Optimal design of spatial distribution networks

Michael T. Gastner[1,2] and M. E. J. Newman[2,3]

[1]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

[2]*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*

[3]*Center for the Study of Complex Systems, University of Michigan, Ann Arbor, Michigan 48109, USA*

(Received 13 April 2006; published 24 July 2006)

We consider the problem of constructing facilities such as hospitals, airports, or malls in a country with a nonuniform population density, such that the average distance from a person's home to the nearest facility is minimized. We review some previous approximate treatments of this problem that indicate that the optimal distribution of facilities should have a density that increases with population density, but does so slower than linearly, as the two-thirds power. We confirm this result numerically for the particular case of the United States with recent population data using two independent methods, one a straightforward regression analysis, the other based on density-dependent map projections. We also consider strategies for linking the facilities to form a spatial network, such as a network of flights between airports, so that the combined cost of maintenance of and travel on the network is minimized. We show specific examples of such optimal networks for the case of the United States.

DOI: 10.1103/PhysRevE.74.016117 PACS number(s): 89.65.−s, 01.75.+m, 89.75.Da, 89.75.Hc

I. INTRODUCTION

Suppose we are given the population density $\rho(\mathbf{r})$ of a country or province, by which we mean the number of people per unit area as a function of geographical position $\mathbf{r}$. And suppose we are charged with choosing the sites of p facilities, such as hospitals, post offices, supermarkets, gas stations, or schools, so that the mean distance to the nearest facility averaged over the population is minimized. In most countries, population density is highly nonuniform, in which case a uniform distribution of facilities would be a poor choice: it benefits us little to build a lot of facilities in sparsely populated areas. A more sensible choice would be to distribute facilities in proportion to population density, so that a region with twice as many people has twice as many facilities. But this distribution too turns out to be suboptimal, because we also gain little by having closely spaced facilities in the highly populated areas—when facilities are closely spaced the typical person is not much farther from their second-closest facility than from their closest, so one or the other can often be removed with little penalty and substantial savings.

Although an exact analytic solution to this optimal location problem has yet to be found, a variety of approximate treatments have been given, which suggest that the ideal solution lies somewhere between these two extremes, with the density of facilities increasing as the two-thirds power of population density, a prediction that we verify here using simulations and visualizations based on cartograms, with actual population data for the United States. In addition, one is often interested in connections between facilities, such as flights between airports [1] or transmission lines between power stations [2]. In the second half of this paper, we generate networks based on a simple model that optimizes network topology with respect to the cost of maintaining and traveling across the network. Depending on the benefit function chosen, we find structures ranging from completely decentralized networks to hub-and-spoke networks.

II. OPTIMAL DISTRIBUTION OF FACILITIES

We wish to distribute p facilities over a two-dimensional area A such that the objective function

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

is minimized. Here $\{\mathbf{r}_1, \ldots, \mathbf{r}_p\}$ is the set of positions of the facilities and $\rho(\mathbf{r})$ is the population density within the region A of interest. This objective function is proportional to the mean distance that a person will have to travel to reach their nearest facility.

Seemingly simple, this so-called *p-median problem* has been shown to be NP-hard [3], so in practice most studies rely either on approximate numerical optimization or approximate analytic treatments [4]. A number of different approaches have been used [5–9]; the calculation given here is essentially that of Gusein-Zade [10].

Our p facilities naturally partition the area A into Voronoi cells. (The Voronoi cell V_i for the ith facility is defined as the set of points that are closer to $\mathbf{r}_i$ than to any other facility.) Let us define $s(\mathbf{r})$ to be the area of the Voronoi cell to which the point $\mathbf{r}$ belongs. In two dimensions, a person living at point $\mathbf{r}$ will on average be a distance $g[s(\mathbf{r})]^{1/2}$ from the nearest facility, where g is a geometric factor of order 1, whose exact value depends on the shape of the Voronoi cell, but which will in any case drop out of the final result. The distance to the nearest facility averaged over all members of the population is proportional to

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

where we are making an approximation by neglecting variation of the geometric factor g between cells.

