Optimal design of spatial distribution networks

Michael T. Gastner, Santa Fe Institute
collaborator: M.E.J. Newman, University of Michigan

The problem

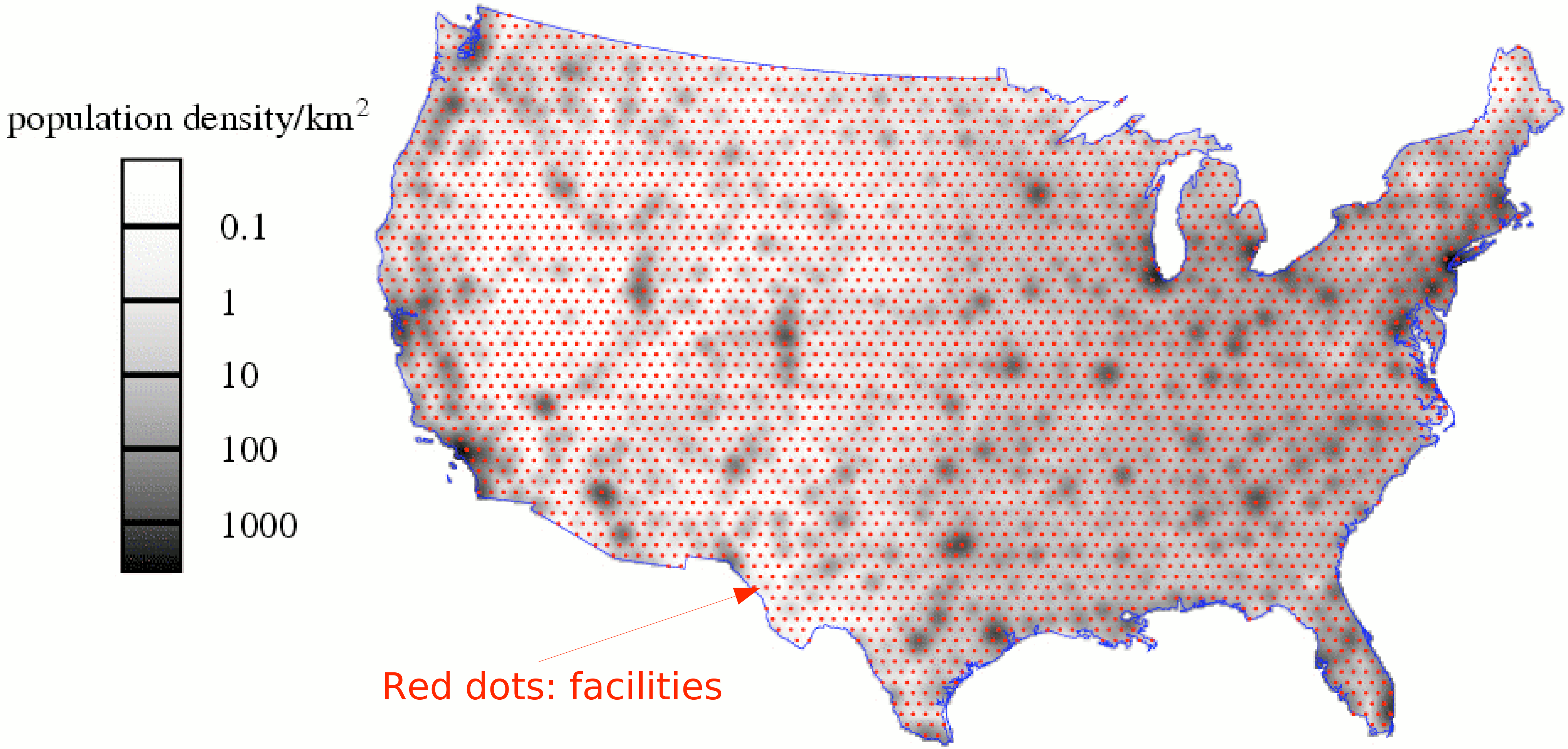
- Suppose we are charged with choosing the sites of p facilities (hospitals, post offices, supermarkets etc.) within a country or province.

- Furthermore, suppose we are given the population density $\rho(\mathbf{r})$ in this region.

- Where do we place the facilities such that the average distance from a person's home to the nearest facility is minimized?

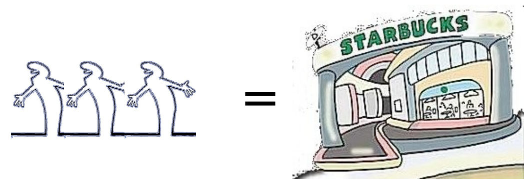
- Often the facilities are interconnected to form networks (e.g. airports, warehouses). How do we optimally connect the facilities to optimize performance of the system as a whole?

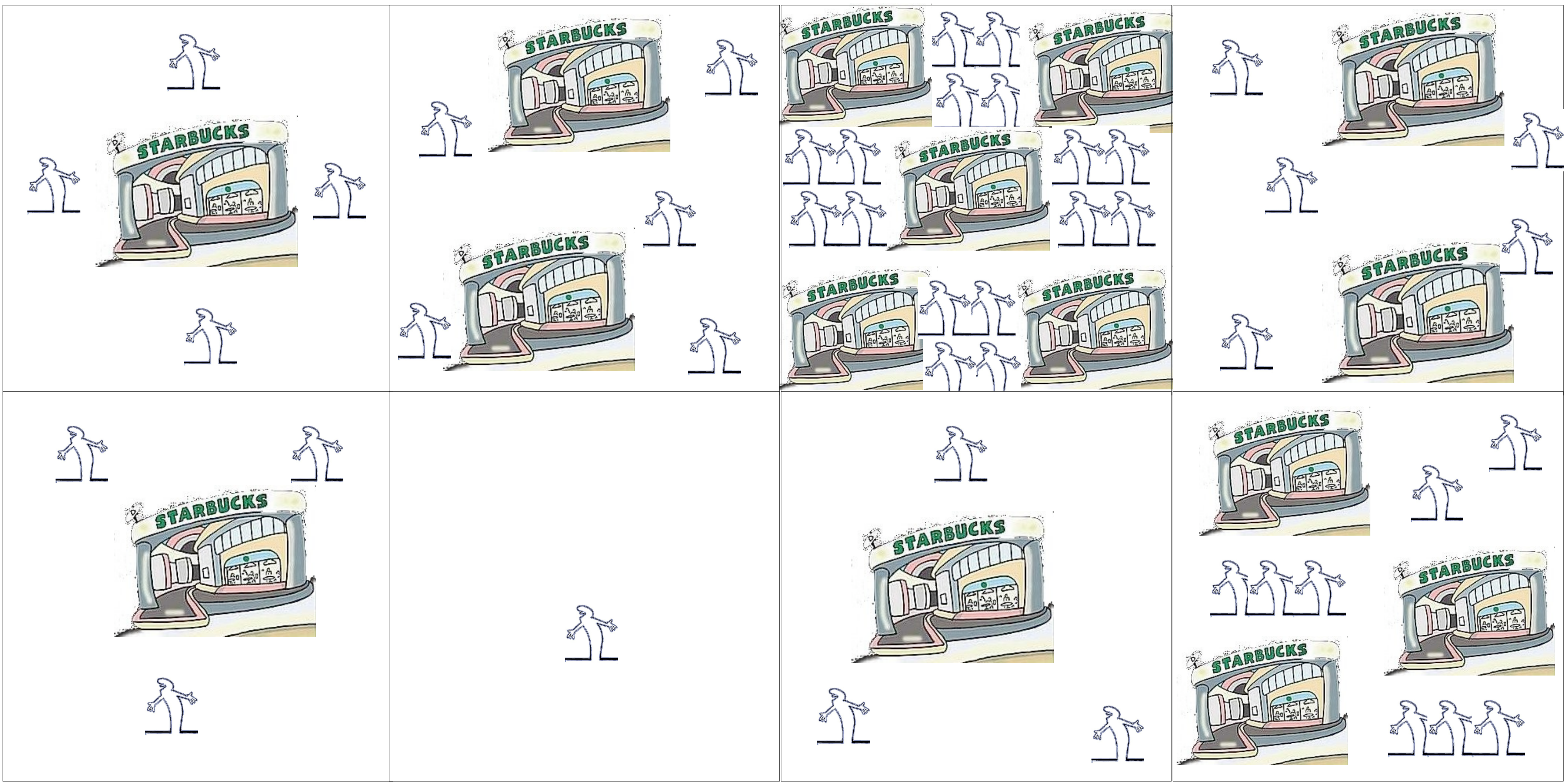
Why is it difficult to locate the facilities?



In most countries, population density is highly non-uniform. $\longrightarrow$
A uniform distribution of facilities would be a poor choice:
it gains us little to build a lot of facilities in sparsely populated areas.

An apparently more sensible choice:
distribute facilities in proportion to population density. $\longrightarrow$
A region with twice as many people has twice as many facilities.

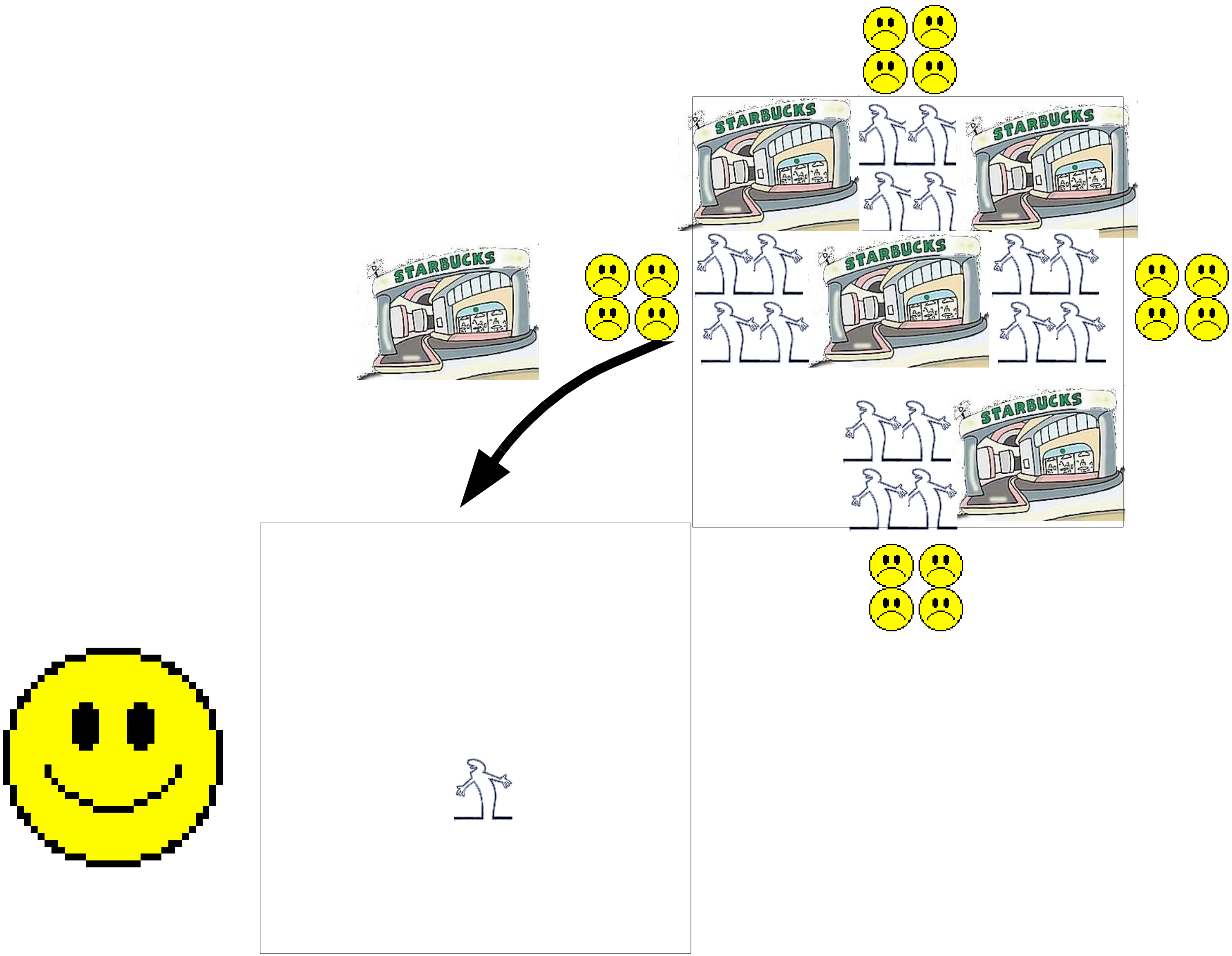




Do we gain anything by having closely spaced facilities in the highly populated areas?

There the second-closest facility is not much farther than the closest.

⟶ One of them might be removed with little penalty?



