Spatial Distribution Networks

Daniel R. Wuellner

April 14, 2009

Why these papers {a OR we}re worth your time

- Infrastructure networks' design emerges from simple optimization
- Flows on networks: network structure/function determines user preferences which in turn have nontrivial effects on network function
- (Cool pictures)

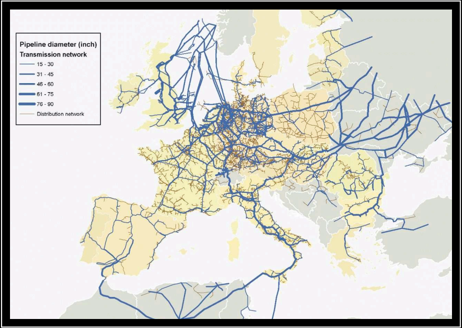
big questions

What's special about spatial networks?

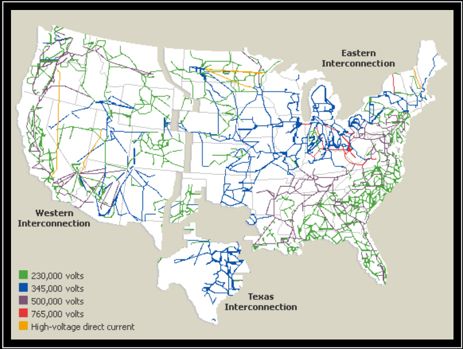
'Networks': all things that look like $G = (V, E)$

- 'Spatial' networks embedded in a network-independent 'space' governing V and E
- $P(i,j) = f((x_i, y_i, z_i), (x_j, y_j, z_j), \ldots)$
- Planar graphs: edges only intersect at nodes
- Usually k_{max} is constrained

Technological spatial network examples

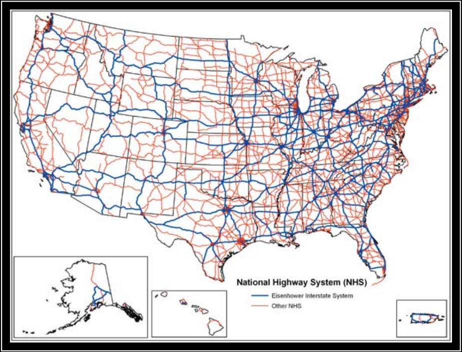


Carvalho et al: arXiv:0903.0195v1 [physics.soc-ph]

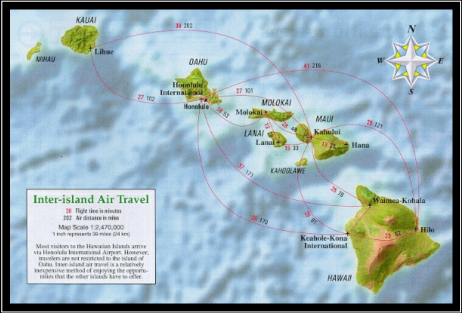


images.encarta.msn.com

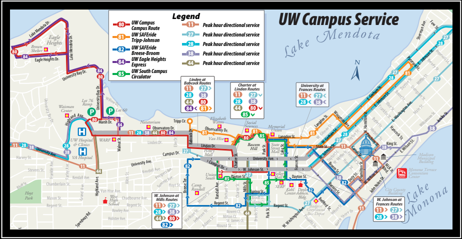
Technological spatial network examples



fhwa.dot.gov

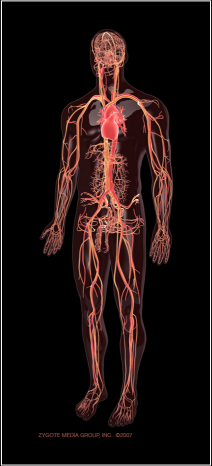


cheaptravel.net



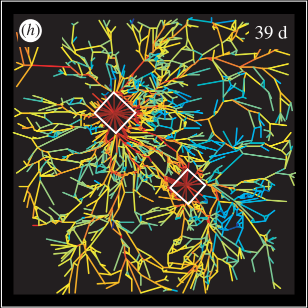
cityofmadison.com/metro

Biological spatial network examples



3dscience.com

davidlnelson.md/Cazadero/CazImages



Bebber et al, Proc R Soc B 2007 **274** 2307-2315

livescience.com/imageoftheday/siod 060508.html

What did I miss?

- Social networks: the farther away we are, the less likely we connect
- Ad-hoc spatial networks: wireless, dynamic traffic networks (as in Zac's model)
- Non-'spatial' 'spatial' networks: hidden metric spaces

What's special about technological networks?