The value of $s(\mathbf{r})$ is constrained by the requirement that there be p facilities in total. Noting that $s(\mathbf{r})$ is constant and equal to $s(\mathbf{r}_i)$ within Voronoi cell V_i, we see that the integral of $[s(\mathbf{r})]^{-1}$ over V_i is

$$\int_{V_i} [s(\mathbf{r})]^{-1} d^2r = [s(\mathbf{r}_i)]^{-1} \int_{V_i} d^2r = 1. \tag{3}$$

Summing over all V_i, we can then express the constraint on the number of facilities in the form

$$\int_A [s(\mathbf{r})]^{-1} d^2r = p. \tag{4}$$

Subject to this constraint, optimization of the mean distance f above gives

$$\frac{\delta}{\delta s(\mathbf{r})} \left[g \int_A \rho(\mathbf{r})[s(\mathbf{r})]^{1/2} d^2r - \alpha \left(p - \int_A [s(\mathbf{r})]^{-1} d^2r \right) \right] = 0, \tag{5}$$

where α is a Lagrange multiplier. Performing the functional derivatives and rearranging for $s(\mathbf{r})$, we find $s(\mathbf{r}) = [2\alpha/(g\rho(\mathbf{r}))]^{2/3}$. The Lagrange multiplier can be evaluated by substituting into Eq. (4), and we arrive at the result

$$D(\mathbf{r}) = \frac{1}{s(\mathbf{r})} = p \frac{[\rho(\mathbf{r})]^{2/3}}{\int [\rho(\mathbf{r})]^{2/3} d^2r}, \tag{6}$$

where we have introduced the notation $D(\mathbf{r}) = [s(\mathbf{r})]^{-1}$ for the density of the facilities.

Thus, if facilities are distributed optimally for the given population distribution, their density should increase with population density but it should do so slower than linearly, as a power law with exponent $\frac{2}{3}$ [29]. In addition to the argument given here, which roughly follows Ref. [10], this result has also been derived previously by a number of other methods [5–9], although all are approximate.

Equation (6) places most facilities in the densely populated areas where most people live while still providing reasonable service to those in sparsely populated areas where a strictly population-proportional allocation might leave inhabitants with little or nothing. Its derivation makes two approximations: it assumes that the geometric factor g is the same for all Voronoi cells and that $s(\mathbf{r})$ is a continuous function. Neither assumption is strictly true, but we expect them to be approximately valid if ρ varies little over the typical size of a Voronoi cell. As a test of these assumptions, we have optimized numerically the distribution of $p=5000$ facilities over the lower 48 states of the United States (Fig. 1) using population data from the most recent U.S. Census [11], which counts the number of residents within more than 8 million blocks across the study region. To create a continuous density function ρ, we convolved these data with a normalized Gaussian distribution of width 20 km [30]. The facility locations were then determined by optimizing the full p-median objective function (1) by simulated annealing [12].

The relation $D \propto \rho^{2/3}$ can be tested as follows. First, we determine the Voronoi cell around each facility. Then we calculate $D(\mathbf{r})$ as the inverse of the area of the corresponding cell and ρ as the number of people living in the cell divided by its area. Figure 2 shows a scatter plot of the resulting data on doubly logarithmic scales. If the anticipated $\frac{2}{3}$-power relation holds, we expect the data to fall along a line of slope $\frac{2}{3}$. And indeed a least-squares fit (solid line in the figure) yields a slope 0.66 with $r^2=0.94$.

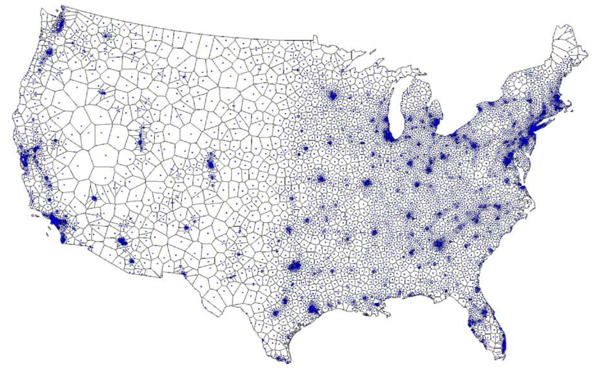

FIG. 1. (Color online) Facility locations determined by simulated annealing and the corresponding Voronoi tessellation for $p=5000$ facilities located in the lower 48 United States, based on population data from the U.S. Census for the year 2000.