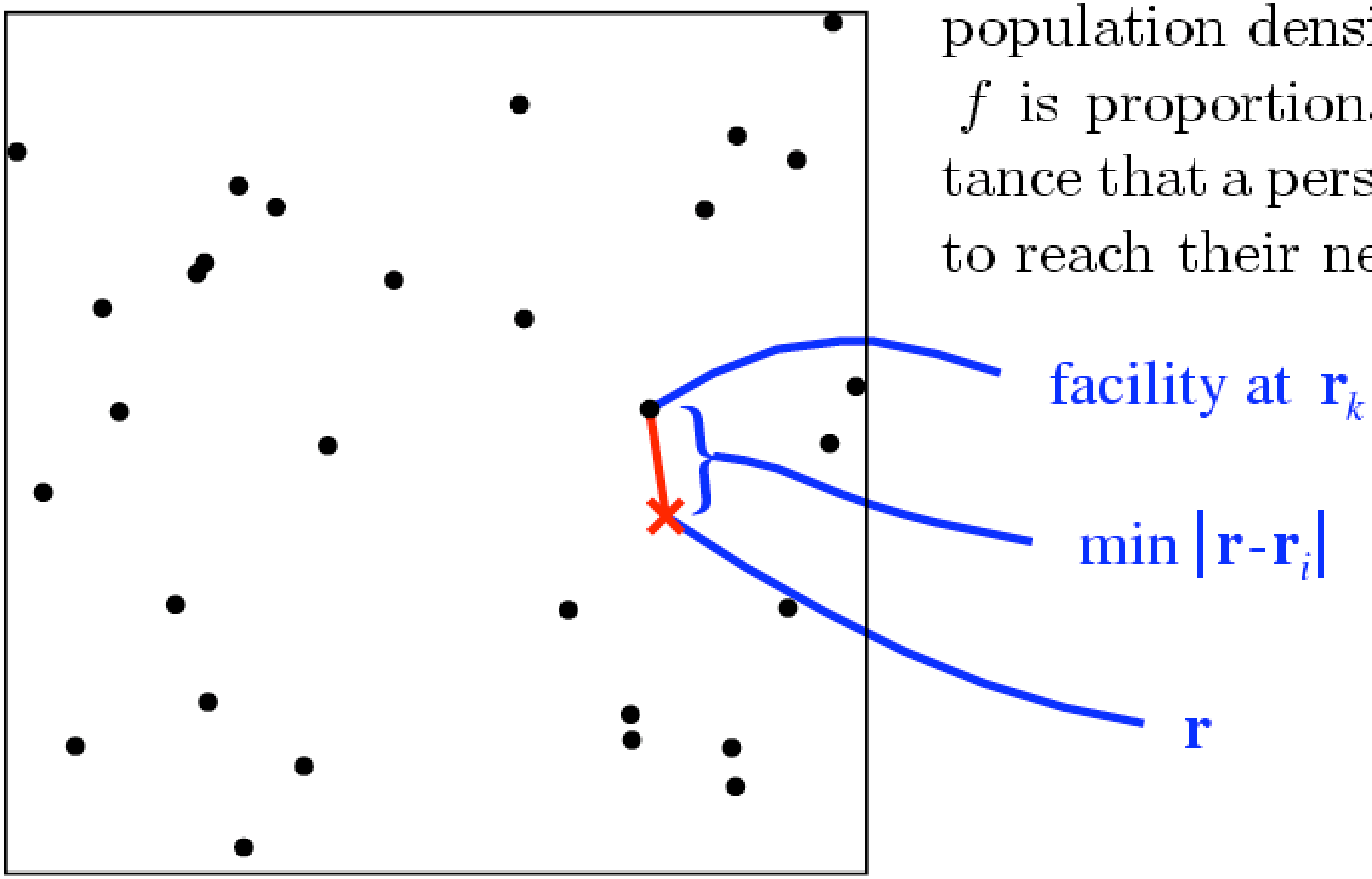
The objective function f

We wish to distribute p facilities over an area A such that

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

is minimized. Here $\{\mathbf{r}_1, \ldots, \mathbf{r}_p\}$ are the facility positions and $\rho(\mathbf{r})$ the population density.

f is proportional to the mean distance that a person will have to travel to reach their nearest facility.

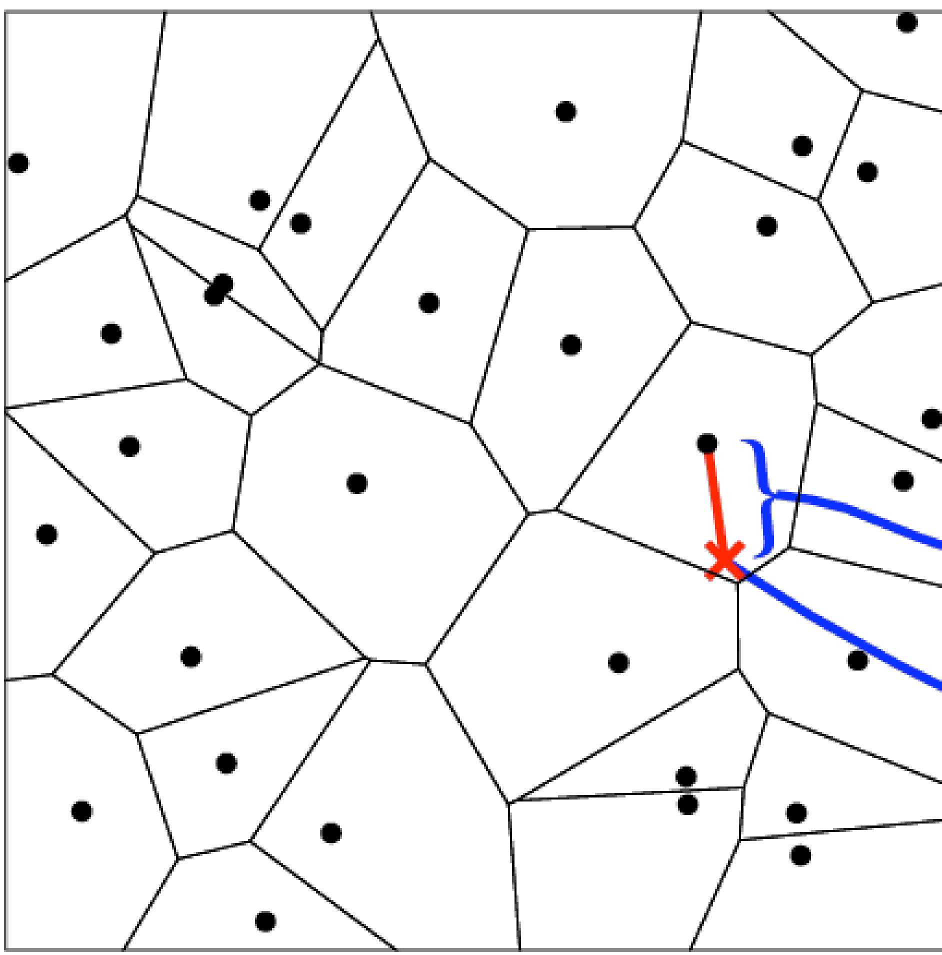


Computational complexity

Finding the exact minimum of

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

is known as the *p-median problem*. It has been shown to be NP-hard, so in practice we must rely on approximate numerical optimization or approximate analytical treatments.



One way to obtain an analytical result is by analyzing the *Voronoi tessellation*.

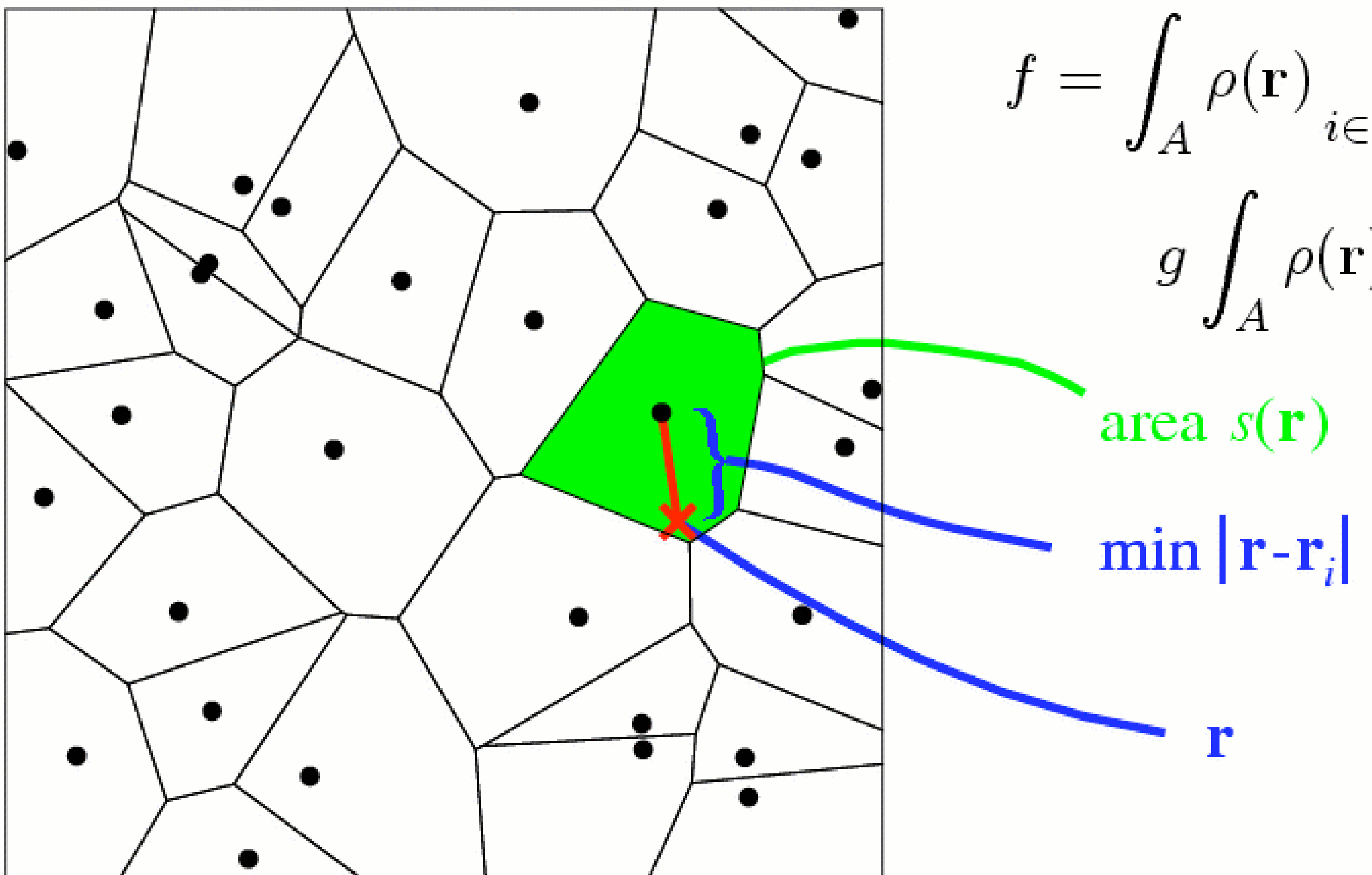
Scaling analysis

Define $s(\mathbf{r})$ = the area of the Voronoi cell to which the point $\mathbf{r}$ belongs.

$\longrightarrow$ In two dimensions:

Distance from a point $\mathbf{r}$ to the nearest facility $= g[s(\mathbf{r})]^{1/2}$,

where g is a geometric factor of order 1.



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

$$g \int_A \rho(\mathbf{r}) [s(\mathbf{r})]^{1/2} \, \mathrm{d}^2 r.$$

Scaling analysis

Since there are p facilities:
$$\int_A [s(\mathbf{r})]^{-1}\,\mathrm{d}^2 r = p. \tag{1}$$

Optimizing f subject to this constraint gives

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

where α is a Lagrange multiplier.

$$\rightarrow \quad s(\mathbf{r}) = [2\alpha/g\rho(\mathbf{r})]^{2/3}.$$

α can be evaluated by substituting into Eq. (1):

$$D(\mathbf{r}) = \frac{1}{s(\mathbf{r})} = p\frac{[\rho(\mathbf{r})]^{2/3}}{\int [\rho(\mathbf{r})]^{2/3}\,\mathrm{d}^2 r},$$

where $D(\mathbf{r}) = [s(\mathbf{r})]^{-1}$ is the density of the facilities.

What does this mean?

$$D(\mathbf{r}) \propto \rho(\mathbf{r})^{2/3} \tag{1}$$

If facilities are distributed optimally, their density should increase with population density, but it should do so slower than linearly, as a power law with exponent $\frac{2}{3}$.

Equation (1) is a compromise between

- a population-proportional allocation, $D(\mathbf{r}) \propto \rho(\mathbf{r})^1$, (2) and
- a spatially homogeneous distribution, $D(\mathbf{r}) \propto \rho(\mathbf{r})^0$. (3)

(1) — similar to (2) — places most facilities in the densely populated areas where most people live.
But (1) — similar to (3) — still proves reasonable service to those in sparsely populated areas.

What does this mean?

Our calculation can also be carried out in general dimension d,

$$D \propto \rho^{d/(d+1)} \propto (\text{area occupied by one person})^{-d/(d+1)}. \tag{1}$$

In plant ecology, the population density of a species, D, scales with the plant size as

$$D \propto \text{size}^{-3/4}. \tag{2}$$

Equation (2) is the result of competition among plants for spatially limited resources and optimal use of these resources (West et al., Nature, 1998).

Note that plants are three-dimensional objects, so the scaling exponent in Eq. (1) and (2) are equal.

Numerical optimization

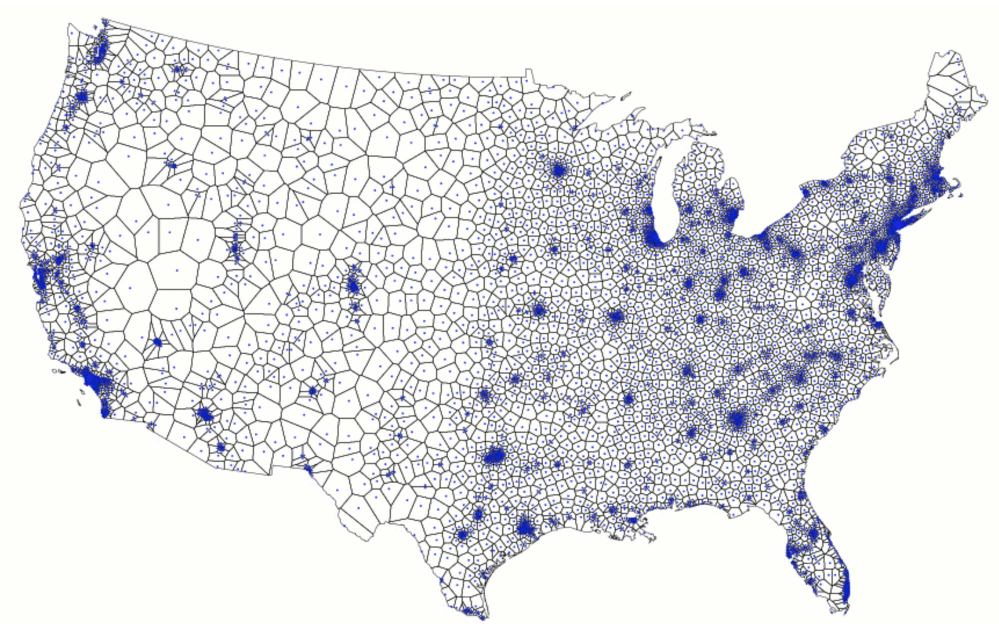


FIG. 1: Facility locations determined by simulated annealing and the corresponding Voronoi tessellation for $p = 5000$ facilities located in the lower 48 United States, based on population data from the US Census for the year 2000.