- Intelligent design!
  - Design: technological
  - Selection: biological, social (and selection criteria need not have anything to do with a particular network representation)
  - Chance: RG models
- In particular: usually satisfy one (or few) simple goals: Efficient transportation of something (water, gas, electricity, information, people) from A to B
- Infer connections between structure and function

Aside

'Designed' networks: people have been thinking about transportation and distribution networks for a long time (and there's big money in getting it right)

The challenge for the 'complex networks' community in this domain is to uncover something new, or to 'generalize' existing research, or to identify what's unique or universal in this class of networks

Classic text: Network Flows: Theory, Algorithms and Applications, by Ahuja, Magnanti, Orlin

Spatial Distribution Network Design Problem

SDNDP

Given a set of things that 'want' to move between a set of spatial points V, what's the cheapest way to get them there?

Gastner & Newman (2006) wants us to think about two specific parts of this question as applied to transportation networks:

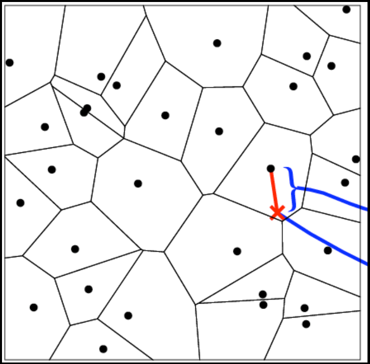
- How do we determine the points V anyway?
- What do we mean by 'cheapest'? Does how we measure cost change the optimal structure?

Optimal facility location

- Assume non-uniform population density $\rho(\mathbf{r})$
- Distribute p facilities to $\{\mathbf{r}_1 \ldots \mathbf{r}_p\}$ to minimize:

$$f(\mathbf{r}_1 \ldots \mathbf{r}_p) = \int_A \rho(\mathbf{r}) \min_{i \in \{1 \ldots p\}} |\mathbf{r} - \mathbf{r}_i| d^2 r$$

- Voronoi cell at $\mathbf{r}_i$: set of points closer to $\mathbf{r}_i$ than to any other $\mathbf{r}_j, j \neq i$



from Gastner slides.pdf

Optimal facility location

- SOLUTION: Optimal density of facilities increases sublinearly with population density:

$$D(\mathbf{r}) = \frac{p}{\int [\rho(\mathbf{r})]^{2/3} d^2 r} [\rho(\mathbf{r})]^{2/3}$$

- Balance between
  - $D(\mathbf{r}) \propto [\rho(\mathbf{r})]^0 =$ constant: Sparse and dense get same
  - $D(\mathbf{r}) \propto [\rho(\mathbf{r})]^1$: Dense get lots, sparse get very few

(math)

- Minimize average distance to nearest facility ($s(\mathbf{r})$ is area of $\mathbf{r}$'s Voronoi cell):

$$f = g \int_A \rho(\mathbf{r})[s(\mathbf{r})]^{1/2} d^2 r$$

- Subject to p-facility constraint :

$$\int_A [s(\mathbf{r})]^{-1} d^2 r = p$$

- Lagrange multiplier optimization:

$$\frac{\delta}{\delta s(\mathbf{r})} \left[g \int_A \rho(\mathbf{r})[s(\mathbf{r})]^{1/2} d^2 r - \alpha \left(p - \int_A [s(\mathbf{r})]^{-1} d^2 r \right) \right] = 0$$

$$\Rightarrow s(\mathbf{r}) = [2\alpha / g\rho(\mathbf{r})]^{2/3}, [2\alpha/g]^{2/3} = \frac{\int_A [\rho(\mathbf{r})]^{2/3} d^2 r}{p}$$

Optimal facility location

Simulated annealing solution

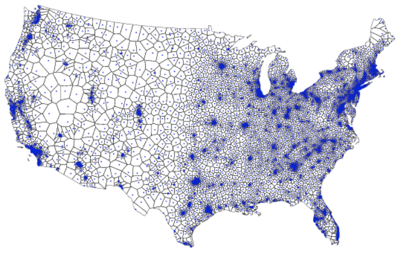


FIG. 1: Facility locations determined by simulated annealing and the corresponding Voronoi tessellation for $p = 5000$ facilities located in the lower 48 United States, based on population data from the US Census for the year 2000.

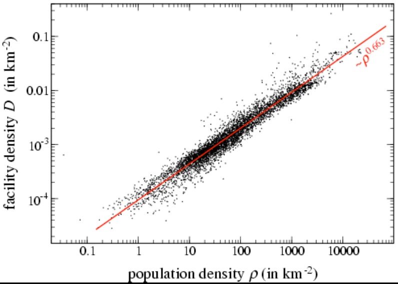


FIG. 2: Facility density D from Fig. 1 versus population density ρ on a log-log plot. A least-squares linear fit to the data gives a slope of 0.663 ± 0.002 (solid line).