Some statistical concerns might be raised about this method. First, we used the Voronoi cell area to calculate both D and ρ, so the measurements of x and y values in the plot are not independent, and one might argue that a positive slope could thus be a result of artificial correlations between the values rather than a real result [13]. Second, it is known that estimating the exponent of a power law such as Eq. (6) from a log-log plot can introduce systematic biases [14,15]. In the next section, we introduce an entirely different test of Eq. (6) that, in addition to being of interest in its own right, suffers from neither of these problems.

III. DENSITY-EQUALIZING PROJECTIONS

If we neglect finite-size effects, it is straightforward to demonstrate that optimally located facilities in a uniformly

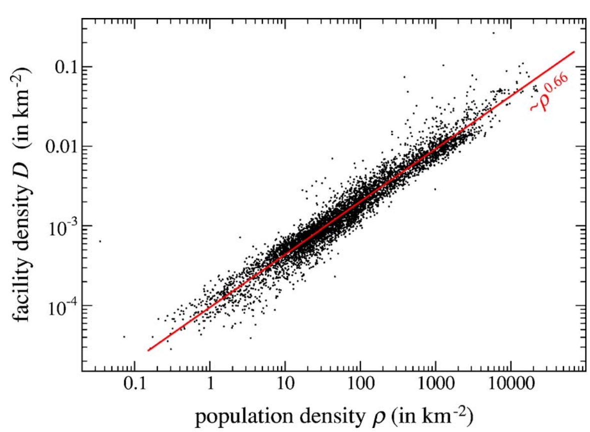

FIG. 2. (Color online) Facility density D from Fig. 1 vs population density ρ on a log-log plot. A least-squares linear fit to the data gives a slope of 0.66 (solid line, $r^2=0.94$).

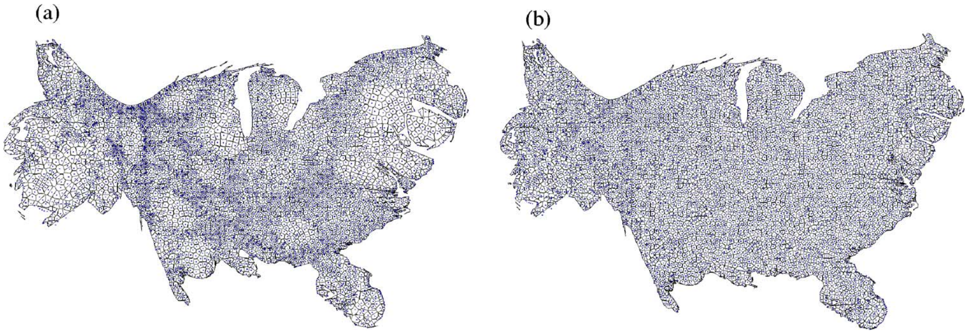

FIG. 3. (Color online) Near-optimal facility location on (a) a cartogram equalizing the population density ρ and (b) a cartogram equalizing $\rho^{2/3}$.

populated space lie on the vertices of a regular triangular lattice [16]. It has been conjectured that for a non uniform population there is a general class of map projections that will transform the pattern of facilities to a similarly regular structure [17]. The obvious candidate projections are population density equalizing maps or *cartograms*, i.e., maps in which the sizes of geographic regions are proportional to the populations of those regions [18–21]. Densely populated regions appear larger on a cartogram than on an equal-area map such as Fig. 1, and the opposite is true for sparsely populated regions. Since most facilities are located where the population density is high, a cartogram projection will effectively reduce the facility density in populated areas and increase it where the population density is low. Therefore, one might expect that a cartogram leads to a more uniform facility density than that shown in Fig. 1. And indeed some authors have used population density equalizing projections as the basis for facility location methods [22,23].

In Fig. 3(a), we show the facilities of Fig. 1 on a population density equalizing cartogram created using the diffusion-based technique of [24]. Although the *population* density is now equal everywhere, the *facility* density is obviously far from uniform. A comparison between Figs. 1 and 3(a) reveals that we have overshot the mark since the facilities are now concentrated in areas where there are few in actual space.

