The Predictive Power of Shortest-Path Weight Inference

ANDREW COYLE, MIRO KRAETZL,
OLAF MAENNEL, MATTHEW ROUGHAN

<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

Rocketfuel

THE UNIVERSITY OF ADELAIDE AUSTRALIA

Let's go a reverse engineering, hey!

- do a bunch of traceroutes from as many places, to as many places as possible
- compile them together
- infer
  - topology
  - routing

  of a single AS (at a time)
- invaluable
  - scientific interest
  - for simulations

Network Tomography

Network Tomography can be generally applied to mean solving inverse problems in communications networks.

- link performance (from end-to-end measurements)
- end-to-end traffic matrices (from link loads)
- topology
- routing

Routing Policy Inference

- current routing is implicit in traceroute measurements
- but of limited utility
- doesn't tell you what will happen if something changes
  - thats where the money is
  - also useful for understanding the mind of the "network engineer"
- really need to infer routing policies
  - simplest case is shortest-path routing
  - infer weights

Shortest-path weight inference

Intuition: measured paths must be shortest-paths

Write as optimization problem (actually a LP)

$$\text{minimize } f = \sum_{e \in E} \varepsilon_e,$$

subject to

$$w_e - \varepsilon_e \le d_e, \quad \forall e \in E,$$

$$w_e + \varepsilon_e \ge d_e, \quad \forall e \in E,$$

$$\sum_{e \in \hat{\mu}_{ij}} w_e \le \sum_{e \in \mu} w_e, \quad \forall i, j \in N, \text{ and } \forall \mu \in P_{ij},$$

$$w_e, \varepsilon_e \ge 0, \quad \forall e \in E,$$

where

- w_e are the link weights
- d_e are the links' geographic distances

Rocketfuel as tomography

Rocketfuel technique is a really a type of tomography

- inverse problem
- constraints imposed by measurements
- problem is underconstrained