Aside on simulated annealing

Kirkpatrick et al, "Optimization by Simulated Annealing". *Science* 1983

- Large search space with local minima
- 'Steepest descent' methods get stuck in local minima
- Solution: allow 'energetically unfavorable' moves provided a 'temperature' is high enough
- Cooling schedule and 'neighbor selection'

Population-equalizing cartograms

- Population density within each cell equalized (by diffusion process)
- Verification of 2/3 relation: minimizes variation of cell area

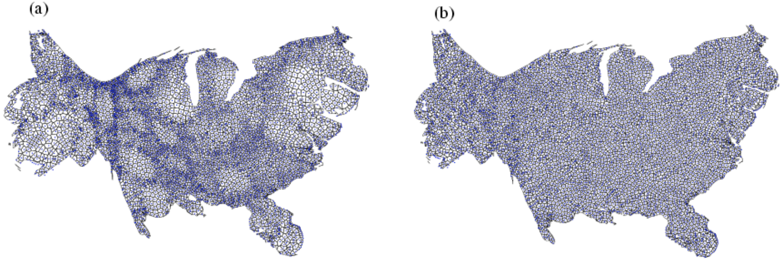
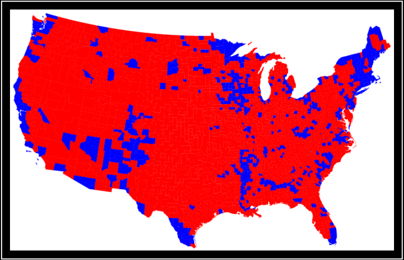
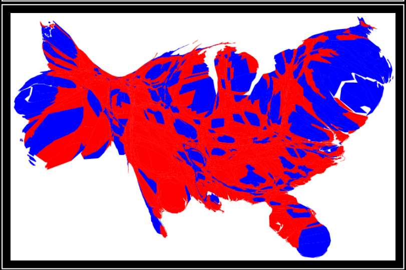


FIG. 3: Near-optimal facility location on (a) a cartogram equalizing the population density ρ and (b) a cartogram equalizing $\rho^{2/3}$.

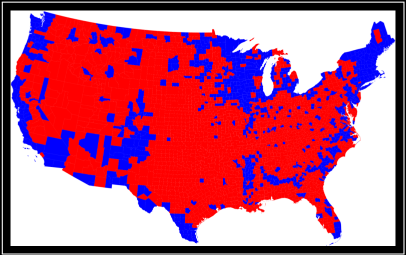
Aside on cartogram applications

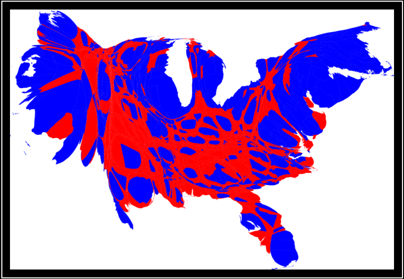
2004 U.S. Presidential Election





2008 U.S. Presidential Election





all from www-personal.umich.edu/~mejn/election

Optimal networks of facilities

- Now that we have a way to determine the 'optimal' set of facilities V, what is the optimal way to wire them?
- Use minimization of 'competing objectives' as in Fabrikant, Koutsoupias, Papadimitriou (2002) or D'Souza et al (2007)

Optimal networks of facilities

- l_{ij} is the shortest geographic distance between i and j along edges ($l_{ij} = \infty$ if no path)
- Total length of edges or network 'maintenance cost':

$$T = \sum_{i<j} A_{ij} l_{ij}$$

- 'Travel cost':

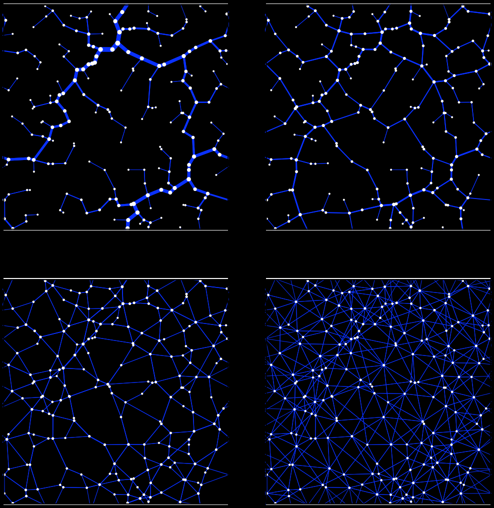
$$Z = \sum_{i<j} w_{ij} l_{ij}$$

where

$$w_{ij} = \int_{V_i} \rho(\mathbf{r}) d^2 r \int_{V_j} \rho(\mathbf{r}') d^2 r'$$