Equation (6) makes clear what is wrong with this approach. Since D grows slower than linearly with ρ, a projection that equalizes ρ will necessarily overcorrect the density of facilities. On the other hand, based on our earlier result, we would expect a projection equalizing $\rho^{2/3}$ instead of ρ to spread out the facilities approximately uniformly. Hence, one way to determine the actual exponent for the density of facilities is to create cartograms that equalize ρ^x, $x \geq 0$, and find the value of x that minimizes the variation of the Voronoi cell sizes on the cartogram. This approach does not suffer from the shortcomings of our previous method based on the doubly logarithmic plot in Fig. 2, since we neither use the Voronoi cells to calculate the population density nor take logarithms. One might argue that the Voronoi cells on the cartogram are not equal to the projections of the Voronoi cells in actual space, which is true—the cells generally will not even remain polygons under the cartogram transformation. The difference, however, is small if the density does not vary much between neighboring facilities.

In Fig. 4, we show the measured coefficient of variation (i.e., the ratio of the standard deviation to the mean) for Voronoi cell sizes on ρ^x cartograms as a function of the exponent x (solid curve). As the figure shows, the minimum is indeed attained at or close to the predicted value of $x=\frac{2}{3}$. Figure 3(b) shows the corresponding cartogram for this exponent. This projection finds a considerably better compromise between regions of high and low population density than either Fig. 1 or Fig. 3(a).

For comparison, we have also made the same measurement for 5000 points distributed randomly in proportion to population. Since the density of these points is by definition equal to ρ, we expect the minimum standard deviation of the cell areas to occur on a cartogram with $x=1$. Our numerical results for this case (dashed curve in Fig. 4) agree well with this prediction. Comparing the solid and the dashed curves in the plot, we see that not only the positions of the minima differ, but also the minimal values themselves. The lower standard deviation for the p-median distribution indicates that optimally located facilities are not randomly distributed with a density $\propto \rho^{2/3}$. Instead, the optimally located facilities occupy space in a relatively regular fashion reminiscent of the triangular lattice of the uniform population case [16,25]. We can confirm this observation by measuring the interior angles formed by the edges of the Voronoi cells. The Voronoi cells of a triangular lattice are regular hexagons and hence all the interior angles are 120°. Figure 5 shows a histogram of the angles for the cells in the equal-area projection of Fig. 1. Since the population is nonuniform, the cells are not exactly regular hexagons, but, as Fig. 5 shows, the angles are none-

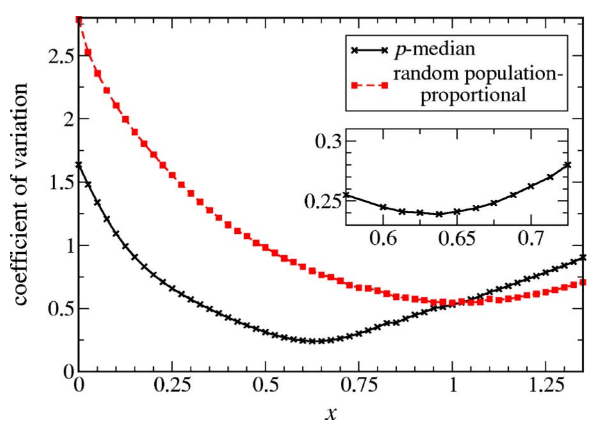

FIG. 4. (Color online) The coefficient of variation (i.e., the ratio of the standard deviation to the mean) for Voronoi cell areas as they appear on a cartogram, against the exponent x of the underlying density ρ^x for a p-median (solid curve) and a random population-proportional distribution (dashed curve). Inset: An expanded view of the minimum for the p-median distribution.

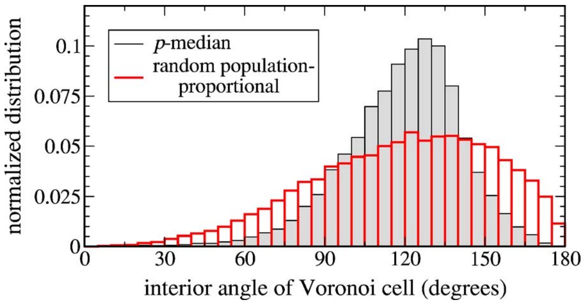

FIG. 5. (Color online) The distribution of angles in the Voronoi diagram for a p-median and a random population-proportional distribution of facilities.

theless narrowly distributed around 120°—more so than for the cells of the random distribution.

IV. OPTIMAL NETWORKS OF FACILITIES

In many cases of practical interest, finding the optimal location of facilities is only half the problem. Often facilities are interconnected to form networks, such as airports connected by flights or warehouses connected by truck deliveries. In these cases, one would also like to find the best way to connect the facilities so as to optimize the performance of the system as a whole.