Need side-information

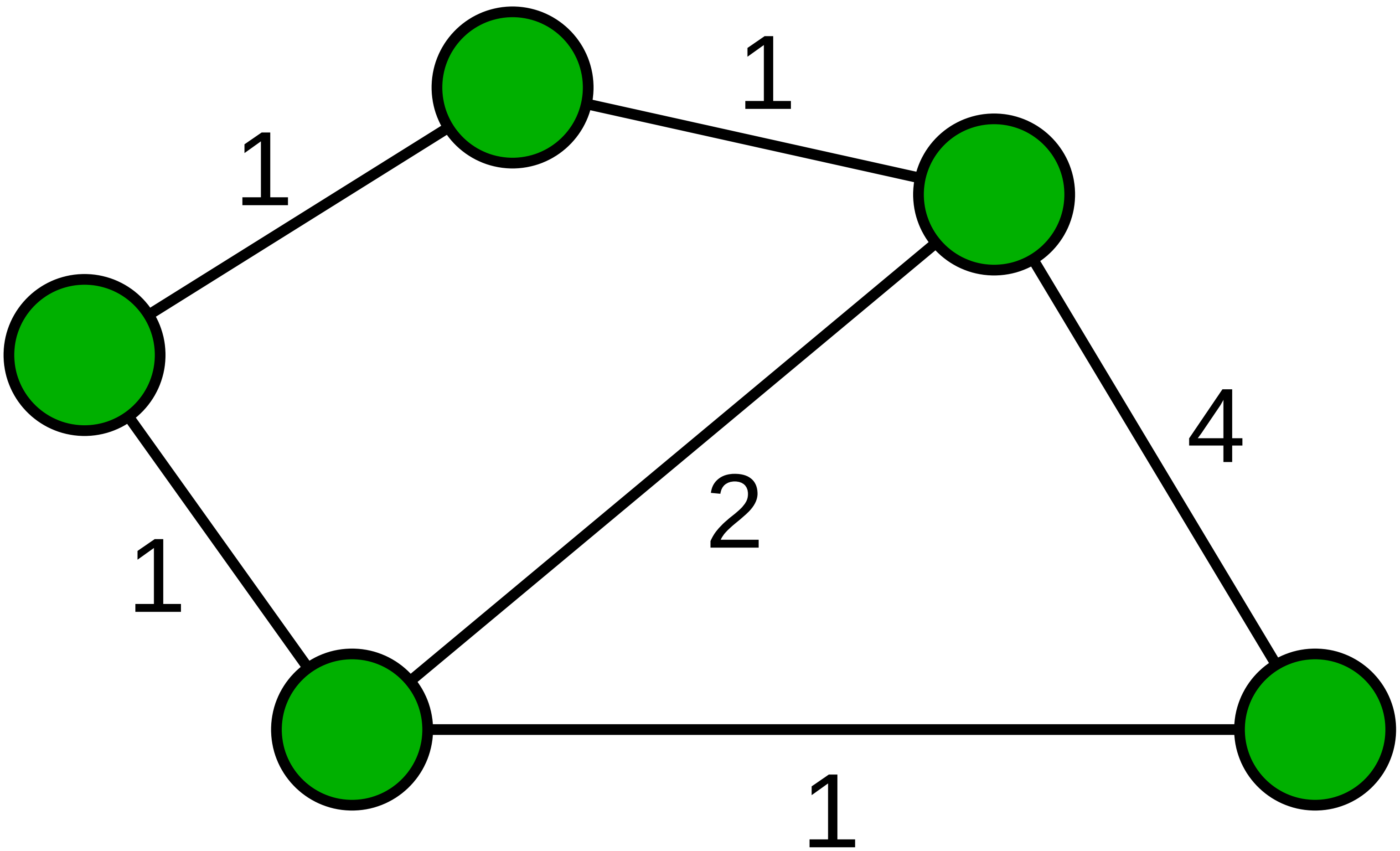
- often called a “prior”
- The Rocketfuel prior is distance proportionality
  - in absence of other information, shortest-path weights should be geographic distance
  - but we know this is wrong

Does it work anyway?

Does it work?

How would we know if it worked?

