src, dst:

Throughput

Bytes

For each epoch, src, dst:

Path (PoP-level)

e.g., iPlane

Video player
measurements

Topology &
Routing Info

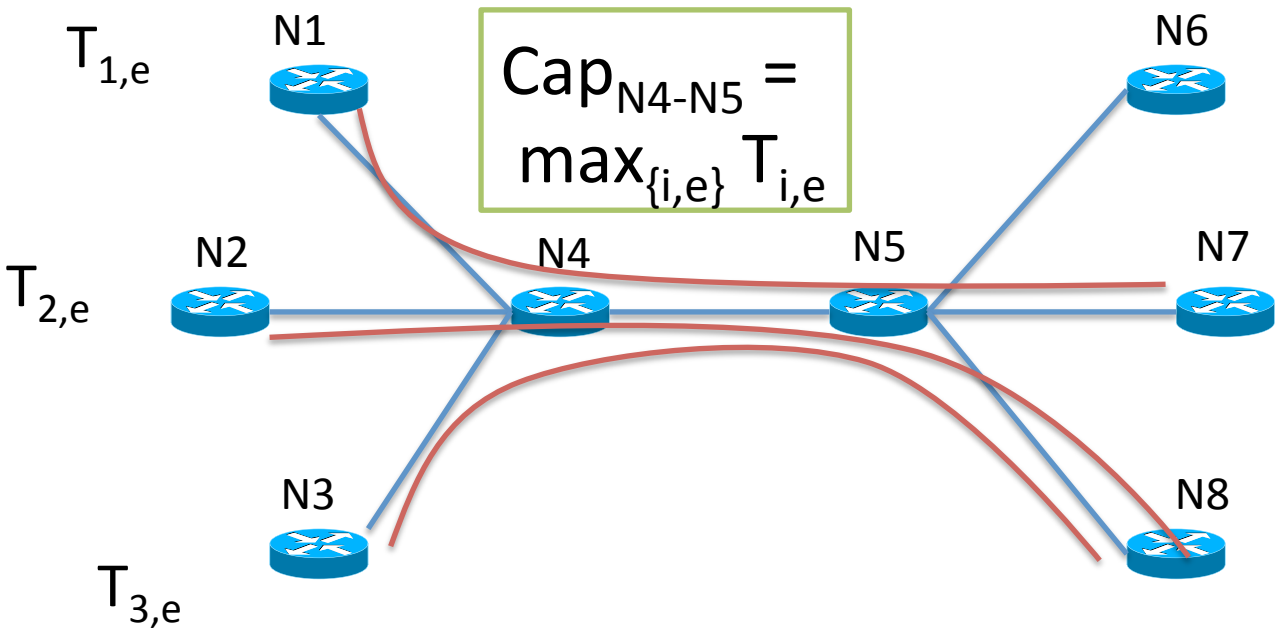
Inference
Algorithms

***P1: InferCapacity
Link → Capacity***

**P2: InferUtilization
Link, Epoch → Utilization**

Strawman 1: Max Estimator

$T_{i,e}$ = throughput measurement in epoch e



Underestimate?

Background?

Discrete values?

Just use max observed throughput?

Strawman 2: Tomography

Underestimate?

Background?