Consider then a situation in which our facilities form the nodes or vertices of a network and connections between them form the edges. The efficiency of this 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 nodes or distribution of mail or cargo). 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, while a network with a large number of direct links is usually expensive to build and operate. The optimal solution lies somewhere between these extremes.

Let us define l_{ij} to be the shortest geographic distance between two vertices i and j measured along the edges in the network. If there is no path between i and j, we formally set $l_{ij}=\infty$. Introducing the adjacency matrix $\mathbf{A}$ with elements $A_{ij}=1$ if there is an edge between i and j and $A_{ij}=0$ otherwise, we can write the total length of all edges as

$$T=\sum_{i<j} A_{ij} l_{ij}. \tag{7}$$

We assume this quantity to be proportional to the cost of maintaining the network. Clearly this assumption is only approximately correct; networked systems in the real world will have many factors affecting their maintenance costs that are not accounted for here. It is, however, the obvious first assumption to make and, as we will see, can provide us with good insight about network structure.

The typical cost of shipping a commodity or traveling through the network depends on the distances l_{ij} as well as the amount of traffic w_{ij} (e.g., weight of cargo, number of passengers, etc.) that flows between vertices i and j [26]. In a spirit similar to our assumption about maintenance costs, we assume that the total travel cost is proportional to

$$Z=\sum_{i<j} w_{ij} l_{ij}. \tag{8}$$

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

$$w_{ij}=\int_{V_i} \rho(\mathbf{r}) d^2r \int_{V_j} \rho(\mathbf{r}') d^2r' \tag{9}$$

in appropriate units. And the total cost of running the network is proportional to the sum $T+\gamma Z$ with $\gamma \geq 0$ a constant that measures the relative importance of the two terms. Then the optimal network is the one minimizing this sum [27,28].

Using again the contiguous 48 states of the United States as an example, we have first determined the optimal placement of $p=200$ facilities, which we then try to connect together optimally. The number of edges in the network depends on the parameter γ. If $\gamma \to 0$, the cost of travel γZ vanishes and the optimal network is the one that simply minimizes the total length of edges. That is, it is the minimum spanning tree, with exactly $p-1$ edges between the p vertices. Conversely, if $\gamma \to \infty$, then Z dominates the optimization, regardless of the cost T of maintaining the network, so that the optimum is a fully connected network or clique with all $\frac{1}{2}p(p-1)$ possible edges present. For intermediate values of γ, finding the optimal network is a nontrivial combinatorial optimization problem. The number of edges increases with γ, but it is difficult to determine the exact set of edges optimizing the cost. Nevertheless, we can derive good, though usually not perfect, solutions using again the method of simulated annealing [31].

There is, however, another complicating factor. In Eq. (8), we assumed that travel costs are proportional to geometric distances between vertices, which is a plausible starting point. In a road network, for example, the quickest and cheapest route is usually not very different from the shortest route measured in kilometers. But in other networks, travel costs can also depend on the number of legs in a journey. In an airline network, for instance, passengers often spend a lot of time waiting for connecting flights, so that they care both about the total distance they travel and the number of stopovers they have to make. Similarly, the total time required for an Internet packet to reach its destination depends on two factors, the propagation delay proportional to the physical distance between vertices (computers and routers) and the store and forward delays introduced by the routers, which grow with the number of intermediate vertices.

To account for such situations, we generalize our definition of the length of an edge and assign to each edge an effective length