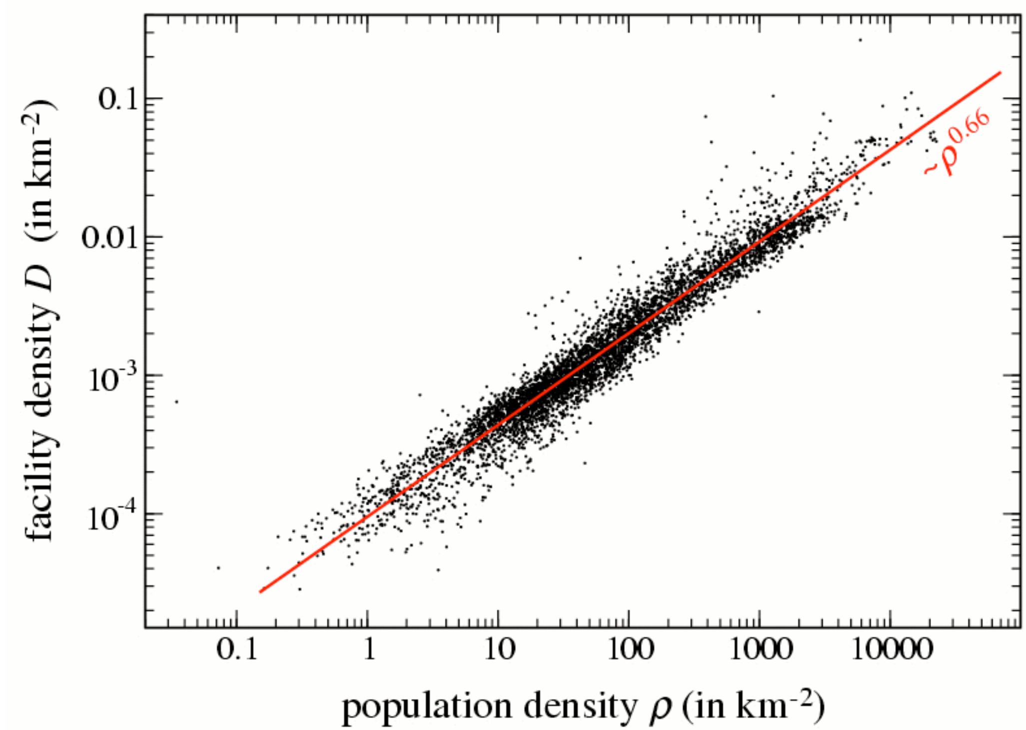
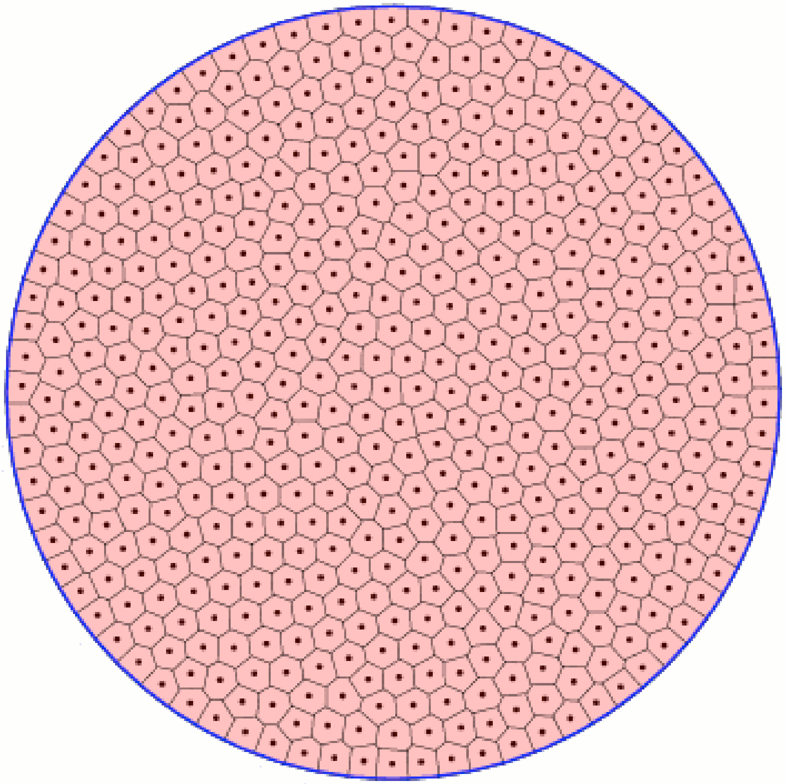


FIG. 2: Facility density D from Fig. 1 versus population density ρ on a log-log plot. A least-squares linear fit to the data gives a slope of 0.66 (solid line, $r^2 = 0.94$).

Is there a geometric solution?

If we neglect finite-size effects, optimally located facilities in a *uniformly* populated space have hexagonal catchment areas.



Is there a map projection that will transform the pattern of facilities for a *nonuniform* population to a similarly regular structure?

Cartograms

Facility locations depend on the population density. $\longrightarrow$

If we want the *facility* density to appear homogeneous, we need a projection which corrects for variations in the *population* density.

Such projections are called *cartograms*.

A cartogram rescales areas such that

- densely populated regions become larger and
- sparsely populated regions become smaller.

Cartograms

William Bunge: Patterns of Location (1964)

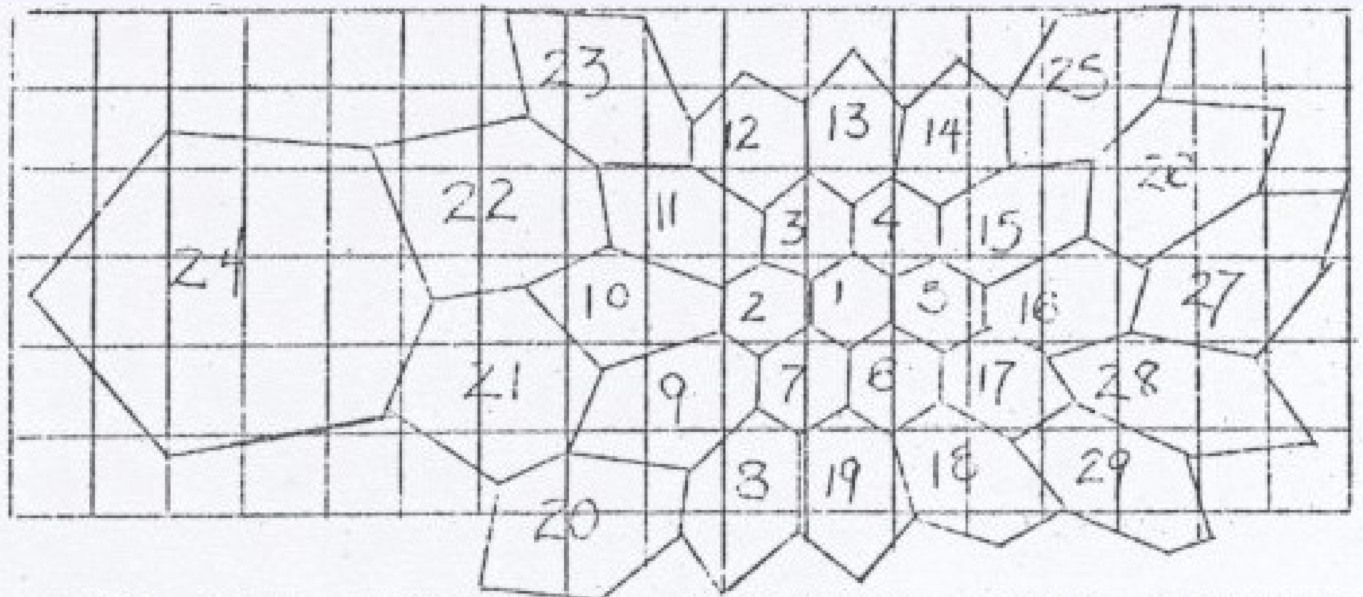


Figure 35. An approximation of a Christaller solution applied to an area of disuniform rural population. (Some errors)

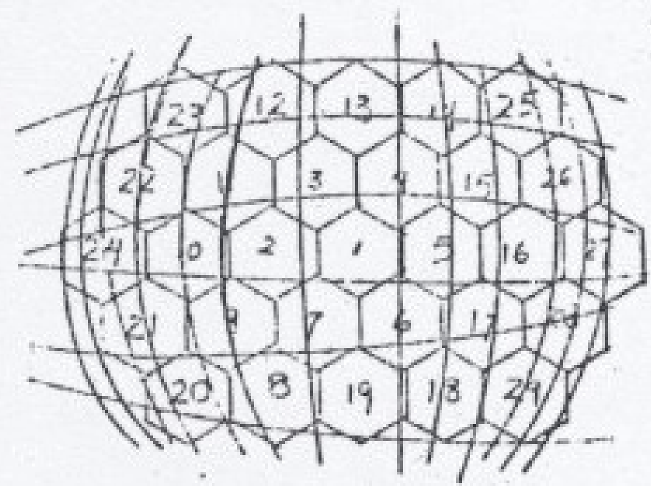
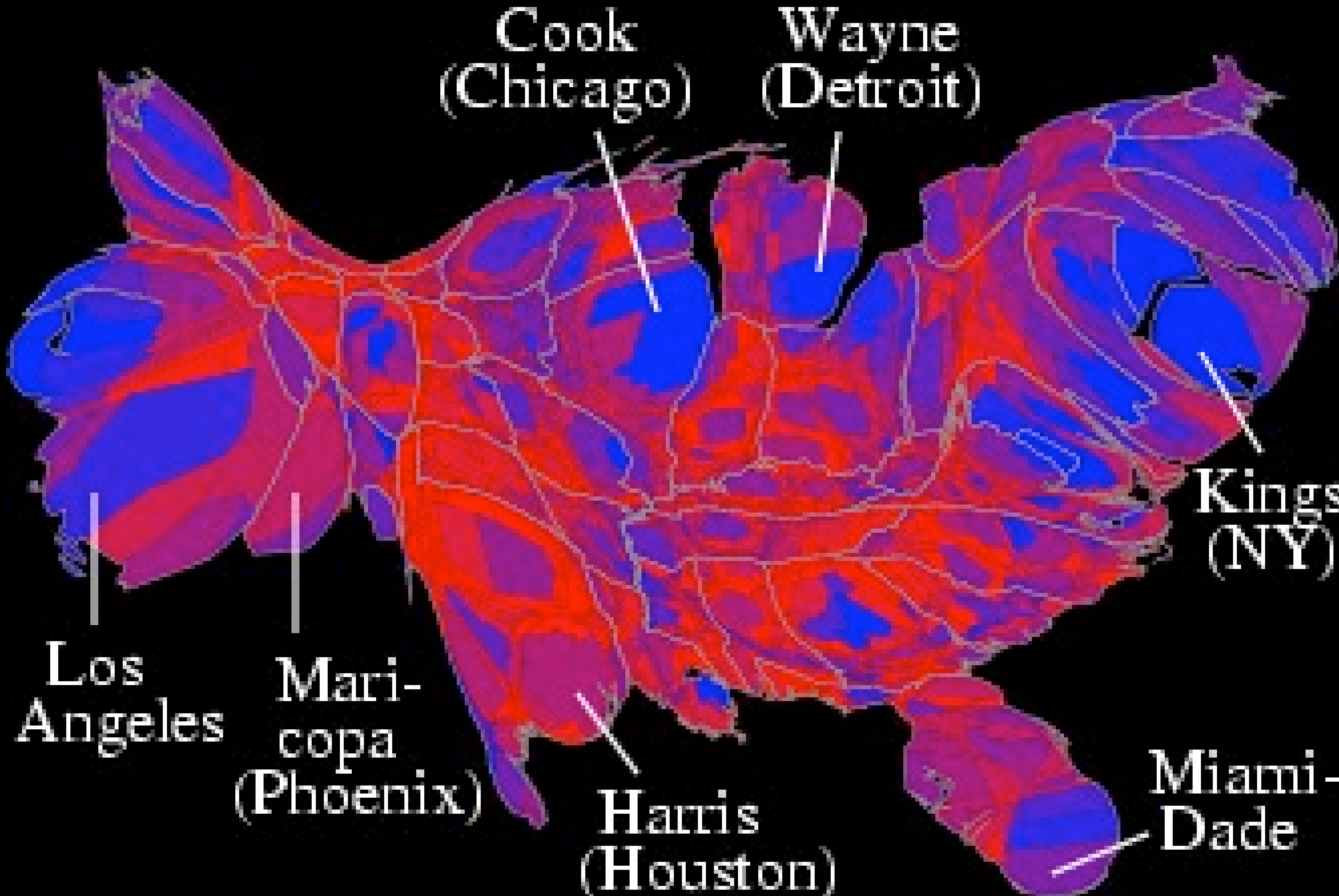
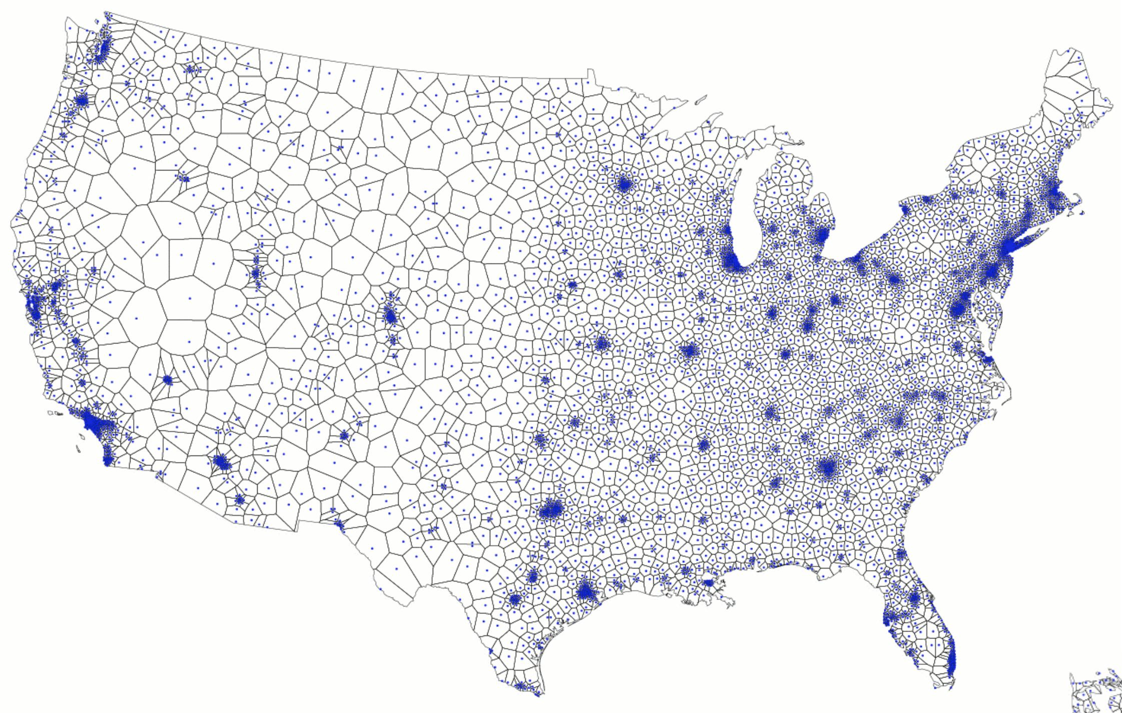


