Title	Telecommunication networks and multifractal analysis of human population distribution
Author(s)	Mannersalo, Petteri; Koski, Antti; Norros, Ilkka
Citation	COST257 Technical Document COST257TD(98)02, 17 p.
Date	1998
Rights	This article may be downloaded for personal use only

Telecommunication networks and multifractal analysis of human population distribution

Petteri Mannersalo* Antti Koski† Ilkka Norros‡

Abstract

Multifractal analysis is applied to the population distribution of Finland. A relation between generalized q-dimensions and minimal cable length needed to interconnect the whole population by a star network is demonstrated. Furthermore, a cost estimate suitable for dimensioning hierarchical networks is presented.

1 Introduction

Fractal properties of human population distributions have been studied for several years. For example, it has been shown that in many countries their population distribution exhibits multifractal scaling over a large range of resolutions [App96, Adj97]. Steve Appleby from British Telecom might be the first to combine multifractal analysis of population distributions and telecommunication network planning [App94, App95]. Loosely speaking, his idea was that in some minimizing problems solutions behave according to simple power laws depending only on parameters related to scaling of a population distribution.

After obtaining very accurate data about the population distribution of Finland, we have examined whether Appleby's results are valid in this case

*VTT Information Technology, P.O.Box 1202, FIN-02044 VTT, Finland. E-mail: petteri.mannersalo@vtt.fi

†VTT Information Technology, P.O.Box 1202, FIN-02044 VTT, Finland. E-mail: antti.koski@vtt.fi

‡VTT Information Technology, P.O.Box 1202, FIN-02044 VTT, Finland. E-mail: ilkka.norros@vtt.fi

too. First we show that, indeed, the population distribution of Finland exhibits multifractal scaling. Then we consider how to optimally build one-layer star networks when the number of nodes is given. By combining the results from the above topics, we show that the minimal cable length needed to connect all the inhabitants of Finland as a function of number of nodes is a linear function on a log-log scale with the slope equal to the reciprocal of the box-counting dimension of the population distribution. This is exactly the same result that Appleby found when studying the population distributions of Great Britain and the United States. Finally, assuming that this relation holds for all consecutive layers, we state a cost estimate that could be used for dimensioning hierarchical star networks.

2 Multifractal analysis and fractal q-dimensions

The material needed to analyze discrete data is introduced in this section. More comprehensive presentations are found, e.g., in [Fal90, Man88, Rie95]

Consider a compactly supported probability measure μ sampled by a mesh consisting of quadrants with diameter δ. Let $I_\delta^{i,j}$ denote quadrant (i,j) at scale δ and $\Lambda_\delta = \left\{ I_\delta^{i,j} \ : \ \mu(I_\delta^{i,j}) > 0 \right\}$, the set of non-empty quadrants. In order to check whether multifractal scaling exists we calculate the partition sum

$$S_\delta(q) = \sum_{I_\delta^{i,j} \in \Lambda_\delta} \mu \left(I_\delta^{i,j} \right)^q , \qquad q \in \mathbb{R}$$

at different scales δ. If $S_\delta(q)$ is a linear function of δ in some region in a log-log scale we say that the region in question is the scaling region and the measure exhibits multifractal scaling there. Evidently, the partition sum is influenced by the choice of origin for the quadrants, i.e., the position of the grid relative to the distribution. Natural choice would be to place the origin such that the partition sum is minimized. Nevertheless, usually the difference is significant only if the quadrant size is about the same order as that of the whole support.

The generalized q-dimensions are determined as

$$D(q) = \frac{\tau(q)}{(q-1)}, \tag{1}$$

where

$$\tau(q) = \liminf_{\delta \to 0} \frac{\log S_\delta(q)}{\log \delta}.$$

In practical applications, this limit is unattainable and we need to approximate the partition function $\tau(q)$ by solving the equation

$$\log S_\delta(q) \approx \tau(q) \log \delta + \text{const} \tag{2}$$

for $\tau(q)$ in the least square sense in the scaling region. After solving $\tau(q)$, the corresponding generalized q-dimension is obtained by equation (1).

The case $q = 1$ needs a bit extra care. First notice that

$$\lim_{q \to 1} \frac{\log S_\delta(q)}{q-1} = \sum_{I_\delta^{i,j} \in \Lambda_\delta} \mu\left(I_\delta^{i,j}\right) \log\left(\mu\left(I_\delta^{i,j}\right)\right) = -H_\delta,$$

where H_δ is known as the Shannon information or entropy. Thus, instead of equations (1) and (2), $D(1)$ is found by solving

$$H_\delta \approx -D(1) \log \delta + \text{const}$$

in the least square sense.