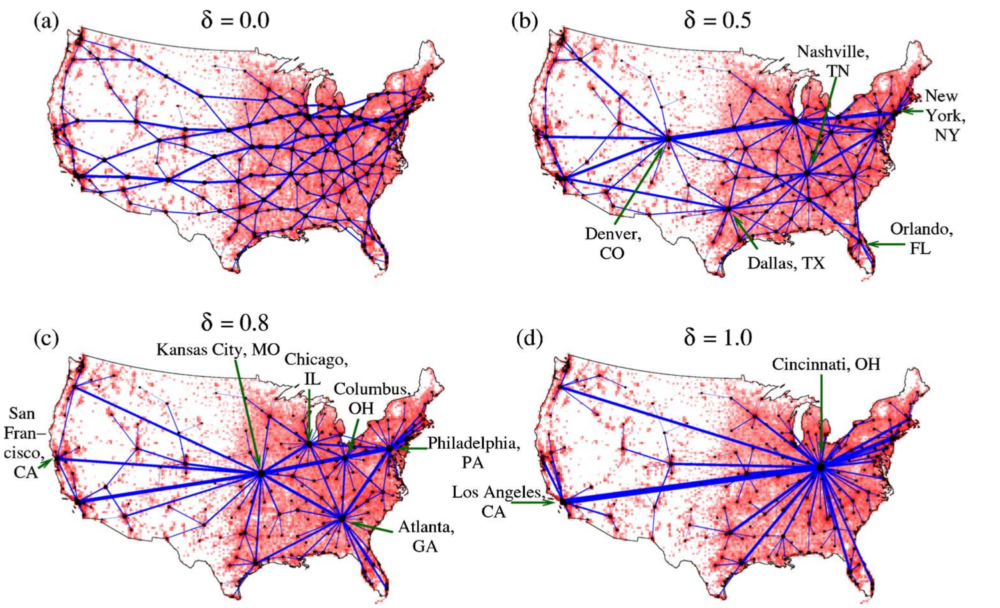

FIG. 6. (Color online) Optimal networks for the population distribution of the United States with $p=200$ vertices and $\gamma=10^{-14}$ for different values of δ.

$$\tilde{l}_{ij} = (1-\delta) l_{ij} + \delta \tag{10}$$

with $0 \leq \delta \leq 1$. The parameter δ determines the user's preference for measuring distance in terms of kilometers or legs. Now we define the effective distance between two (not necessarily adjacent) vertices to be the sum of the effective lengths of all edges along a path between them, minimized over all paths. The travel cost is then proportional to the sum of all effective path lengths

$$Z = \sum_{i<j} w_{ij} \tilde{l}_{ij}, \tag{11}$$

and the optimal network for given γ and δ is again the one that minimizes the total cost $T+\gamma Z$. Since the second term in Eq. (10) is dimensionless, we normalize the length appearing in the first term by setting the average "crow flies" distance between a vertex and its nearest neighbor equal to 1.

What is a realistic value for γ? We can make an order of magnitude estimate as follows. The sum in Eq. (7) has m nonzero terms, where m is the number of edges in the network. Most real networks are sparse, with $m=O(p)$. Furthermore, edges are of typical length 1 in our length scale, so that $T=O(p)$, with $p \simeq 200$ in the examples studied here. The sum in Eq. (11), on the other hand, 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 $Z=O(P^2) \simeq 10^{17}$ for the U.S. with a current population of $P \simeq 2.8 \times 10^8$. Assuming that our investments in maintenance and travel costs are of the same order of magnitude and setting $T \approx \gamma Z$ then leads to an estimate for γ of order 10^{-15}or 10^{-14}.

In Fig. 6, we show the results for $\gamma=10^{-14}$. When $\delta=0$, passengers (or cargo shippers) care only about total kilometers traveled and the optimal network strongly resembles a network of roads, such as the U.S. interstate network. As δ increases, the number of legs in a journey starts playing a more important role and the approximate symmetry between the vertices is broken as the network begins to form hubs. Around $\delta=0.5$, we see networks emerging that constitute a compromise between the convenience of direct local connections and the efficiency of hubs, while by $\delta=0.8$ the network is dominated by a few large hubs in Philadelphia, Columbus, Chicago, Kansas City, and Atlanta that handle the bulk of the traffic. On the highly populated California coast, two smaller hubs around San Francisco and Los Angeles are visible. In the extreme case $\delta=1$, where the user cares only about number of legs and not about distance at all, the network is dominated by a single central hub in Cincinnati, with a few smaller local hubs in other locations such as Los Angeles.

V. CONCLUSIONS

We have studied the problem of optimal facility location, also called the p-median problem, which consists of choosing positions for p facilities in geographic space such that the mean distance between a member of the population and the nearest facility is minimized. Analytic arguments indicate that the optimal density of facilities should be proportional to the population density to the two-thirds power. We have confirmed this relation by solving the p-median problem numerically and projecting the facility locations on density-equalizing maps. We have also considered the design of optimal networks to connect our facilities together. Given optimally located facilities, we have searched numerically for the network configuration that minimizes the sum of maintenance and travel costs. A simple two-parameter model allows us to take different user preferences into account. The model gives us intuition about a number of situations of practical interest, such as the design of transportation networks, parcel delivery services, and the Internet backbone.