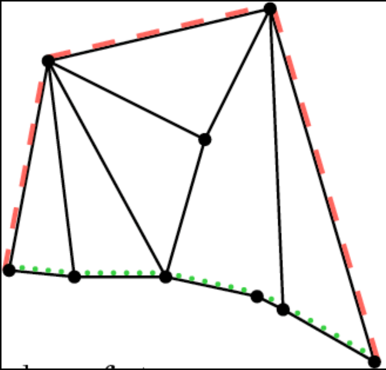
Optimal networks of facilities

- Minimize 'total cost' $T + \gamma Z$ for parameter $\gamma \geq 0$
- $\gamma = .0002, .002, .02, .2$

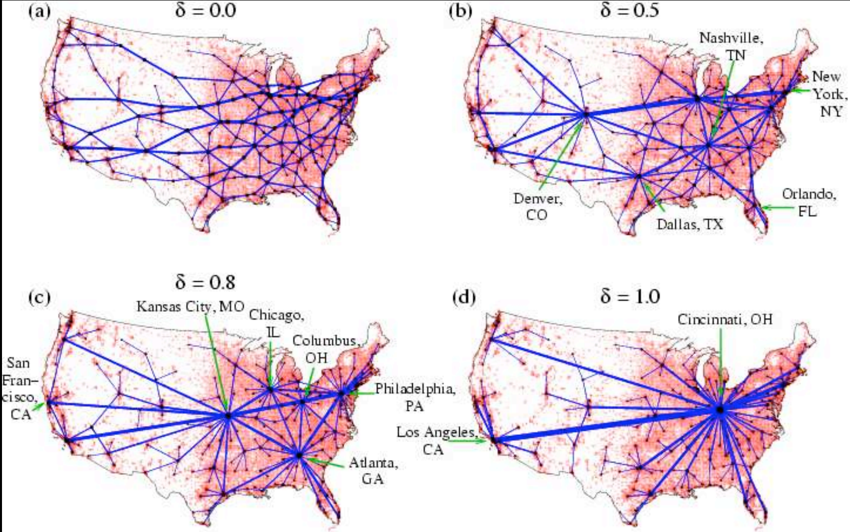


How do we measure 'distance'?

- Geographic distance (km) vs. graph distance (hopcount)
- Rescale edge length: $\widetilde{l}_{ij} = (1 - \delta) l_{ij} + \delta$ for $0 \leq \delta \leq 1$
- Modify travel cost term: $Z = \sum_{i<j} w_{ij} \widetilde{l}_{ij}$
- (actually that $\widetilde{l}_{ij}$ is the shortest rescaled *path* between i and j)



Optimal networks of facilities



Optimal networks of facilities

What's missing (and/or what's wrong)?

Optimal networks of facilities

What's missing (and/or what's wrong)?

- Pictures and analysis

Optimal networks of facilities

What's missing (and/or what's wrong)?

- Pictures and analysis
- Scaling of γ: want $O(T) = O(\gamma Z)$?

Optimal networks of facilities

What's missing (and/or what's wrong)?

- Pictures and analysis
- Scaling of γ: want $O(T) = O(\gamma Z)$?
- Congestion effects: depend on δ

Optimal networks of facilities

What's missing (and/or what's wrong)?

- Pictures and analysis
- Scaling of γ: want $O(T) = O(\gamma Z)$?
- Congestion effects: depend on δ
- Minimize whose costs: operators, users, system?

Optimal networks of facilities

What's missing (and/or what's wrong)?

- Pictures and analysis
- Scaling of γ: want $O(T) = O(\gamma Z)$?
- Congestion effects: depend on δ
- Minimize whose costs: operators, users, system?
- How realistic is cost minimization?

Optimal networks of facilities

What's missing (and/or what's wrong)?

- Pictures and analysis
- Scaling of γ: want $O(T) = O(\gamma Z)$?
- Congestion effects: depend on δ?
- Minimize whose costs: operators, users, system?
- How realistic is cost minimization?

PoA in transportation nets (Youn et al 2008)

How do we allocate flows f_{ij} on a (fixed) network?

- Social optimum: cheapest to the system

$$\min \sum_{u \in \text{users}} C_u(f_{ij})$$

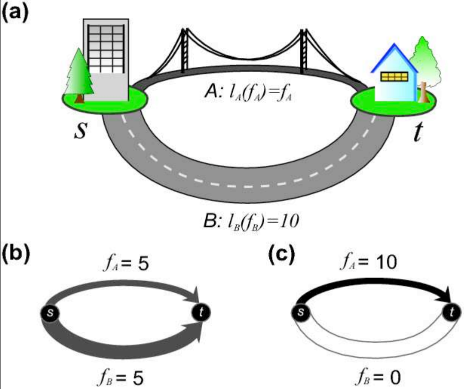
- Nash equilibrium: strategy which can't be improved unilaterally. The u-th user chooses a path taking f_{ij} as fixed:

$$\min C_u(f_{ij})$$

- Other users' decisions affect my C_u

Price of anarchy

Let l_{ij} be the 'delay' along link ij, f_{ij} the flow.