Another function characterizing scaling properties is the Legendre spectrum f_L which is also defined by $\tau(q)$:

$$f_L(\alpha) = \inf_{q \in \mathbb{R}} (\alpha q - \tau(q)).$$

Loosely speaking, f_L measures asymptotic probability to observe a coarse Hölder exponent α when picking a quadrant at random: for self-similar measures

$$\mathrm{P}\left(\alpha(I_\delta(x)) \approx \alpha\right) \sim \delta^{D(0) - f_L(\alpha)},$$

where

$$\alpha(I_\delta(x)) = \frac{\log \mu(I_\delta(x))}{\log \delta}$$

is the coarse Hölder exponent of a quadrant $I_\delta(x)$ containing point x. Thus, the probability of observing any exponent $\alpha \neq \alpha_0$, with $f_L(\alpha_0) = D(0)$, tends to zero as the quadrant size gets smaller.

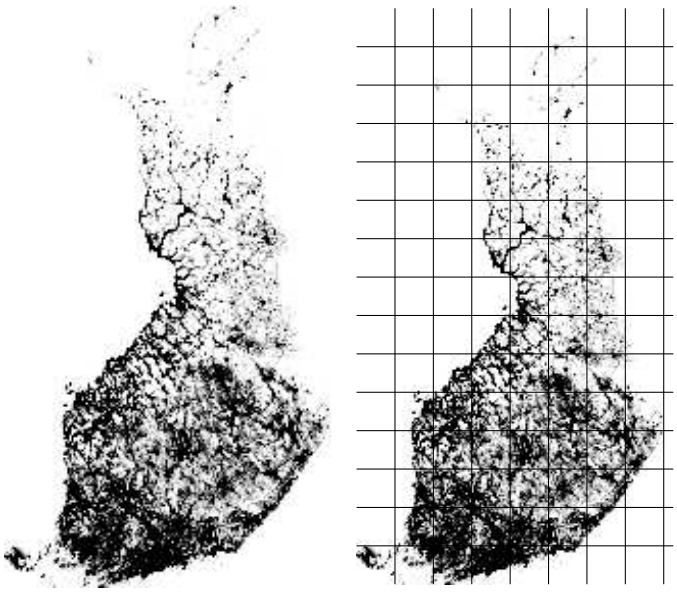

Figure 1: The population distribution of Finland and a mesh of quadrants of size 80×80 km^2.

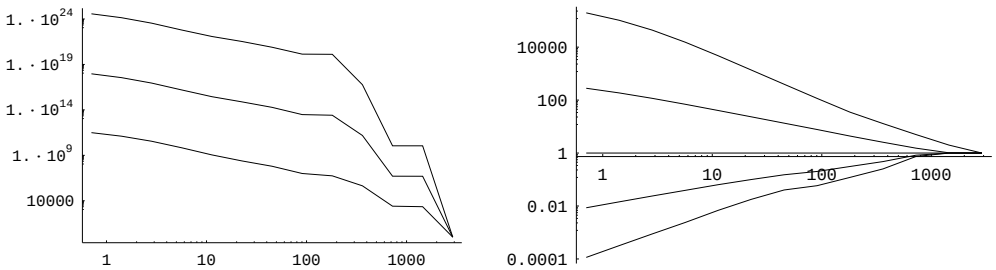

Figure 2: Multifractal scaling of the population distribution of Finland. Partition sum against quadrant size: on the left $q = -3, -2, -1$ and on the right $q = 0, 0.5, \ldots, 2$.

2.1 Multifractal analysis of the population distribution of Finland

The population distribution of Finland at resolution 0.5×0.5 km^2 was obtained from Finnish Central Statistical Office. Multifractal scaling was tested by calculating partition sums over resolutions $2^i \times 2^i$ km^2, $i = -1, 0, 1, \ldots, 11$, (see figure 2).

If q is non-negative, the scaling is good over all resolutions, besides few largest ones. When using the resolutions of order 100 km or more, the fact that the partition sums were calculated with the same fixed mesh origin caused that the scaling was no more so fine.

At first glance, the partition sums calculated with negative values of q seem to be also scalable. Unfortunately, this is only an illusion. When considering the linear region, that is, resolutions from 0.5 km to 100 km, the number of quadrants with few, say one or two, habitants inside remains almost constant. On the other hand, behavior of the partition sum for negative q-values is dominated by these quadrants. Thus $\tau(q)$ is approximately constant and multifractal scaling does not exist.