Figure 36 Map transformed into uniform density

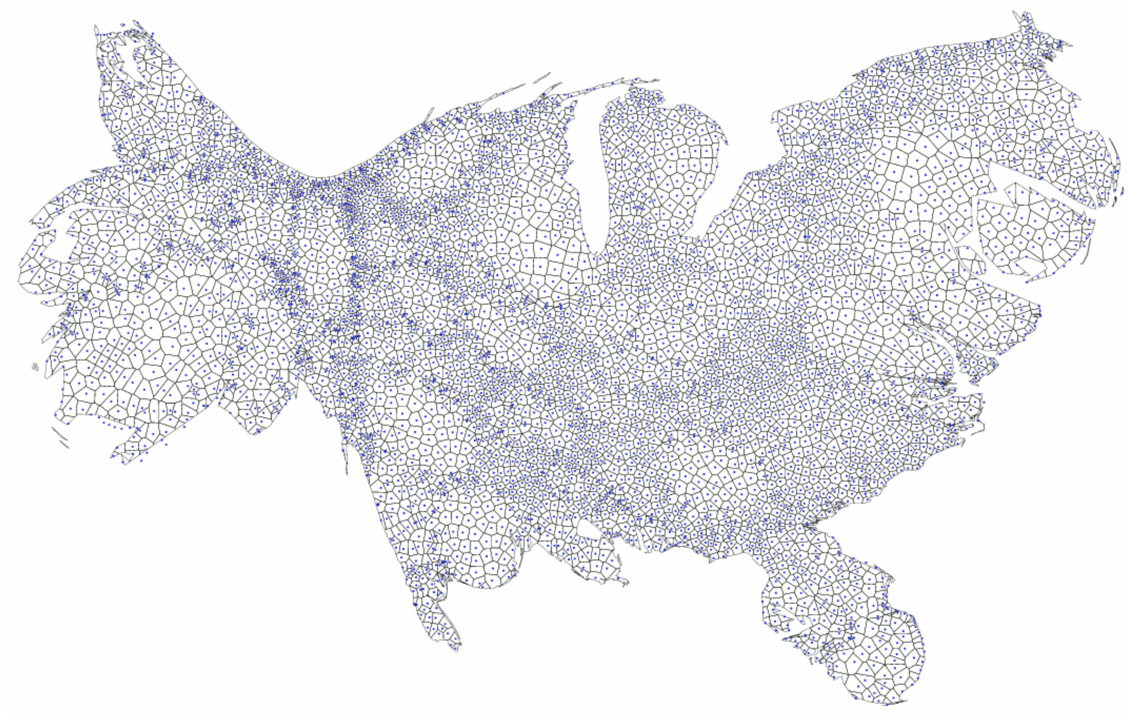

Cook
(Chicago)
Wayne
(Detroit)
Kings
(NY)
Los
Angeles
Mari-
copa
(Phoenix)
Harris
(Houston)
Miami-
Dade

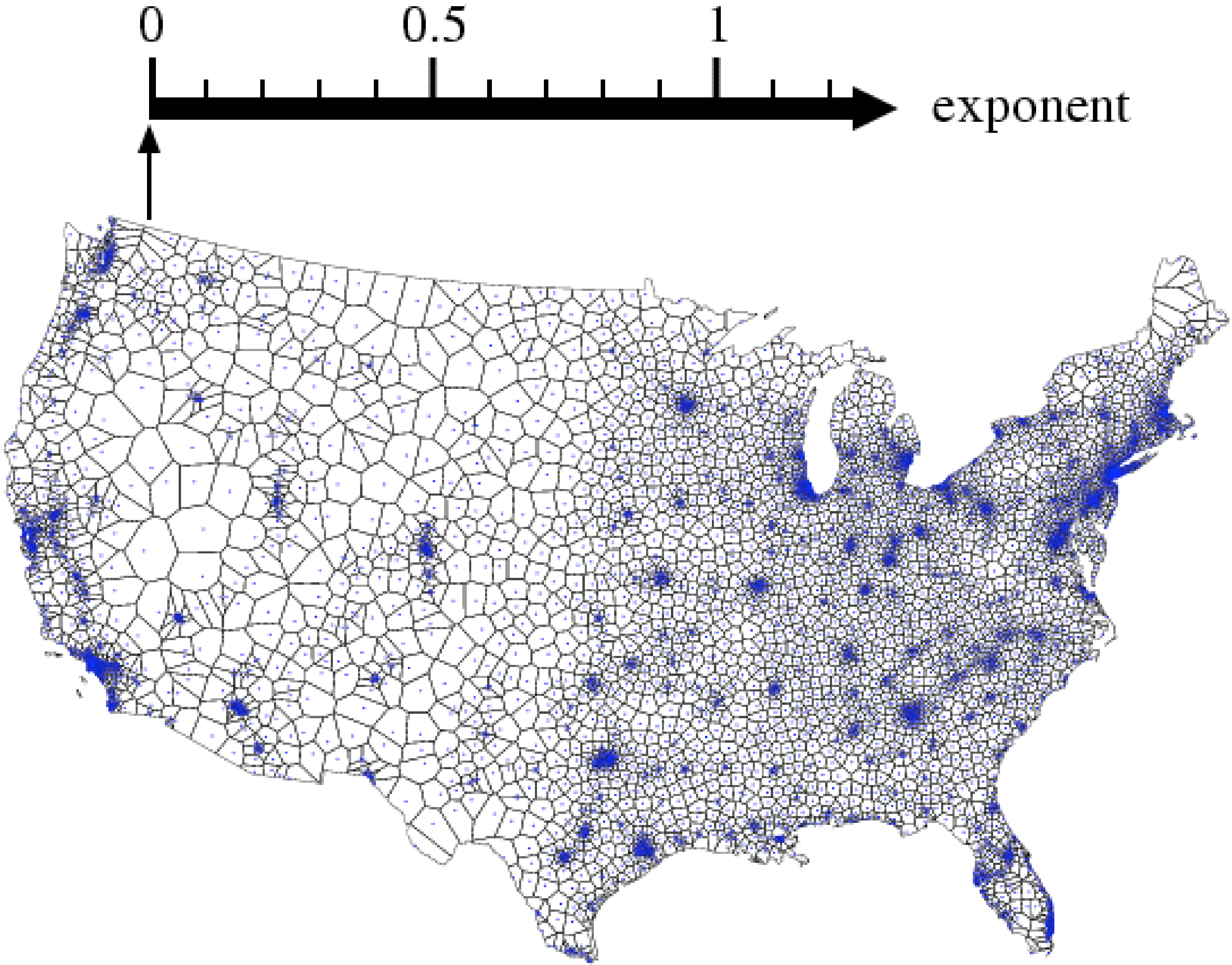
Equal-area vs. equal-population projection