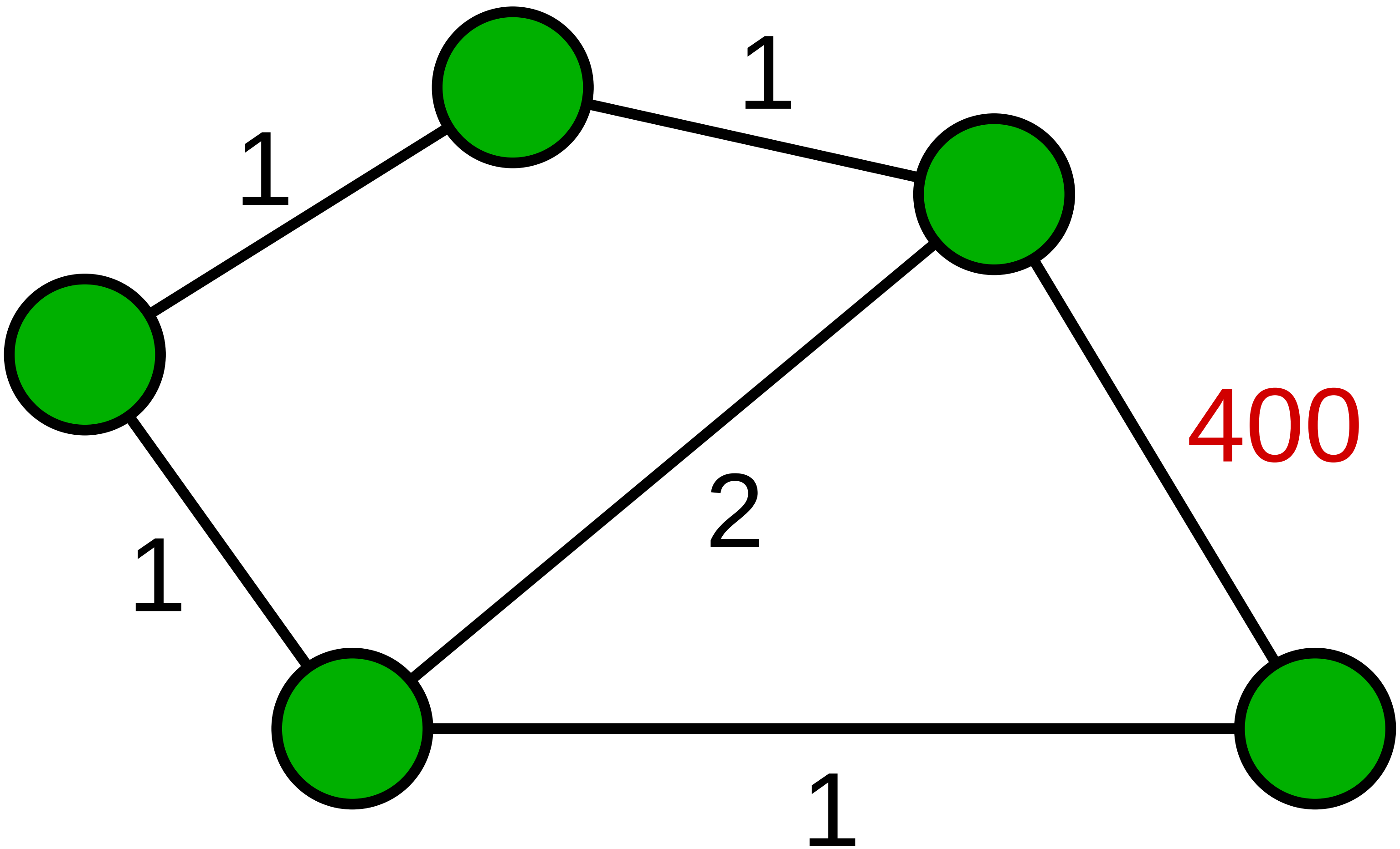
- "accuracy" is meaningless here



Does it work?

How would we know if it worked?