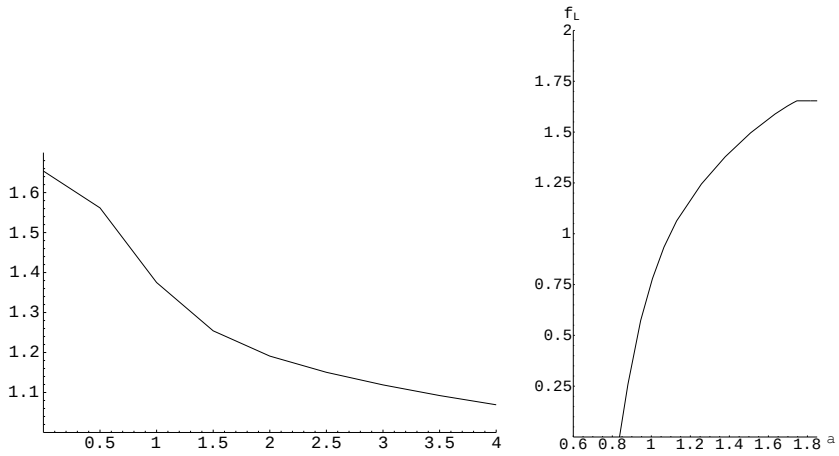

Figure 3: Generalized q dimensions and Legendre spectrum of the population distribution of Finland.

Generalized q-dimensions were approximated from the log-log plots. The values changed a bit depending on which resolutions were used in the least square fitting, for example $D(0) = 1.65 \pm 0.05$. The Legendre spectrum was

calculated using only positive q-values. The maximum of the spectrum was obtained with $\alpha \approx 1.65$ and the corresponding value $f_L(\alpha) \approx D(0)$.

3 Self-similar distributions and minimizing problems

3.1 Strictly self-similar measures

Consider a strictly self-similar set S related to an iterated function scheme (IFS), i.e., a set that is invariant under a collection of similarities $G_1, \ldots, G_m : \mathbb{R}^2 \to \mathbb{R}^2$,

$$|G_i(x) - G_i(y)| = c_i|x - y|, \tag{3}$$

where $0 < c_i < 1$. In other words, S satisfies the equation

$$\bigcup_{i=1}^{m} G_i(S) = S.$$

For convenience, let the $G_i(S)$'s be disjoint. Let $p_1, \ldots, p_m$ be positive numbers with $\sum_{i=1}^{m} p_i = 1$. A self-similar measure μ on S can be defined by setting $\mu(G_{i_1} \circ \ldots \circ G_{i_l}(S)) = p_{i_1} \cdots p_{i_l}$.

Consider the problem of finding optimal placements for nodes $\mathbf{r}_i$, $i = 1, \ldots, n$, such that

$$f^{(\beta)}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \int_{\mathbb{R}^2} \min_{i=1,\ldots,n} \left\{ |\mathbf{x} - \mathbf{r}_i|^\beta \right\} \, d\mu(\mathbf{x}), \quad \beta \in \mathbf{R}^+, \tag{4}$$

is minimized. If we can solve this problem for $n = 1$ then an upper bound for all $n = m^l$ is easily found. Denote $f_n^{(\beta)} = \min f^{(\beta)}(\mathbf{r}_1, \ldots, \mathbf{r}_n)$. Making a

change of variables and using definition (3) we get

$$\begin{aligned} f_m^{(\beta)} &= \min f^{(\beta)}(\mathbf{r}_1, \ldots, \mathbf{r}_m) \le \sum_{i=1}^{m} \min_{\mathbf{r}_i \in \mathbb{R}^2} \int_{G_i(S)} |\mathbf{x} - \mathbf{r}_i|^\beta \; d\mu(\mathbf{x}) \\ &= \sum_{i=1}^{m} \min_{\mathbf{r}_i \in \mathbb{R}^2} \int_S |G_i(\mathbf{y}) - \mathbf{r}_i|^\beta \; p_i \; d\mu(\mathbf{y}) \\ &= \sum_{i=1}^{m} \min_{\mathbf{r}_i \in \mathbb{R}^2} \int_S |\mathbf{y} - G_i^{-1}(\mathbf{r}_i)|^\beta \; c_i^\beta \; p_i \; d\mu(\mathbf{y}) \\ &= \min_{\mathbf{r_1} \in \mathbb{R}^2} \left\{ \int_S |\mathbf{y} - \mathbf{r}_1|^\beta \; d\mu(\mathbf{y}) \right\} \sum_{i=1}^{m} c_i^\beta \; p_i = f_1^{(\beta)} \sum_{i}^{m} c_i^\beta \; p_i. \end{aligned}$$