ACKNOWLEDGMENTS

The authors thank the staff of the University of Michigan's Numeric and Spatial Data Services for their help. This work was funded in part by the National Science Foundation under Grant No. DMS–0405348 and by the James S. McDonnell Foundation.

[1] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, Proc. Natl. Acad. Sci. U.S.A. **102**, 7794 (2005).
[2] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[3] N. Megiddo and K. Supowit, SIAM J. Comput. **13**, 182 (1984).
[4] *Approximation Algorithms for NP-hard Problems*, edited by D. S. Hochbaum (PWS, Boston, 1997).
[5] J. Mycielski and W. Trzeciakowski, J. Regional Sci. **5**, 59 (1963).
[6] J. Viriakis, Ekistics **27**, 362 (1969).
[7] D. S. Palmer, Oper. Res. Q. **24**, 121 (1973).
[8] G. E. Stephan, Science **196**, 523 (1977).
[9] C. Cameron, S. Low, and D. Wei, in Proceedings of the 2002 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (Marina del Rey, 2002), pp. 152–159.
[10] S. M. Gusein-Zade, Geogr. Anal. **14**, 246 (1982).
[11] U. S. Census Bureau, Census 2000: Census of Population and Housing, DVD-ROM (2003).
[12] M. T. Gastner, Ph.D. thesis, University of Michigan, Ann Arbor, MI (2005).
[13] D. R. Vining, C.-H. Yang, and S.-T. Yeh, Science **205**, 219 (1979).
[14] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).
[15] M. L. Goldstein, S. A. Morris, and G. G. Yen, Eur. Phys. J. B **41**, 255 (2004).
[16] A. Lösch, *Die Räumliche Ordnung der Wirtschaft* (Fischer, Jena, 1940).
[17] W. Bunge, Discussion paper 3, University of Michigan Inter-University Community of Mathematical Geographers (1964).
[18] S. M. Guseyn-Zade and V. S. Tikunov, Mapping Sciences and Remote Sensing **31**, 49 (1994).
[19] D. Dorling, Concepts and Techniques in Modern Geography (CATMOG) **59** (1996).
[20] B. D. Dent, *Cartography: Thematic Map Design*, 5th ed. (WCB/McGraw-Hill, Boston, 1999).
[21] W. Tobler, Ann. Assoc. Am. Geogr. **94**, 58 (2004).
[22] A. Getis, Economic Geography **39**, 14 (1963).
[23] G. Rushton, Papers of the Regional Science Association **28**, 111 (1971).
[24] M. T. Gastner and M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **101**, 7499 (2004).
[25] W. Christaller, *Die Zentralen Orte in Süddeutschland*, 3rd ed. (Wissenschaftliche Buchgesellschaft, Darmstadt, 1980).
[26] M. Barthélemy and A. Flammini, e-print physics/0601203.
[27] J. W. Billheimer and P. Gray, Transp. Sci. **7**, 49 (1973).
[28] M. Los and C. Lardinois, Transp. Res., Part B: Methodol. **16**, 89 (1982).
[29] The analogous calculation can also be carried out in general dimension d; the corresponding exponent in that case is $d/(d+1)$.
[30] Since people are discrete objects, the census data must be binned or coarse-grained at some level to define a population *density*. If we average over a length scale comparable to the mean distance between neighboring facilities, ρ varies little within most Voronoi cells, making the approximations in Eqs. (2)–(6) reasonable. These arguments break down if there are only very few facilities to be built because it becomes impossible to define a meaningful density of facilities D. But for hospitals or department store chains in the United States, the number of facilities is on the order of 10^3–10^4, which is large enough to reliably define a density.
[31] In our simulated annealing calculations, we use three different update moves: we randomly add or delete edges or we swap the end points of two randomly chosen edges. We used an exponential cooling schedule with an initial temperature equal to one-tenth of the total cost $T+\gamma Z$ of the minimum spanning tree, lowered by a factor $1-(3\times 10^{-6})$ after each Monte Carlo step. The optimization was stopped after 10^6 consecutive Monte Carlo steps failed to reduce the total cost. We set the starting temperature deliberately high so that all traces of the initial network are quickly erased. Since there are typically many local minima, the algorithm might not always return exactly the same network, even for this rather conservative cooling schedule. The results presented here are the best networks obtained during five independent runs in each case.