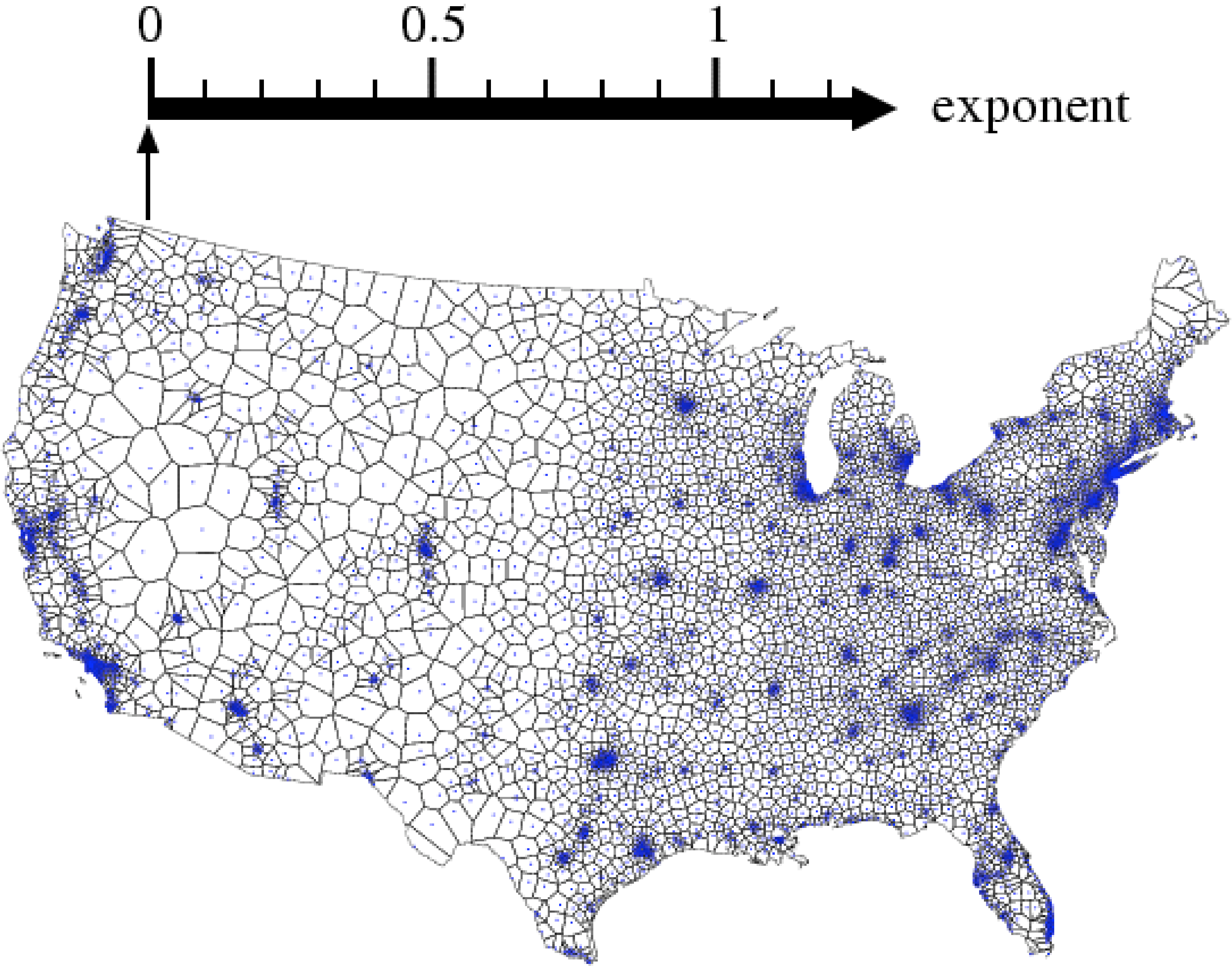


Equal-area projection

Equal-population projection

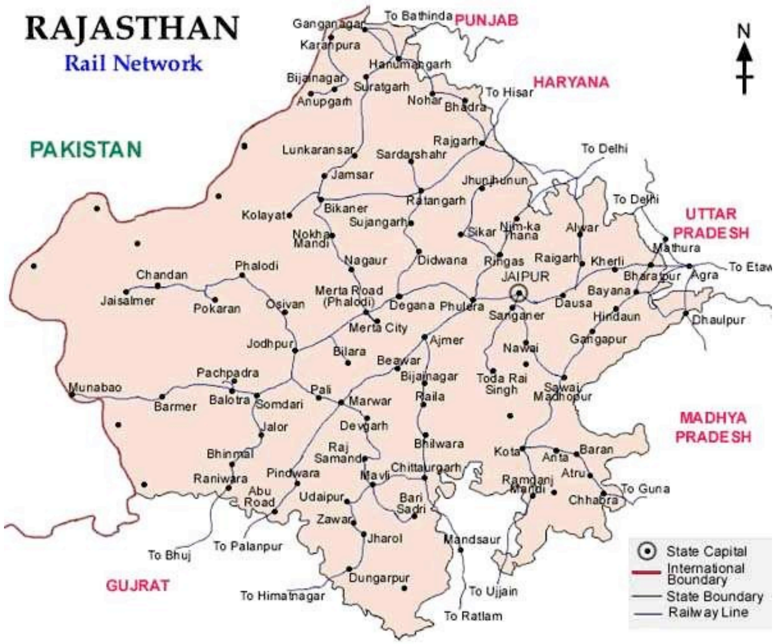


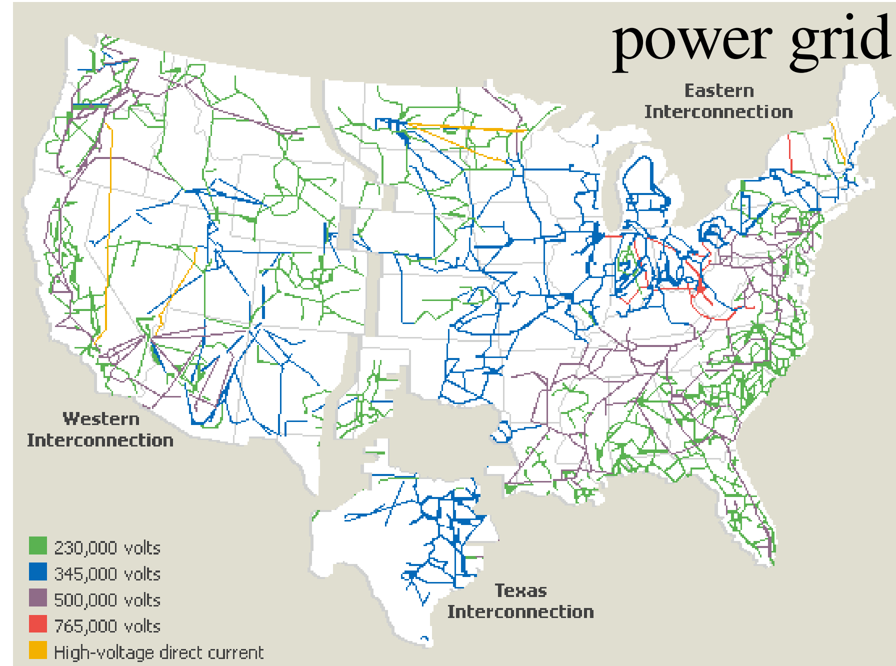

0
0.5
1
exponent

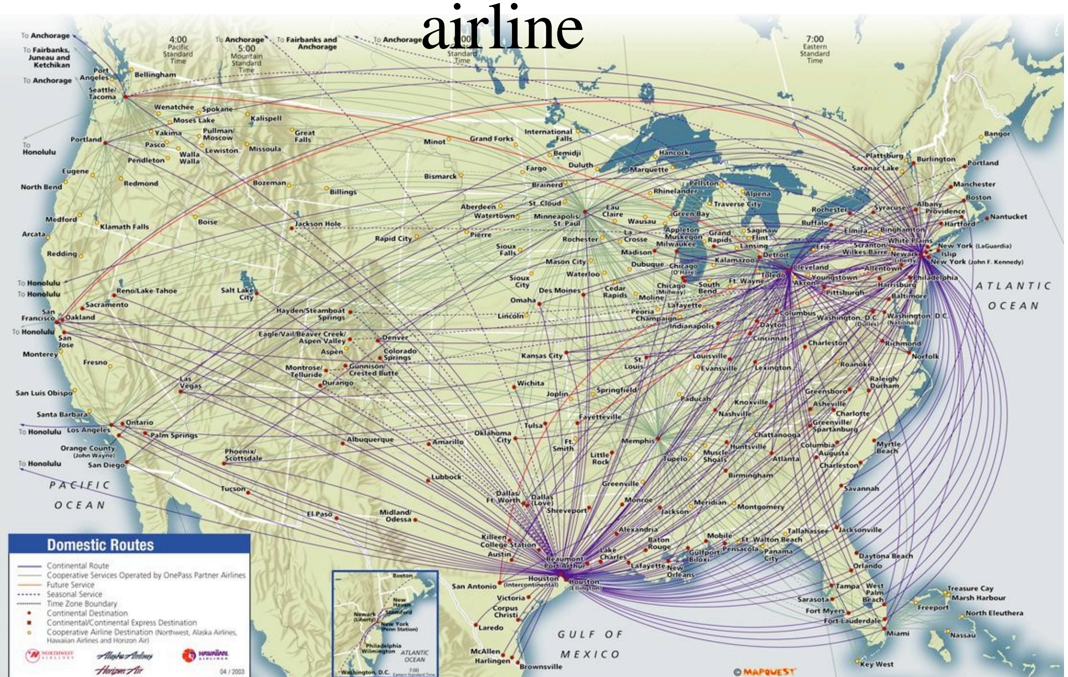

0
0.5
1
exponent

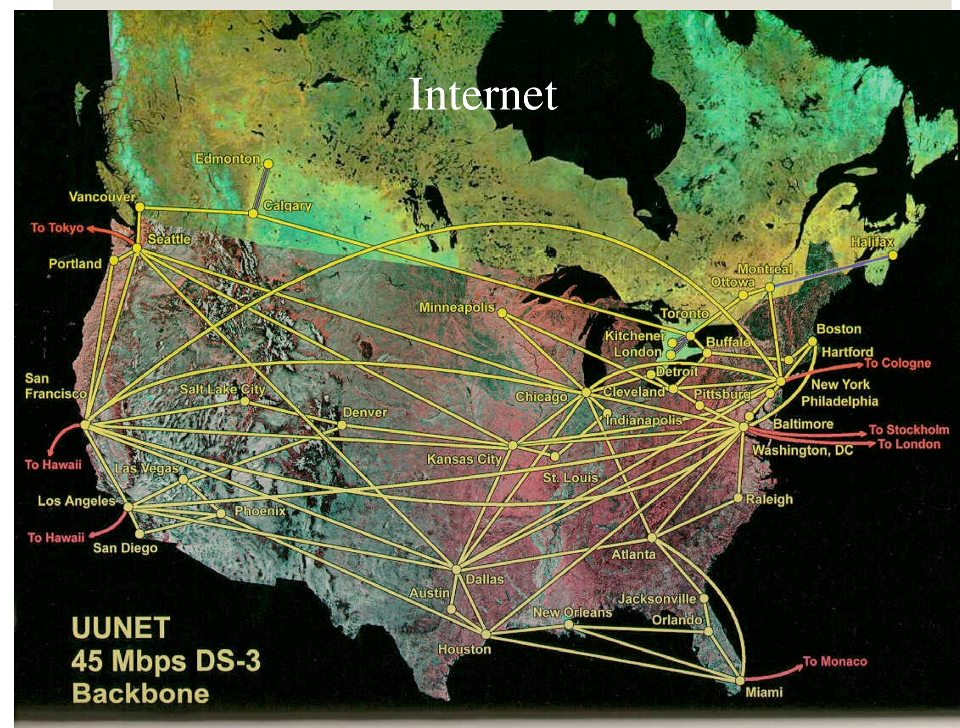
Optimal network of facilities

trains



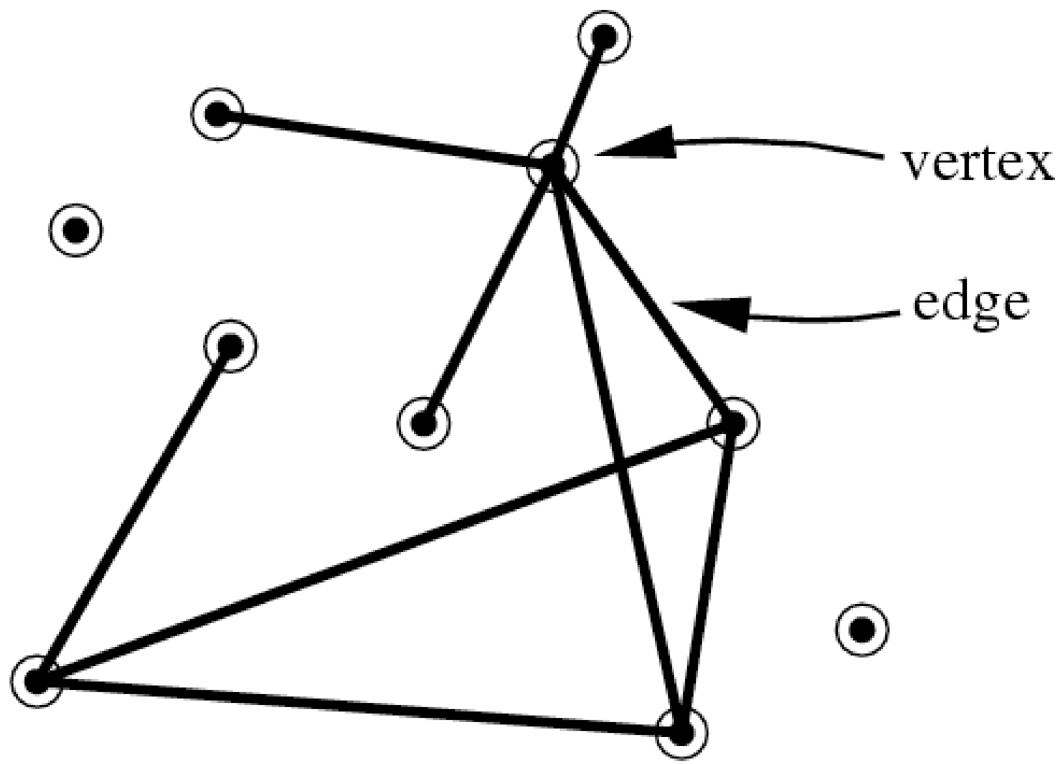






A few definitions

Consider a situation in which the facilities form the vertices of a network and connections between them form the edges.



If there are p facilities in the network, the *adjacency matrix* $\mathbf{A}$ is a $p \times p$ matrix with elements

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between facilities } i \text{ and } j \\ 0 & \text{otherwise.} \end{cases}$$

The efficiency of a network of facilities

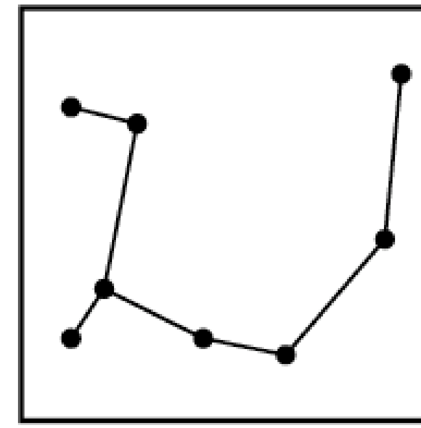
The efficiency of a network, as we will consider it here, depends on two factors.

- On the one hand, the smaller the sum of the lengths of all edges, the cheaper the network is to construct and maintain.
- On the other hand, the shorter the distances through the network between vertices, the faster the network can perform its intended function (e.g. transportation of passengers between vertices or distribution of mail or cargo).

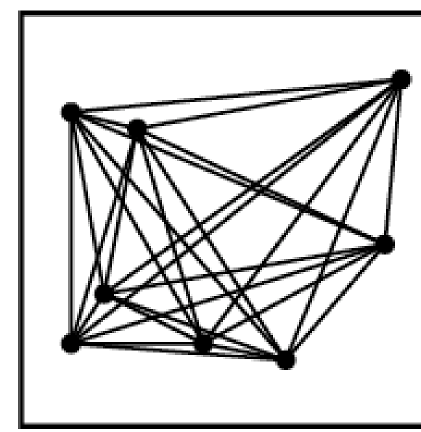
The efficiency of a network of facilities

These two objectives generally oppose each other.

- A network with few and short connections will not provide many direct links between distant points, and paths through the network will tend to be circuitous.



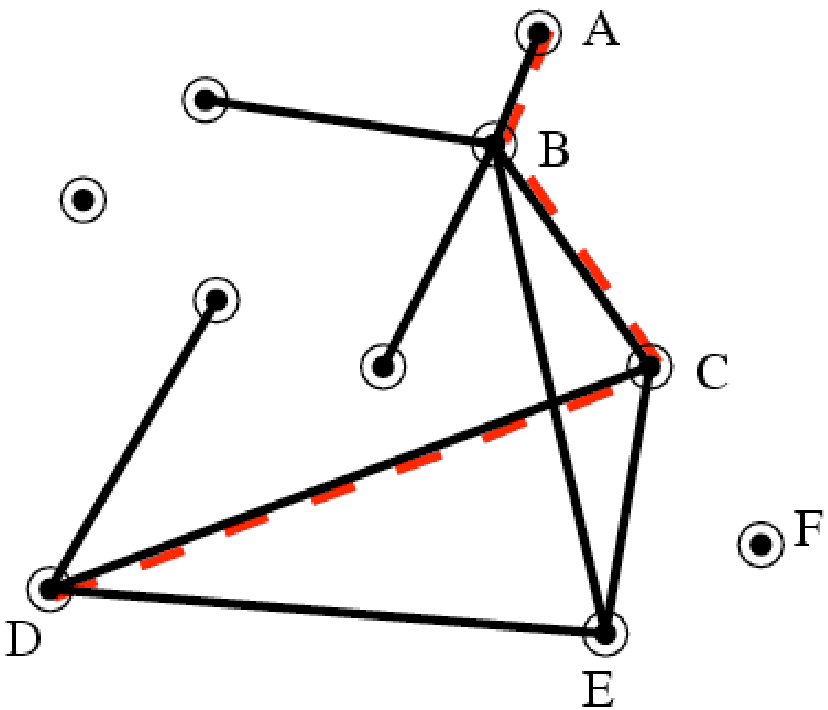
- A network with a large number of direct links is usually expensive to build and operate.



The optimal solution lies somewhere between these extremes.

Construction cost

Let us define l_{ij} to be the shortest geometric distance between two vertices i and j measured along the edges in the network. For example,



$$l_{AD} = l_{AB} + l_{BC} + l_{CD},$$

$$l_{EF} = \infty.$$

The total length of all edges is equal to $T = \sum_{i<j} A_{ij} l_{ij}$.

We assume the cost of building and maintaining the network to be proportional to T.

Travel cost

The cost of traveling or shipping a commodity from vertex i to vertex j depends on

- the distance l_{ij} and
- the amount of traffic w_{ij}.