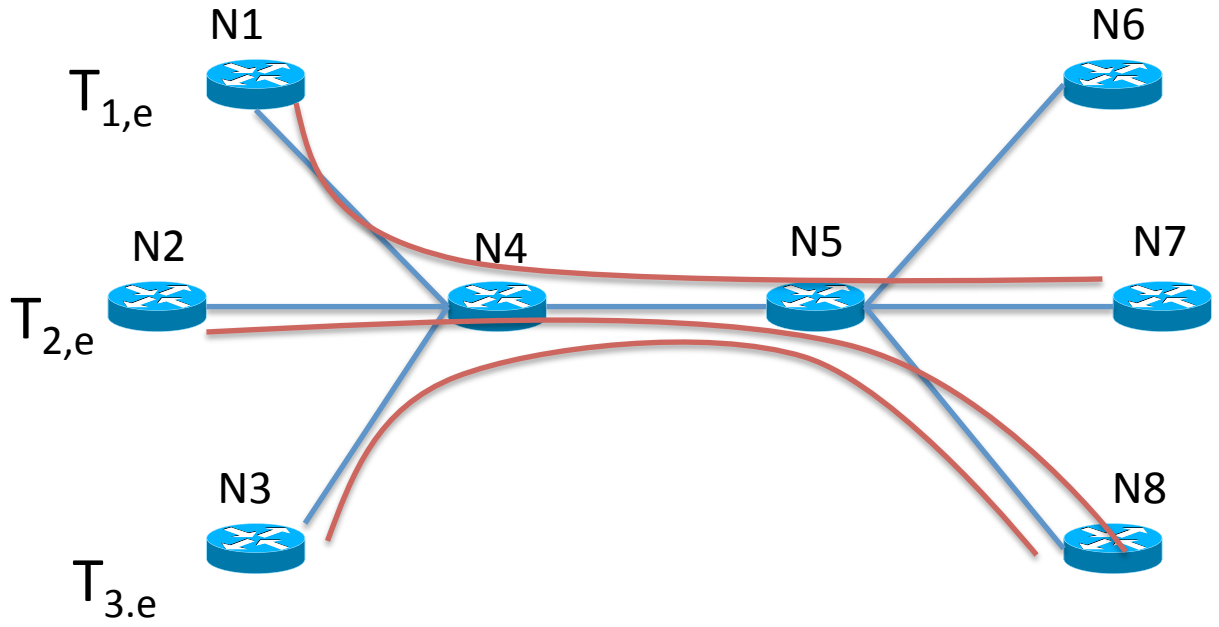
Discrete values?

- 1. Cap_{Ni-Nj} from discrete values
- 2. Hidden background values
 $B_{N\{1,2,3\}-N\{6,7,8\}}$
- 3. Cost-awareness



Constrained Optimization
Output = Cap_{Ni-Nj}
Idea: $Max_e \{B+T, Bytes/Time\}$

Underconstrained!

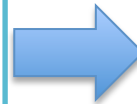


Idea: Add “Side” information

1. “Gravity” assumption
2. B/T ratio
3. Expected overprovisioning



1. Cap_{Ni-Nj} from discrete values
2. Hidden background values
 $B_{N\{1,2,3\}-N\{6,7,8\}}$
3. Cost-awareness



Constrained Optimization
Output = Cap_{Ni-Nj}
Idea: $Max_e \{B+T, Bytes/Time\}$

Underconstrained!

Concrete Problem Definition

For each epoch, src, dst:

Throughput

Bytes

For each epoch, src, dst:

Path (PoP-level)

e.g., iPlane

Video player
measurements

Topology &
Routing Info

Traffic Map
Service

P1: InferCapacity
Link \rightarrow Capacity

P2: InferUtilization
Link, Epoch \rightarrow Utilization

Strawman: Tomography+?