$$f_A^{SO} = f_B^{SO} = 5 \qquad\qquad f_A^{NE} = 10,\ f_B^{NE} = 0$$

$$C^{SO} = 75 \qquad\qquad C^{NE} = 100$$

Price of anarchy

Price of anarchy (PoA) measures the 'inefficiency of decentralization':

$$\text{PoA} = \frac{\sum l_{ij}(f_{ij}^{NE}) \cdot f_{ij}^{NE}}{\sum l_{ij}(f_{ij}^{SO}) \cdot f_{ij}^{SO}}$$

Previous example: $\text{PoA} = 4/3$

This is system-view; some users are worse off than others in SO

Braess's Paradox

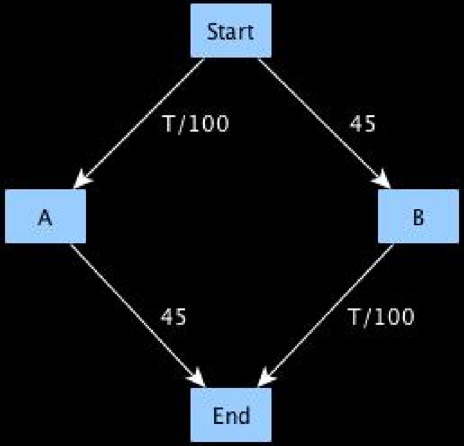
In a user/decentralized optimum, adding additional links may actually make everybody worse off.

Anecdotal examples (yep, from wikipedia.org/wiki/Braess's paradox)

- Seoul, South Korea: a speeding-up in traffic around the city was seen when a motorway was removed
- Stuttgart, Germany, 1969: after investments into the road network, the traffic situation did not improve until a section of newly-built road was closed
- New York City, 1990: closing of 42nd street in New York City reduced the amount of congestion in the area.

Braess's Paradox

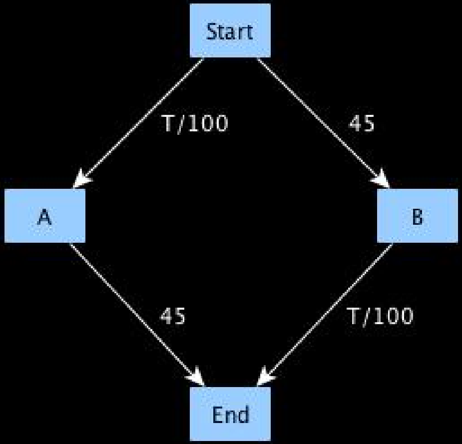
Consider a simple 4-node network.



Braess's Paradox

Consider a simple 4-node network.

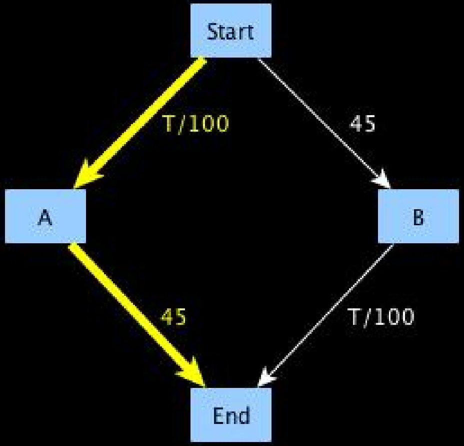
- 4000 drivers want to go from start to end
- Two links depend on flow (T)



Braess's Paradox

Time needed to send A drivers along A-route:

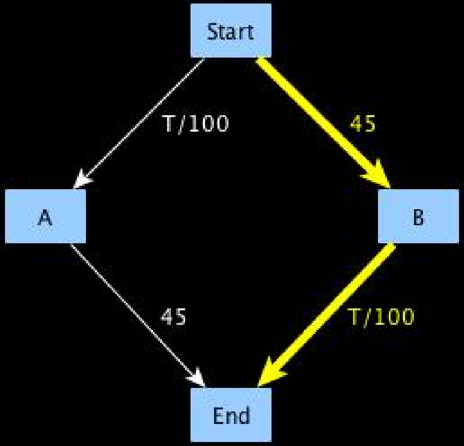
$$l_A = \frac{A}{100} + 45$$



Braess's Paradox

Time needed to send B drivers along B-route:

$$l_B = 45 + \frac{B}{100}$$

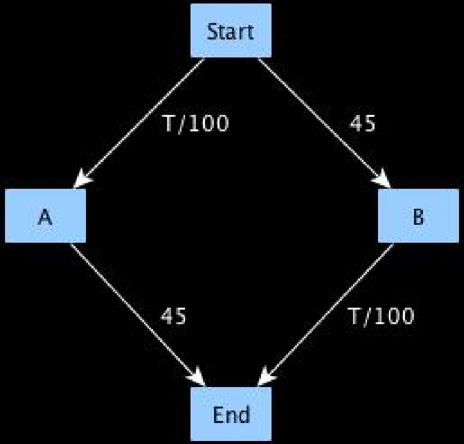


Braess's Paradox

Nash equilibrium (user-optimal) occurs when $l_A = l_B$:

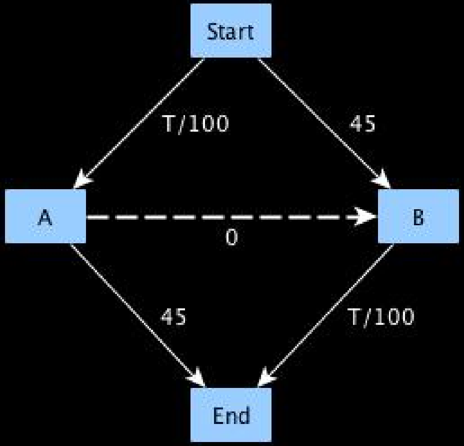
$A = B = 2000$

$l_A = l_B = 65\text{min}$



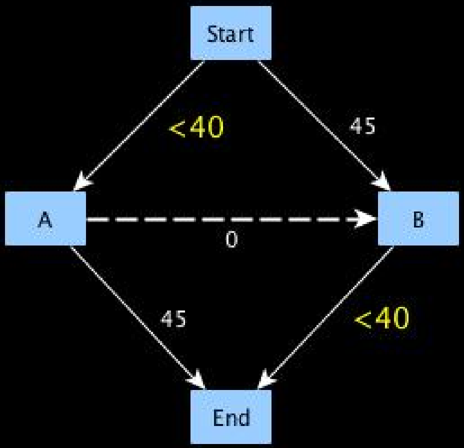
Braess's Paradox

Now add a 'free' shortcut. This can only improve things, right?



Braess's Paradox

Consider the 'worst-case' on flow-dependent links.
Always cheaper than fixed links.

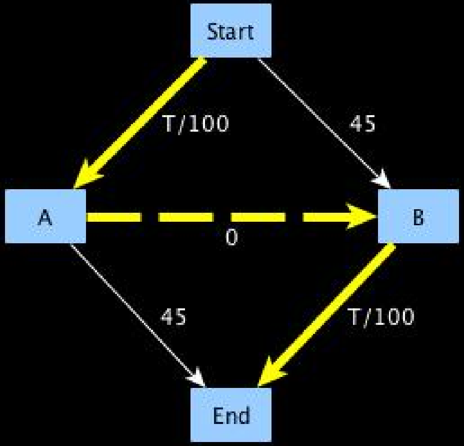


Braess's Paradox

Claim: everyone rationally chooses this route.

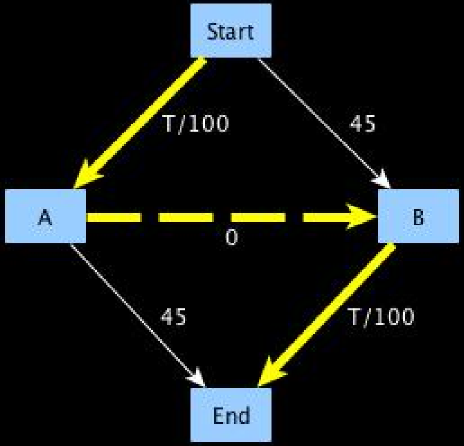
$$I_S = \frac{4000}{100} + \frac{4000}{100}$$

$$= 80 \text{ min}$$



Braess's Paradox

Why? Could I unilaterally do better?

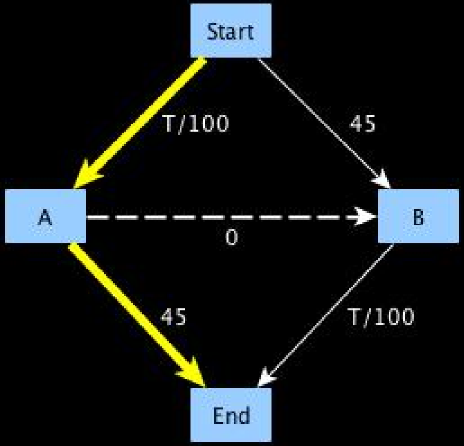


Braess's Paradox

This way costs more:

$$l_A = \frac{4000}{100} + 45$$

$$= 85 \text{ min} > 80$$

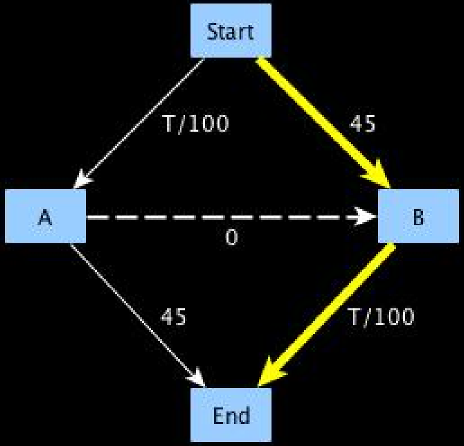


Braess's Paradox

And so does this:

$$l_B = \frac{4000}{100} + 45$$

$$= 85 \text{ min} > 80$$

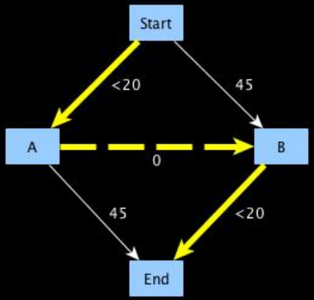


Braess's Paradox

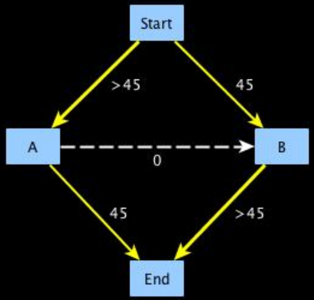
So what? Is this realistic?

- Sensitive to flows and cost functions

Link improves overall flow
$N = 2000$



Link doesn't matter
$N = 9000$



Flows in real-world road networks

- d_{ij} distance of link, $v_{ij} = 35$mph, f_{ij} flow, p_{ij} capacity
- Delays: $l_{ij} = \frac{d_{ij}}{v_{ij}} \left[1 + \alpha \left(\frac{f_{ij}}{p_{ij}} \right)^{\beta} \right]$
- $\alpha = 0.2$, $\beta = 10$ large superlinearity
- Optimal flows:
  - Social optimal flows f_{ij}^{SO} minimize

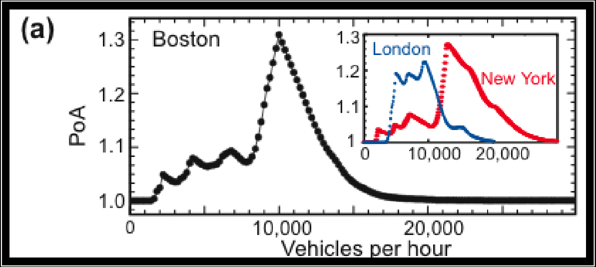
$$\sum_{\text{link}(i,j)} l_{ij}(f_{ij}) f_{ij}$$

  - Nash equilibrium flows f_{ij}^{NE} minimize

$$\sum_{\text{link}(i,j)} \int_0^{f_{ij}} l_{ij}(f') df'$$

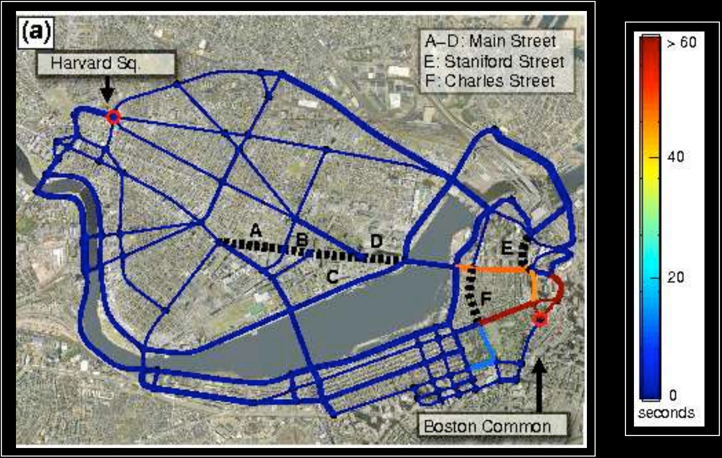
PoA in real-world road networks

- Send flow between one S-D pair
- PoA depends on total flow volume:

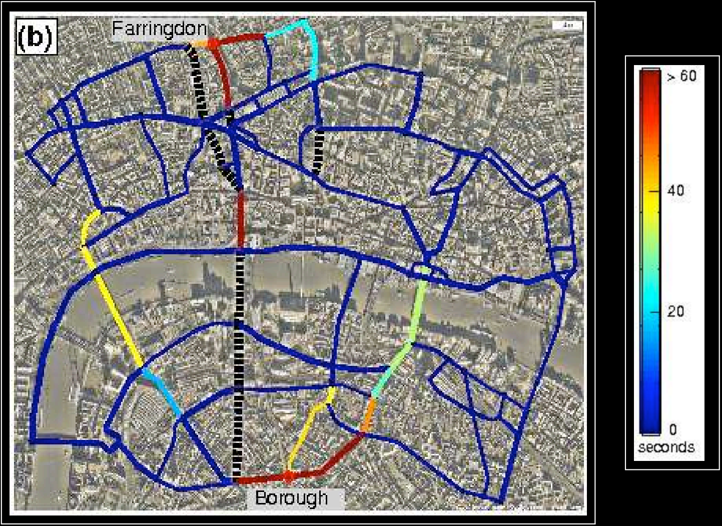


- Qualitatively similar results for different source, destination pairs
- (N, E): Boston: $(88, 246)$ London: $(82, 217)$ NYC: $(125, 319)$

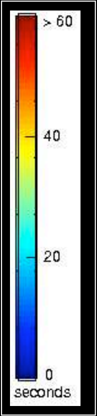
Braess's Paradox: Boston



Braess's Paradox: London



Braess's Paradox: NYC



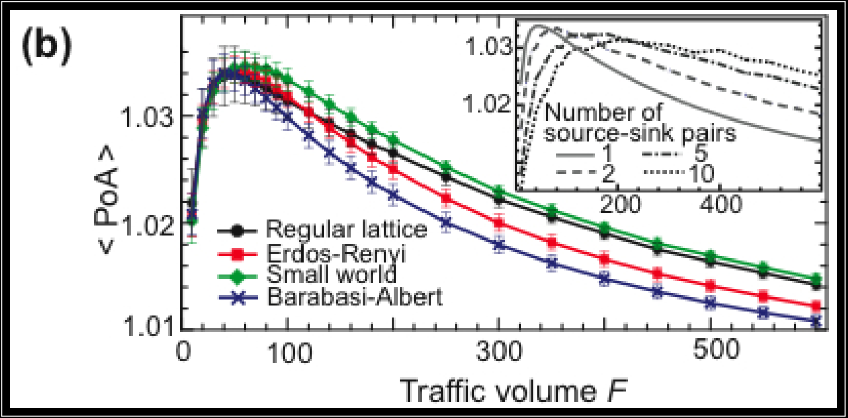
Comparison with random and regular models

Models with $N = 100$, $< k >= 6$

- 1-D lattice w/ links to 3rd nearest neighbors, periodic bdys
- ER RG
- SW nets with $p = 0.1$
- BA nets

$l_{ij} = a_{ij} f_{ij} + b_{ij}$
a_{ij} random integer $\in \{1, 2, 3\}$
b_{ij} random integer $\in \{1 \ldots 100\}$

Comparison with random and regular models



SO optimizes: $C = \sum (a_{ij} f_{ij}^2 + b_{ij} f_{ij})$

NE optimizes: $\widetilde{C} = \sum (\frac{1}{2} a_{ij} f_{ij}^2 + b_{ij} f_{ij})$

PoA $\rightarrow 1$ when $f_{ij} \rightarrow 0$ or $f_{ij} \rightarrow \infty$

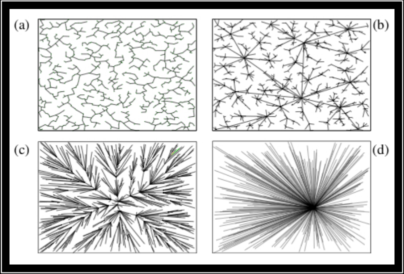
What's missing, what's wrong, what's next?

- Are there more frequent 'paradox'-links?
- What does the PoA of 'multicommodity' (many source, many destination) flow look like?
- Why is there such a 'striking' similarity between RG models? (just a function of $< k >$?)

Optimal traffic networks (Barthélemy and Flammini 2006)

- Arbitrary allocation of traffic t_e to minimize $w_e = d_e/t_e$ (s.t. connectivity)?
- Stochastic optimization scheme (metropolis vs. simulated annealing)
- If b_e is edge-betweenness, d_e is length of e, minimize

$$E_{\mu\nu} = \sum_{e \in T} b_e^{\mu} d_e^{\nu}$$



Multilayered Spatial Network Optimization

- Infrastructure networks are composed of multiple entities: independent airline carriers, utility transmission vs. distribution companies
- Optimization problems solved by each layer/entity need not be identical, e.g.:
  - Southwest ('point-to-point') vs. Traditional carriers ('hub-and-spoke')
  - High-capacity transmission layer seeks a robust structure; distribution layer simply minimizes cost to end-users

Why these papers {a OR we}re worth your time

- Infrastructure networks' design emerges from simple optimization
- Flows on networks: network structure/function determines user preferences which in turn have nontrivial effects on network function
- (Cool pictures)