We assume that w_{ij} is proportional to the populations in the Voronoi cells V_i and V_j, so that

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

in appropriate units.

The total travel cost in our model is then $Z = \sum_{i<j} w_{ij} l_{ij}$.

Total network cost

Construction cost $T = \sum_{i<j} A_{ij} l_{ij}$, travel cost $Z = \sum_{i<j} w_{ij} l_{ij}$.

The total cost of running the network is proportional to

$$C = T + \gamma Z,$$

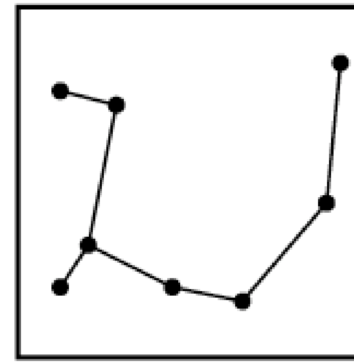
where $\gamma \geq 0$ measures the relative importance of the two terms.

The best network is the one that minimizes the total cost C for given facility locations $\mathbf{r}_1, \ldots, \mathbf{r}_p$ and given γ.

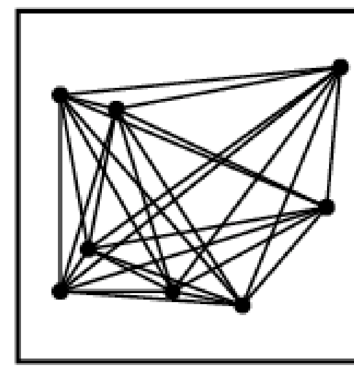
The meaning of γ

$$\text{total cost } C = T + \gamma Z$$

$\gamma \to 0$: The total cost is $\approx T$ (geometric length of all edges).
$\longrightarrow$ minimum spanning tree

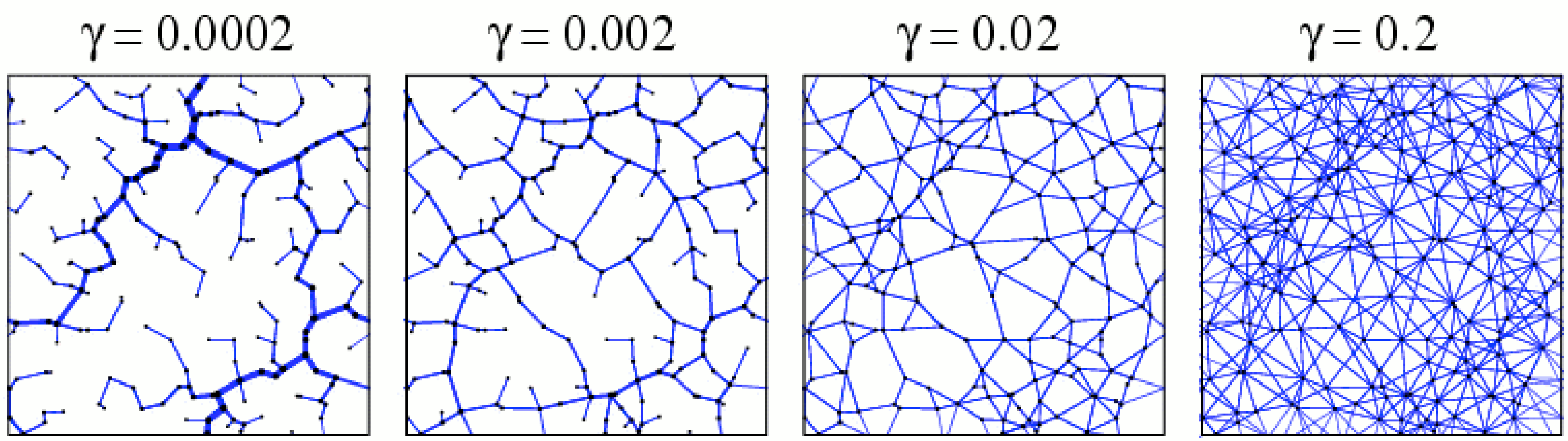


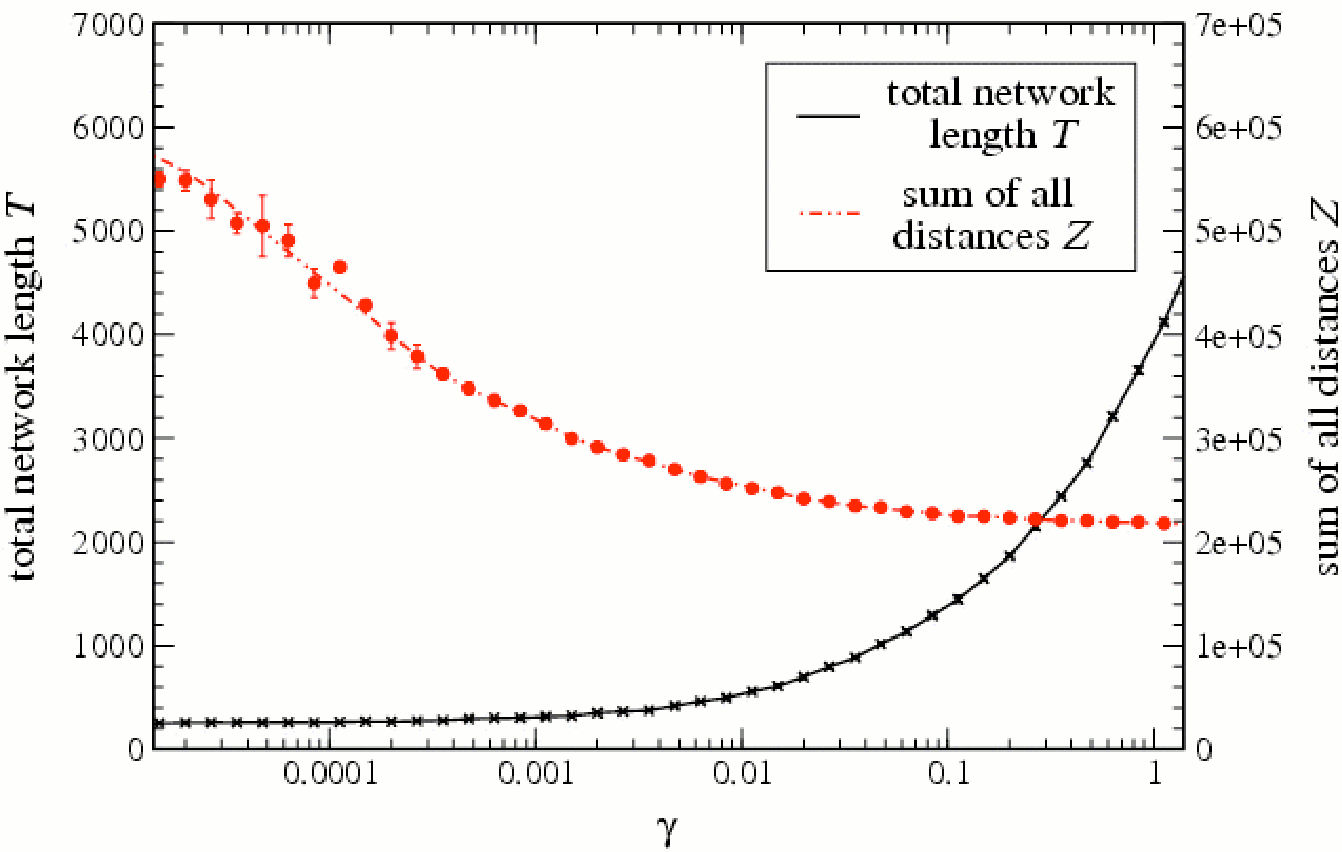
$\gamma \to \infty$: The total cost is approximately $\propto Z$ (sum of effective distances between vertices).
$\longrightarrow$ fully connected network



For intermediate γ: Non-trivial (presumably NP-complete) optimization problem.

200 random vertex positions, periodic boundary conditions, and setting all $w_{ij} = \text{const.}$





It is possible to achieve a small travel cost T for a construction cost Z not much higher than that of the minimum spanning tree.

What is a realistic value for γ?

For simplicity's sake, let us normalize the length scale by setting the average "crow flies" distance between a vertex and its nearest neighbor equal to 1.

We can make an order of magnitude estimate as follows.

The sum $T = \sum_{i<j} A_{ij} l_{ij}$ has m nonzero terms, where m is the number of edges in the network.

Most real networks are sparse, with $m = O(p)$, and edges are of typical length 1 in our length scale, so that $T = O(p)$.

What is a realistic value for γ?

The sum $Z = \sum_{i<j} w_{ij} l_{ij}$ contains $\frac{1}{2}p(p-1) = O(p^2)$ nonzero terms.

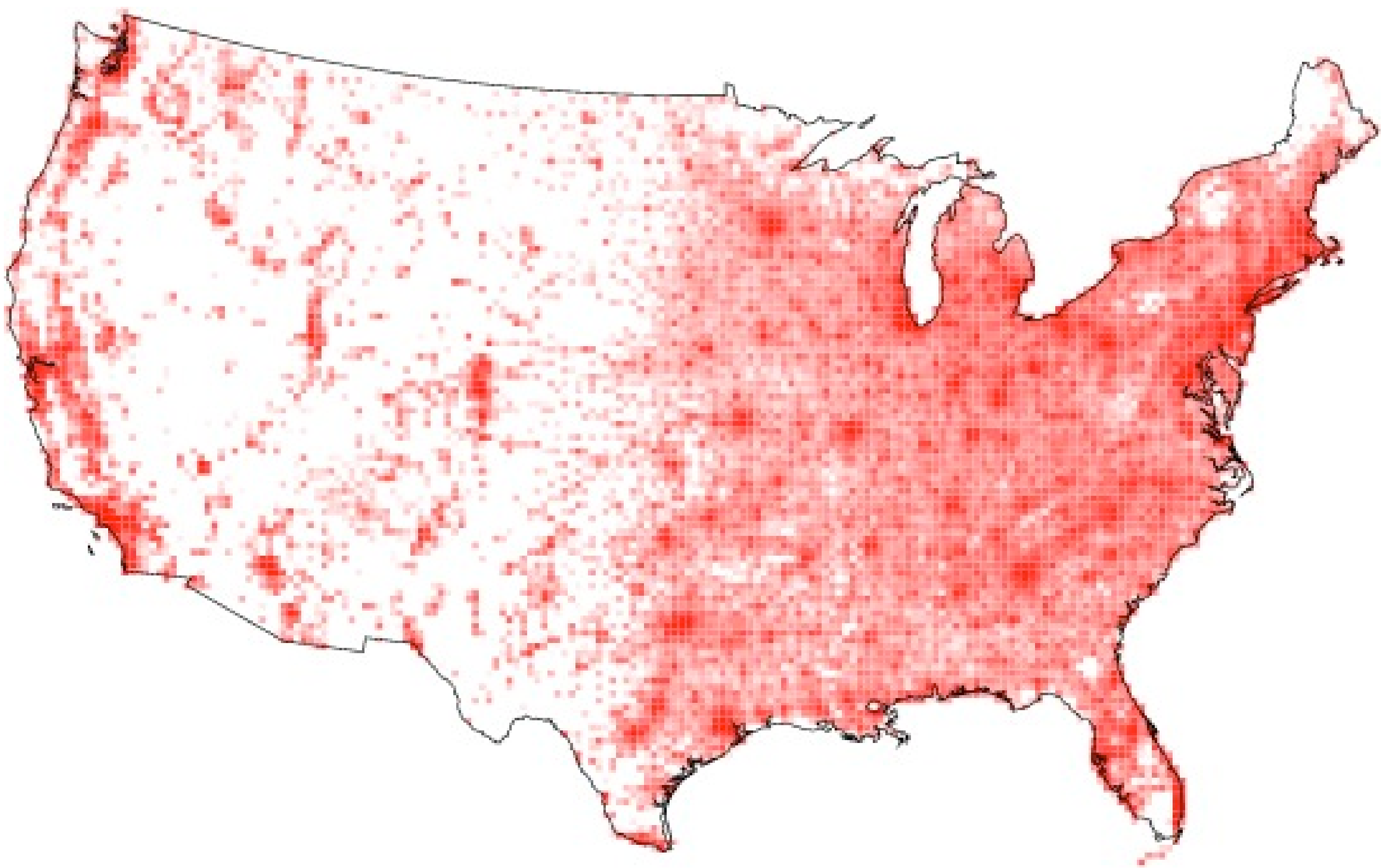
If P is the total population, the weights w_{ij} have typical value $(P/p)^2$. Thus $Z = O(P^2)$.

We assume that investments in maintenance and travel costs are of the same order of magnitude: $T = O(p) \approx \gamma Z$.

In our examples $p = 200$ and $P = 2.8 \times 10^8$ for the U.S. which leads to $\gamma \approx 10^{-14}$.