1. “Gravity” assumption
2. B/T ratio
3. Expected overprovisioning



Constrained Optimization
Output = $B_{N_i - N_j}$

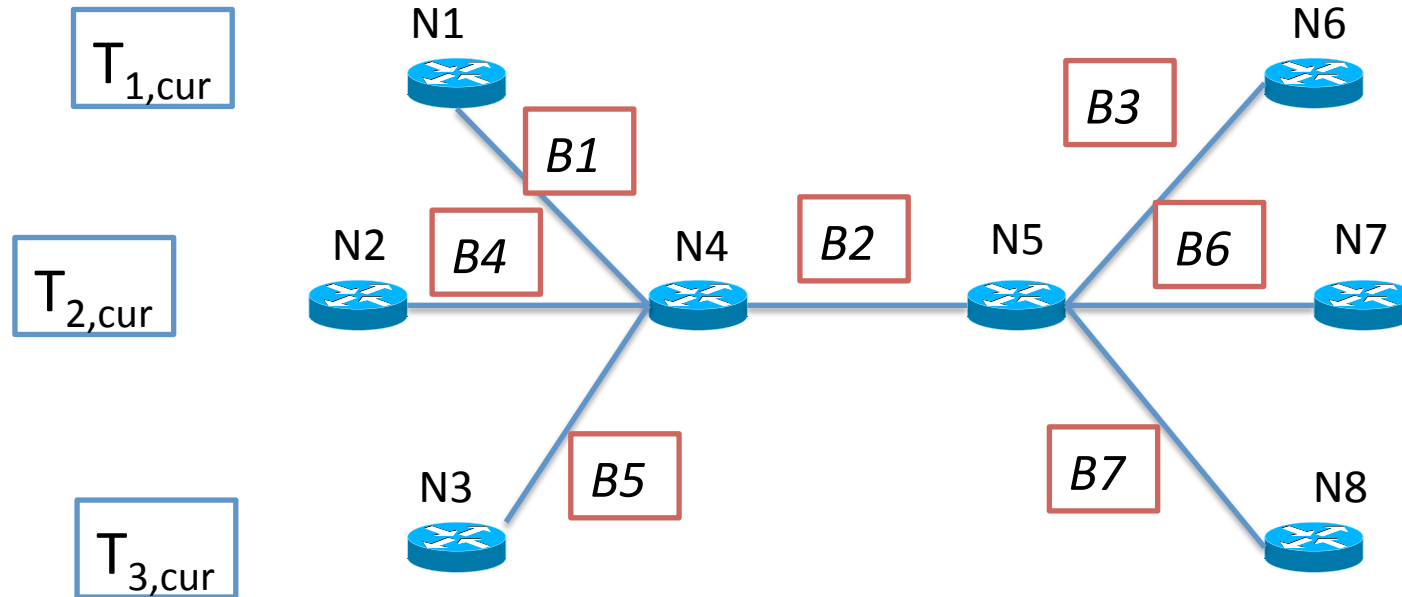
“Aggregate” effects

Don't have history

High-level Idea: Capacity + Max-min fairness

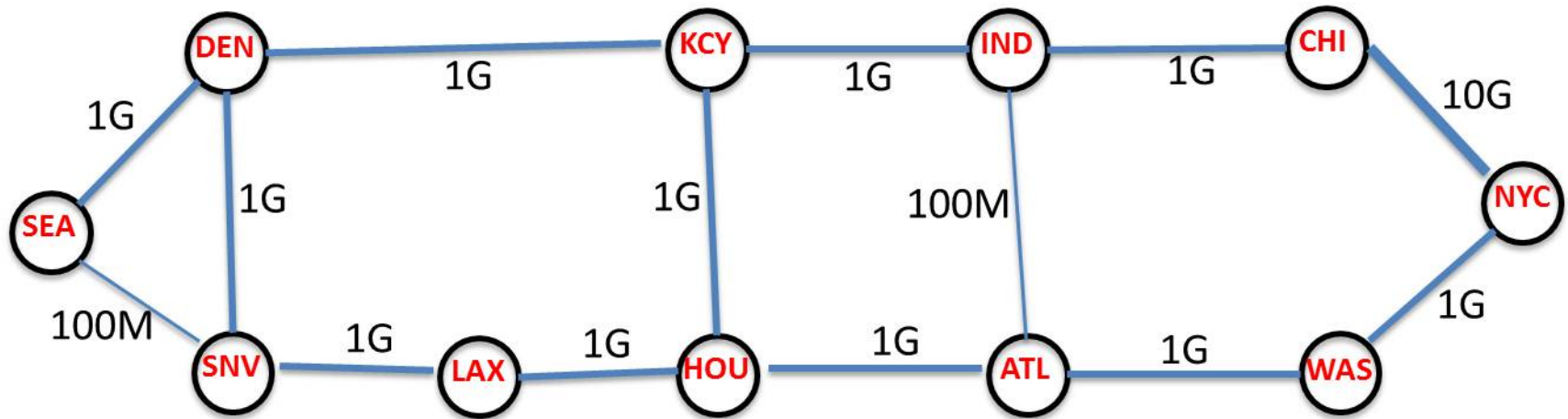
P2: InferUtilization
Link, Epoch \rightarrow Utilization

Since we know capacity from solving P1,
What B values “explain” observed throughputs?