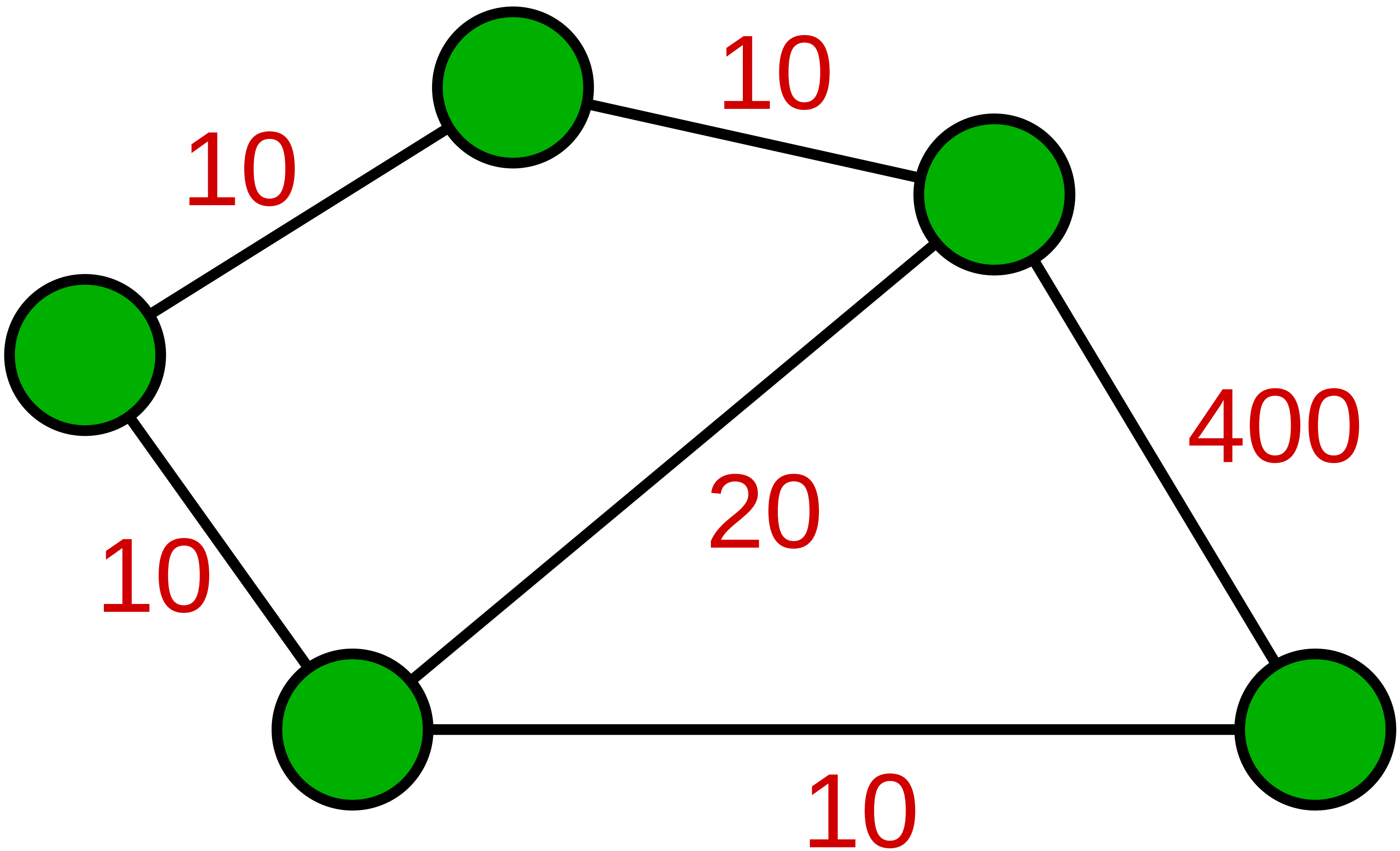
- **we can change a weight, without changing routing**



Does it work?

How would we know if it worked?

- **we can change a whole lot of weights**



Predictive power

What is really interesting is how well we can predict the network behaviour

- obviously has to be behaviour that we don't "see"
  - optimization automatically ensures that weights will fit the observed routing
- two cases considered here
  - unobserved routes (incomplete data)
  - routing after a link failure

Methodology

THE UNIVERSITY OF ADELAIDE AUSTRALIA

- Used real data (Abilene, GEANT)
  - doesn't allow for multiple simulations
  - doesn't allow us to vary real prior
- Combined with simulations
  1. start with a topology (real, or Rocketfuel)
  2. generate a set of traffic
  3. generate sets of weights
     (a) Given weights (some distance proportionality)
     (b) Unit (less distance proportionality)
     (c) Unit plus jitter
     (d) Optimized weights (no distance correlation)
     (e) "Backbone" weights (spanning tree + backup) (very far from proportional)

Unobserved routes results

Predictive power (on average) for 5 (randomly chosen) unobserved routes

Network	weights				
	given	unit	u+j	synthetic	backbone
AS 1	97.3%	95.3%	95.5%	92.9%	78.3%
AS 1239	96.6%	96.4%	96.6%	92.9%	74.2%
GEANT	91.5%	95.4%	94.4%	90.3%	67.8%

- results are reasonable to very good
- real distribution of weight values plays little role, unless it is really extreme