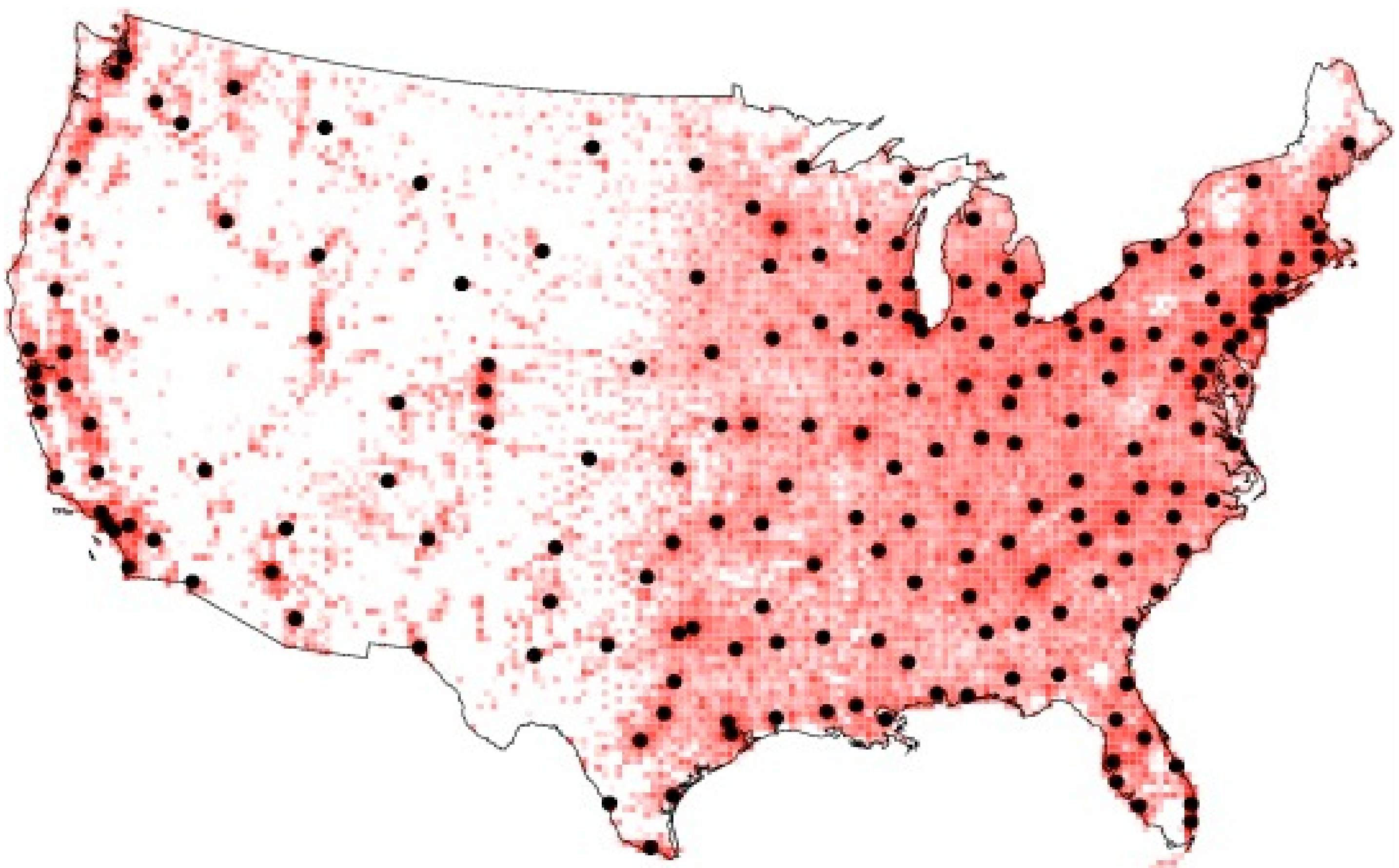
Optimal networks of optimally located facilities

The optimal network design problem then consists of two parts.
First, we distribute p facilities on the map by solving the p-median problem.



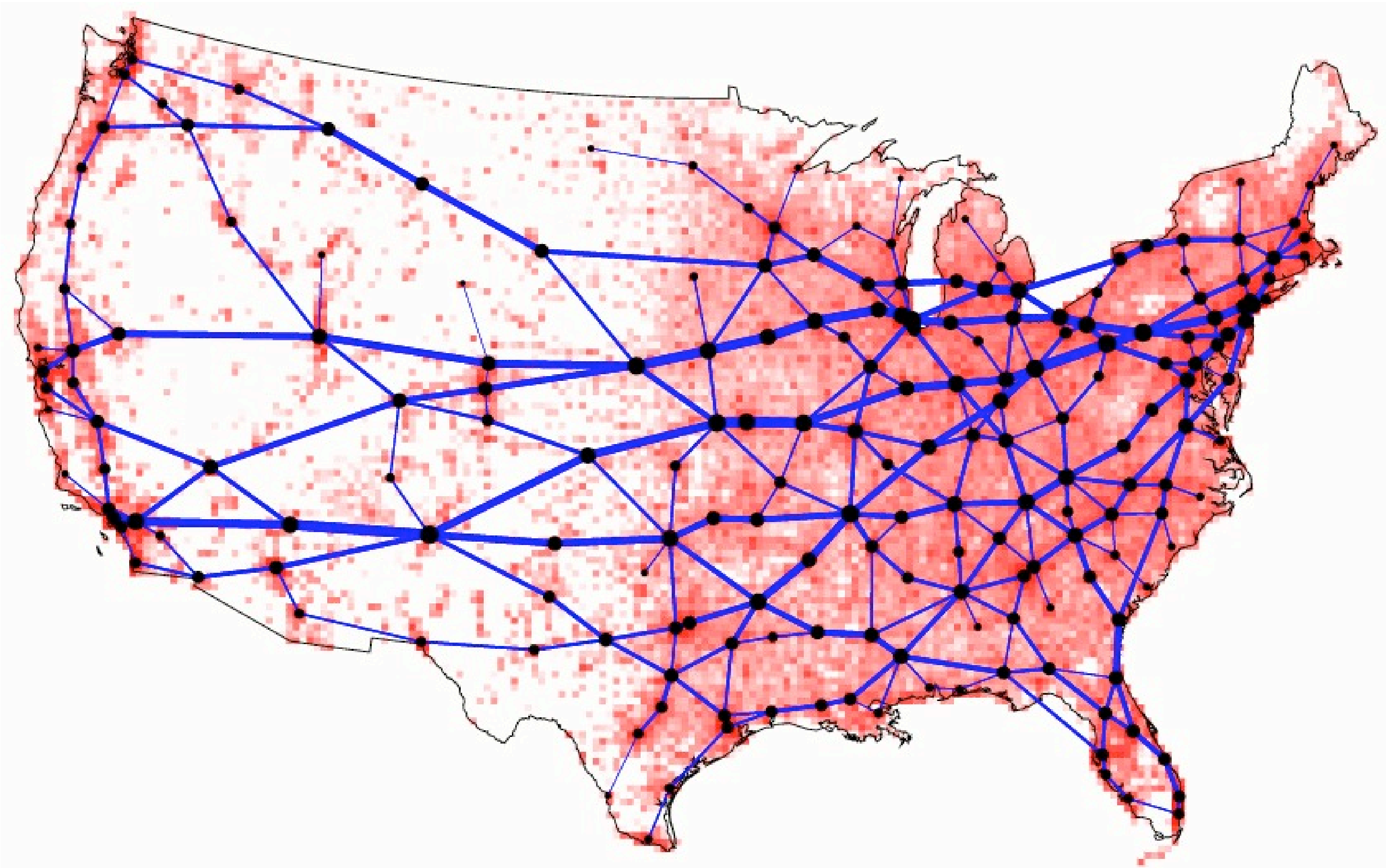
Optimal networks of optimally located facilities

The optimal network design problem then consists of two parts.
First, we distribute p facilities on the map by solving the p-median problem.
Then we find the network minimizing the total cost C.



Optimal networks of optimally located facilities

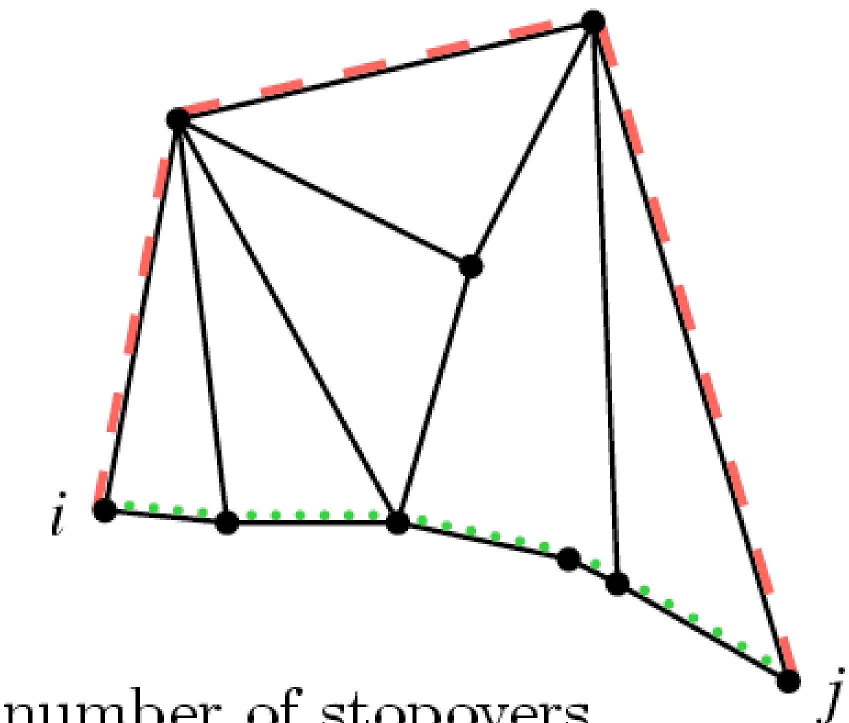
The optimal network design problem then consists of two parts.
First, we distribute p facilities on the map by solving the p-median problem.
Then we find the network minimizing the total cost C.



Different routing strategies

There is another complicating factor. We have assumed that travel costs are proportional to geometric distances.

In some networks, users may not choose the geometrically shortest path, especially if it has many edges.



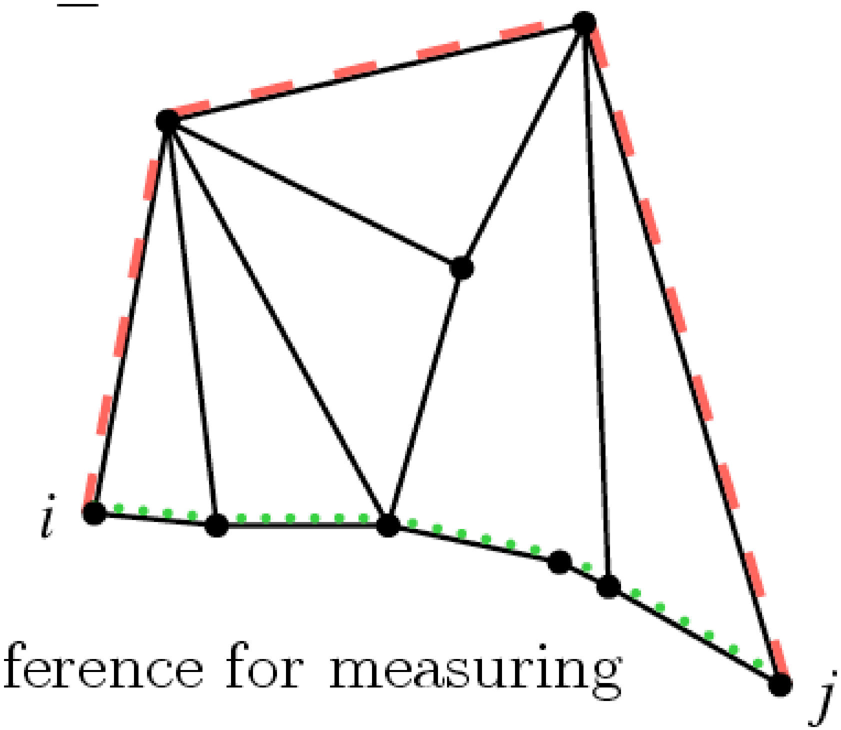
Examples:

- Airline passengers want to limit the number of stopovers.
- Internet packets arrive more quickly and reliably if the number of routers along the way is small.

Different routing strategies

We can account for such situations by using a more flexible notion of distance. We assign to each pair of adjacent vertices an effective length (travel cost)

$$\tilde{l}_{ij} = (1-\delta)l_{ij} + \delta, \quad 0 \le \delta \le 1.$$

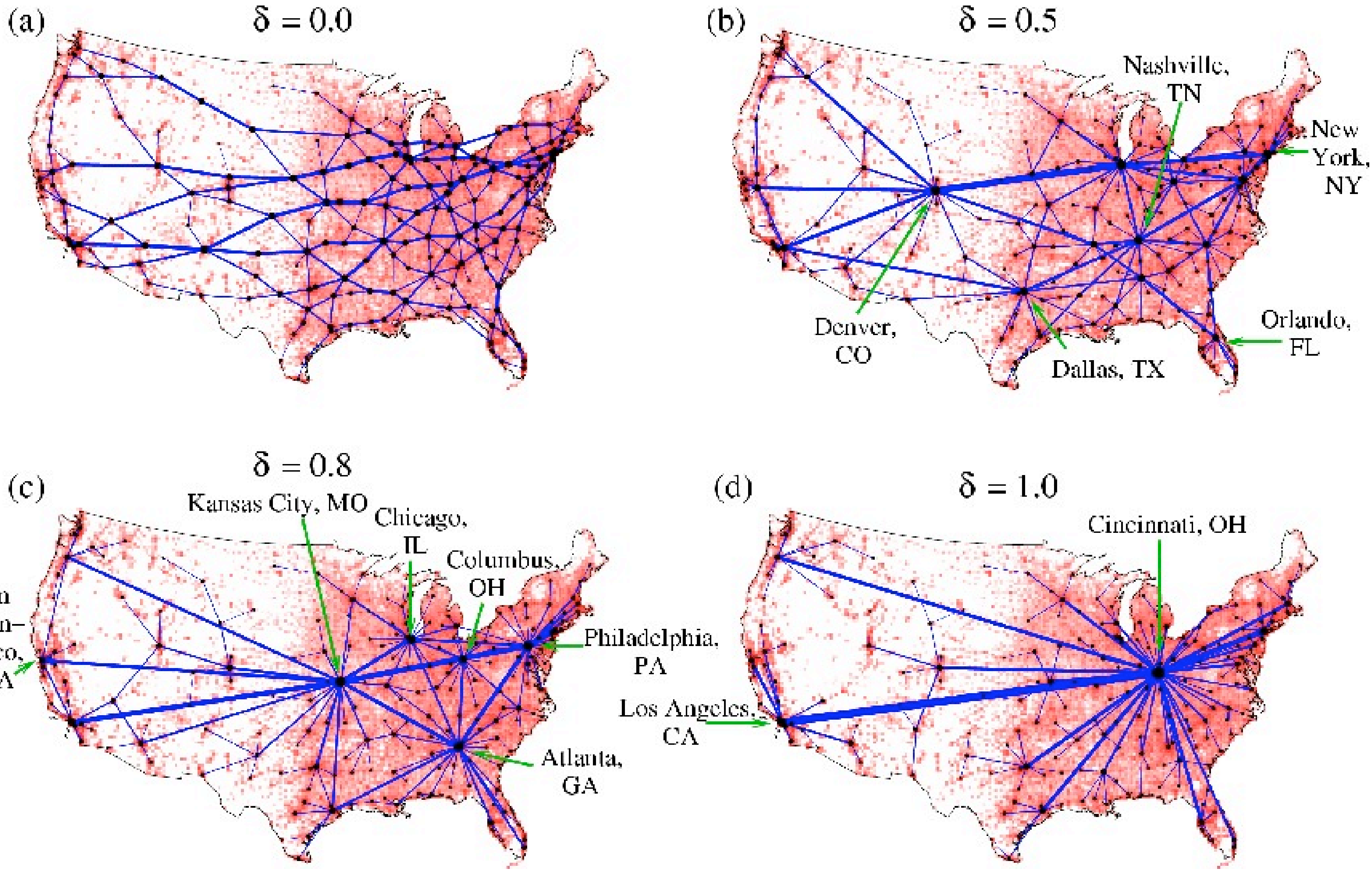


The travel cost is then $Z = \sum_{i<j} w_{ij}\tilde{l}_{ij}$.