Brute-force search to find a max-likelihood estimator
s.t. predicted throughputs “match” the observed

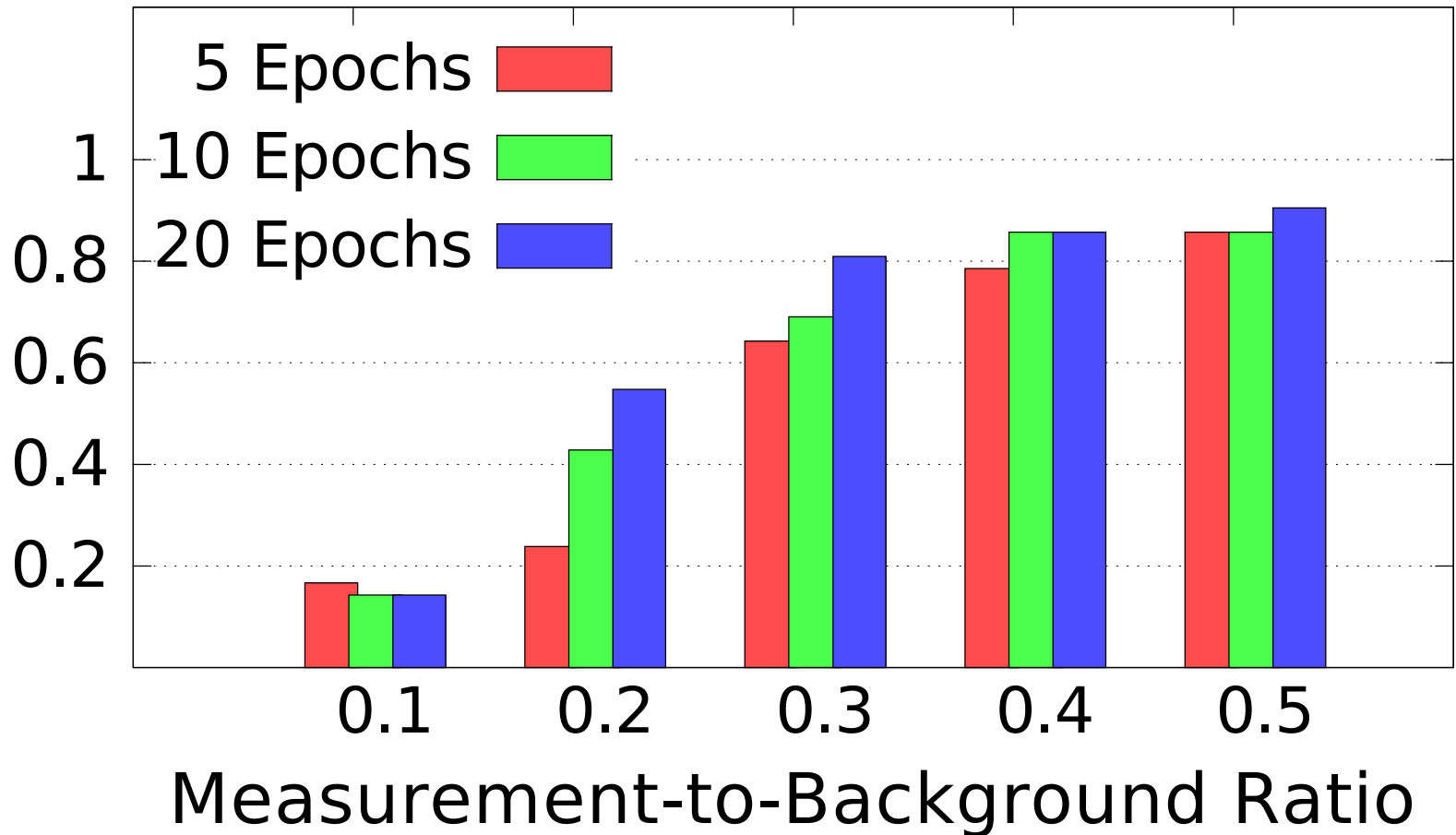
Evaluation Setup



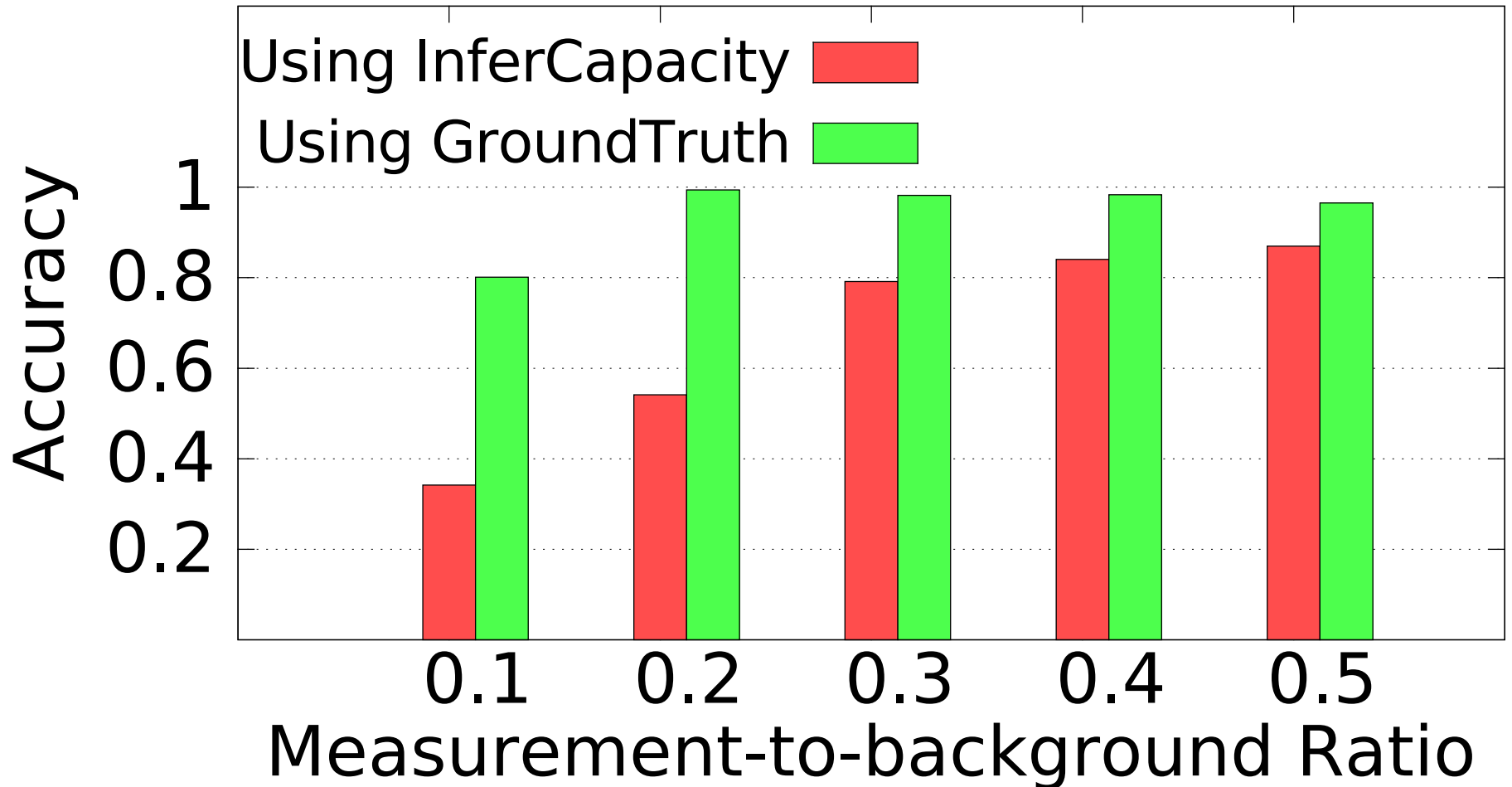
- Custom flow-level simulator
- Sensitivity
 - number of epochs
 - Background vs. measurement traffic ratio
 - Accuracy of capacity inference

Accuracy of Capacity Inference

Fraction of links
with perfect inference



Accuracy of Background Inference



Limitations and Open Issues

- How much side information do we need?
- Other potential “carrier signals”?
- Sensitivity to estimation parameters
- Scalability of the algorithms
- Measurement bias?
- ...

Conclusions

- A real-time *Internet Traffic Map* would be cool
 - Many apps could benefit
- Has always been out of our reach
 - Coverage, Overhead + Real-time view
- Video offers a unique opportunity
 - Capacity and Background Inference via Video
 - Initial results seem promising
- Many limitations, open questions, and future work!