Link failure results

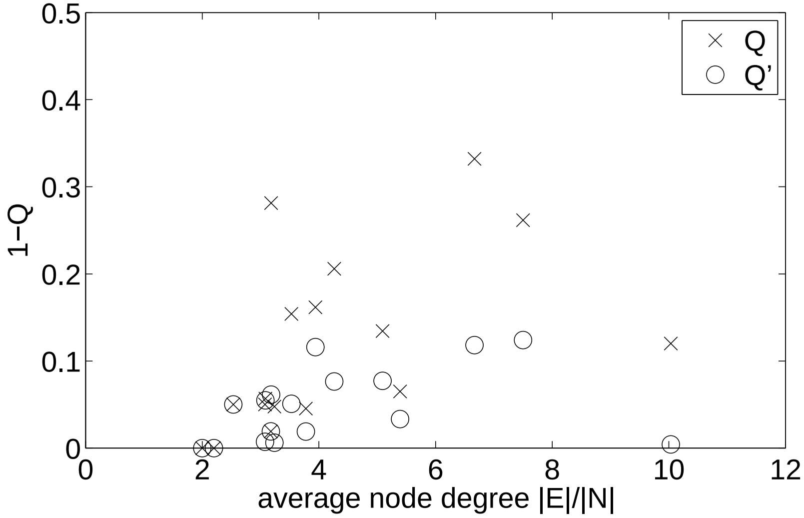
Predictive power for routing after single links failures

Network	weights				
	given	unit	u+j	synthetic	backbone
AS 1	94.4%	99.9%	99.2%	90.9%	69.5%
1239	89.9%	100.0%	94.1%	59.8%	27.3%
GEANT	87.8%	99.7%	94.2%	74.7%	35.5%

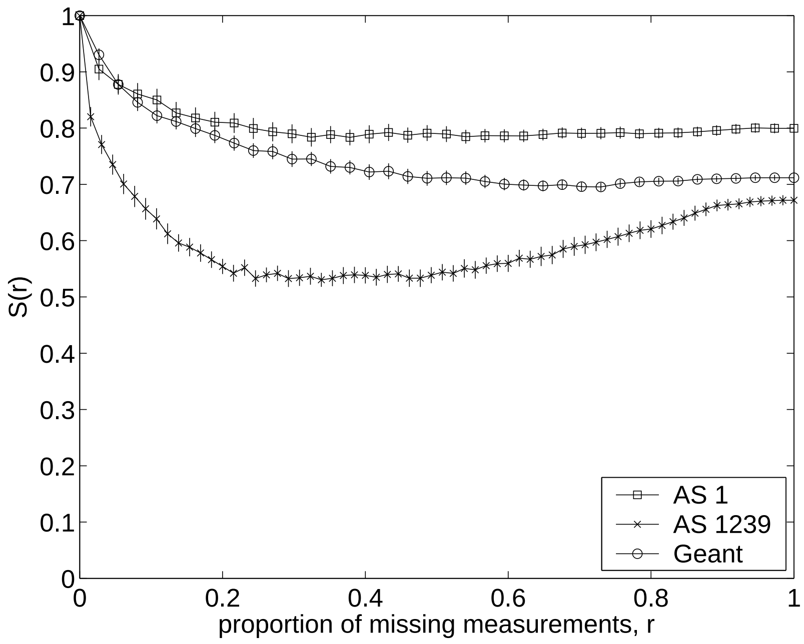
- harder task
  - most cases perform worse than before
- now, weight distribution plays more of a role
  - weights further from distance perform worse

Other results

Topology dependence



Information reversal



Conclusion

THE UNIVERSITY OF ADELAIDE AUSTRALIA

- Rocketfuel approach isn't bad (in the absence of anything better)
- Predictive power is a useful methodology – not just for this problem but for a range of inverse (tomography) problems where outright accuracy isn't really the important feature

Future work

- improved algorithms
- incorporating topology errors
- further investigation of information reversal

ECMP

Equal-Cost Multiple Paths (ECMP) is important

- effects routing
- effects measurements
- effects inference
- effects interpretation of results

paper lists effects $\pm$ effects of ECMP

Computation Time

- 1.8 Ghz Intel PC
- times are $O(|E|^3)$