The parameter δ determines the user's preference for measuring distance in terms of kilometers or edges:

- $\delta = 0$: geometric distance
- $\delta = 1$: number of edges (graph distance)

Different routing strategies



How realistic is our model?

Is there a 2/3 power law for real facilities?

G. Edward Stephan, Science (1977):

Territorial Division: The Least-Time Constraint Behind the Formation of Subnational Boundaries

Such an analysis has been carried out for 98 modern nations (*4*). While the slopes for individual nations vary somewhat around the expected −2/3 value (and in some cases the number of subdivisions within a nation was too small to permit adequate statistical test), the aggregated 1764 political subdivisions did yield a regression slope between log-size and log-density of −0.66, a result which very clearly conforms to the theoretical expectation developed here.

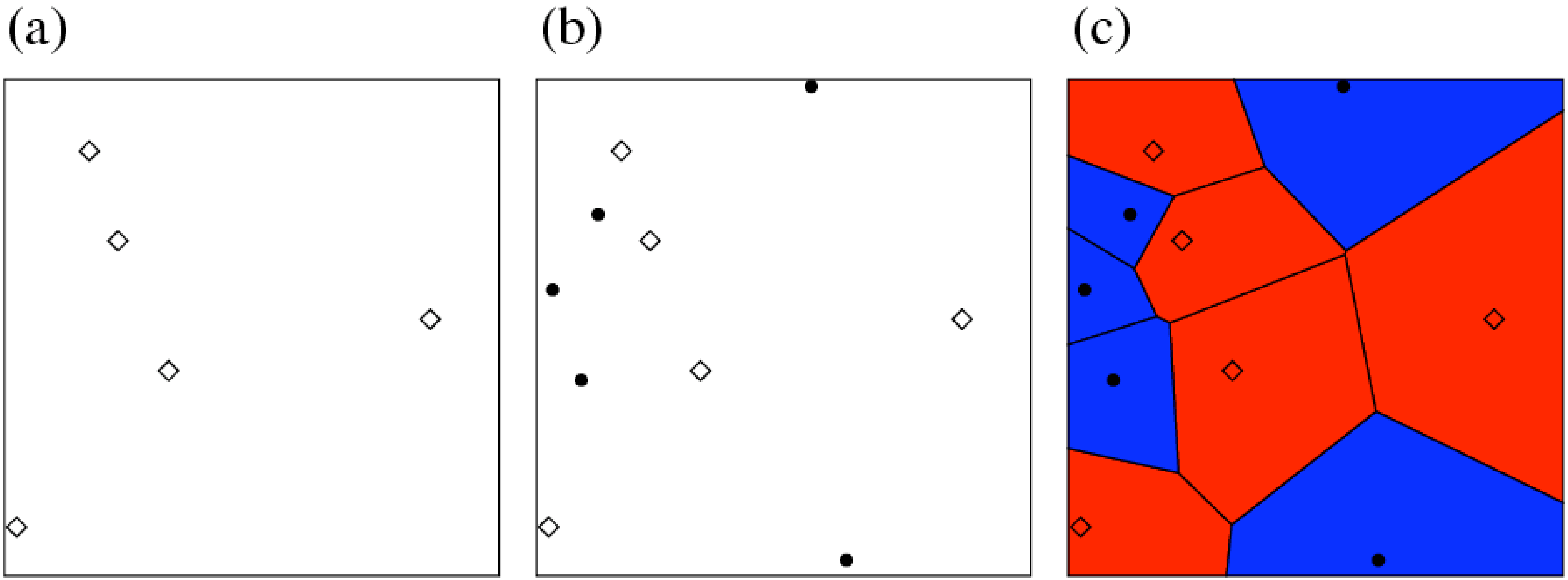
G. Edward Stephan, J. of Reg. Sci. (1988)

	Service Establishments	b	r^2	f	g	α	β	b_a	r^2_a
1	Alcoholic Beverages, *p*	0.92	0.77	0.99	0.13	1.14	0.31	0.87	0.77
2	Alcoholic Beverages	0.93	0.97	0.91	0.03	0.94	0.07	0.97	0.97
3	Bicycle Shops, *p*	0.97	0.82	1.17	0.19	1.45	0.48	0.81	0.86
4	Bicycle Shops	0.88	0.97	0.86	0.08	0.93	0.17	0.92	0.97
5	Collection Agencies, *p*	0.80	0.96	0.82	0.22	1.05	0.56	0.78	0.96
6	Collection Agencies	0.76	0.92	0.68	0.13	0.78	0.29	0.87	0.95
7	Drive-in Theaters, *p*	0.73	0.90	0.68	0.22	0.87	0.57	0.78	0.91
8	Drive-in Theaters	0.87	0.87	0.75	0.04	0.78	0.09	0.96	0.90
9	Motion Pictures, *p*	0.81	0.96	0.76	0.15	0.89	0.35	0.85	0.97
10	Motion Pictures	0.85	0.94	0.73	0.04	0.77	0.09	0.96	0.97
11	Musical Instruments, *p*	0.89	0.97	0.93	0.14	1.08	0.33	0.86	0.97
12	Musical Instruments	0.91	0.96	0.86	0.03	0.89	0.07	0.97	0.97
13	Second-hand Stores, *p*	0.96	0.94	1.04	0.12	1.18	0.27	0.88	0.95
14	Second-hand Stores	0.87	0.92	0.83	0.07	0.90	0.16	0.93	0.92
15	Gas Stations, *p*	0.87	0.99	0.85	0.11	0.95	0.24	0.89	0.99
16	Gas Stations, *d*	0.93	0.99	0.92	0.05	0.97	0.11	0.95	0.99
17	Gas Stations	0.96	0.99	0.94	0.03	0.97	0.06	0.97	0.99
18	Farm Supplies, *p*	0.62	0.60	0.75	0.48	1.43	1.84	0.52	0.64
19	Farm Supplies	0.88	0.90	0.86	0.09	0.95	0.20	0.91	0.90
20	Mobile Home Dealers, *p*	0.66	0.86	0.66	0.34	0.99	1.02	0.66	0.86
21	Mobile Home Dealers	0.96	0.94	0.92	0.01	0.93	0.02	0.99	0.94
22	Apparel and Accessories	0.95	0.97	0.94	0.03	0.97	0.07	0.97	0.97
23	Auto Accessories	1.00	0.99	0.99	0.00	0.99	0.00	1.00	0.99
24	Bookstores	0.81	0.93	0.72	0.04	0.75	0.09	0.96	0.95
25	Building Materials	0.98	0.99	0.99	0.03	1.02	0.07	0.97	0.99
26	Cameras/Photography	0.89	0.94	0.85	0.06	0.91	0.13	0.94	0.94
27	Department Stores	0.82	0.95	0.76	0.09	0.85	0.19	0.91	0.96
28	Dry Cleaners	0.97	0.99	0.96	0.02	0.98	0.04	0.98	0.99
29	Electrical Supplies	0.76	0.91	0.67	0.09	0.75	0.21	0.91	0.92
30	Furniture Stores	0.93	0.97	0.92	0.06	0.97	0.12	0.94	0.97
31	Grocery Stores	0.94	0.84	0.89	0.02	0.91	0.05	0.98	0.84
32	Hardware Stores	0.96	0.98	0.96	0.03	1.00	0.07	0.97	0.98
33	Pharmacies	0.95	0.96	0.90	0.01	0.91	0.01	0.99	0.97
34	Plumbing and Heating	0.76	0.92	0.72	0.16	0.85	0.37	0.84	0.92
35	Radio and TV Stores	0.93	0.97	0.87	0.00	0.88	0.01	1.00	0.97
36	Record Stores	0.87	0.97	0.81	0.05	0.85	0.11	0.95	0.98
37	Restaurants	0.89	0.98	0.85	0.06	0.91	0.14	0.94	0.98
38	Stationery Stores	0.96	0.98	0.93	0.00	0.94	0.01	1.00	0.98
39	Tires and Batteries	0.98	0.98	0.99	0.03	1.02	0.05	0.97	0.98
40	Travel Agencies	0.90	0.97	0.86	0.04	0.90	0.08	0.96	0.97
41	Used-car Dealers	0.96	0.95	0.92	0.00	0.92	0.00	1.00	0.95
42	Variety Stores	0.89	0.97	0.86	0.09	0.95	0.21	0.91	0.97
43	Total Retail Stores	0.94	0.99	0.92	0.05	0.97	0.11	0.95	0.99
44	Baptist Churches, *a*	0.88	0.99	0.88	0.10	0.97	0.22	0.90	0.99
45	Catholic Churches, *a*	0.70	0.93	0.64	0.21	0.81	0.52	0.79	0.94
46	Newspapers	0.69	0.88	0.62	0.23	0.81	0.58	0.77	0.89
47	Hospitals, *p*	0.68	0.95	0.68	0.32	0.99	0.92	0.68	0.95
48	1920 County Seats, *r*	0.64	0.86	0.64	0.36	1.00	1.14	0.64	0.86
49	1970 County Seats, *r*	0.55	0.74	0.59	0.48	1.14	1.84	0.52	0.74
50	1970 County Seats, *p*	0.43	0.55	0.39	0.54	0.84	2.33	0.46	0.56
51	Physicians, *p*	1.09	0.98	1.10	−0.08	1.02	−0.16	1.08	0.98

→ Most commercial facilities follow power laws with exponents closer to 1 than to 2/3.

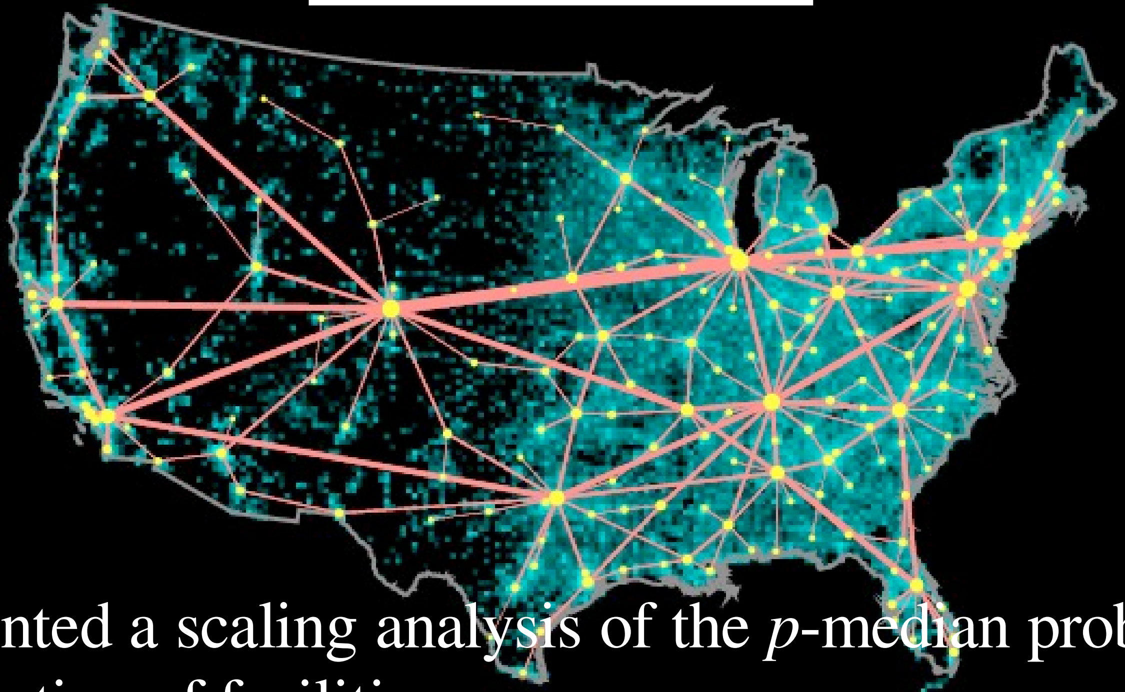
A model for commercial facilities

The *Voronoi game* is a simple model for competitive facility location.



(a) Player $\mathcal{A}$ first places a fixed number of facilities (diamonds).
(b) Then player $\mathcal{B}$, knowing $\mathcal{A}$'s decision, places her facilities (circles).
(c) The square is tessellated into Voronoi cells and each player receives the population in her cells as payoff.

Conclusion



- We have presented a scaling analysis of the *p*-median problem, a model for the optimal location of facilities.
- The density of facilities grows as population density to the 2/3 power.
- Cartograms can be used to verify this relation.
- We have studied a model of optimal spatial network design with tunable parameters for the cost functions involved.
- Depending on the parameters we find structures ranging from decentralized to hub-and-spoke networks.
- We mentioned limitations and extensions of our model.