Respectively,

$$f_{m^l}^{(\beta)} \le f_1^{(\beta)} \sum_{i_1=1}^{m} \cdots \sum_{i_l=1}^{m} p_{i_1} \cdots p_{i_l} \; c_{i_1}^\beta \cdots c_{i_l}^\beta = f_1^{(\beta)} \left(\sum_{i=1}^{m} p_i \; c_i^\beta \right)^l, \quad l = 0, 1, 2 \ldots \; . \tag{5}$$

Though (5) is only an upper bound, it characterizes quite well the behavior of $f_n^{(\beta)}$. Assume that the inequality holds for all integers n, i.e.,

$$\tilde{f}_n^{(\beta)} \le \tilde{f}_1^{(\beta)} \left(\mathrm{E}_{\mathbf{p}}(c^\beta) \right)^{\frac{\log n}{\log m}},$$

where $\mathrm{E}_{\mathbf{p}}(c^\beta) = \sum_{i=1}^{m} p_i c_i^\beta$. Taking logarithm of both sides leads to the inequality

$$\log \tilde{f}_n^{(\beta)} \le \frac{\log \mathrm{E}_{\mathbf{p}}(c^\beta)}{\log m} \; \log n + \log \tilde{f}_1^{(\beta)} = C(\mathbf{c}, \beta) \log n + \text{const}, \tag{6}$$

which means that $\tilde{f}_n^{(\beta)}$ is majorized by a linear function on a log-log scale. If $p_i = 1/m$ and $c_i = c$ for all i, then $C(\mathbf{c}, 1) = \log c / \log m = -1/D(0)$, where $D(0)$ is the box-counting dimension of the set defined by similarities G_i, $i = 1, \ldots, m$, satisfying $|G_i(x) - G_i(y)| = c \; |x - y|$. In this case, (5) is an equality. If the c_i's are identical, the left hand side of (5) reduces to $f_1^{(\beta)} c_1^{l\beta}$, so that the upper bound is insensitive with respect to p_i's.

3.2 An example: Sierpinski triangle

As a concrete example, also appearing in [App95], we consider a self-similar distribution defined on a triangle. In this case, the similarities G_i, $i = 1, 2, 3$, are mappings that shrink a triangle by factor 0.5 and move it onto one of the corners of the original triangle (see figure 4). Suppose that the mass is evenly distributed, i.e, $p_i = 1/3$, $i = 1, 2, 3$, and try to find the minimum cable length needed to interconnect this population to a given number of nodes; determine $f_n^{(1)}$ for arbitrary n.

If the population is served by one node, its position should be somewhat shown on the left in figure 5, and when the number of nodes is increased by three, it is clear that the nodes should be placed at the same location with respect to the half-size triangles as the single node was respect to the full-size triangle. This means that increasing the number of nodes by three decreases the total and average cable length by factor 0.5. Thus $f_n^{(1)}$, $n = 3^i$, satisfies equation (6) with $C((0.5, 0.5, 0.5), 1) = -\log 2/\log 3$, i.e., the total (weighted) cable length is a power law whose exponent is $\log 2/\log 3 = -1/D(0)$, the reciprocal of the box-counting dimension of the Sierpinski triangle. In the case of figure 4, the upper bound given by inequality (5) has exactly the same power law.

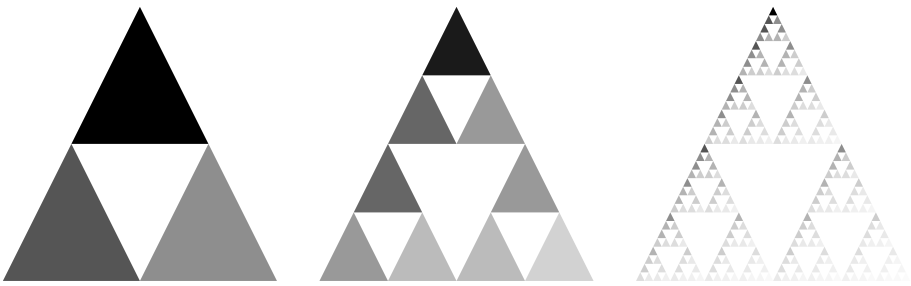

Figure 4: Three different resolutions of a self–similar measure defined on a Sierpinski's triangle. $p_1 = 0.5$, $p_2 = 0.3$ and $p_3 = 0.2$.

3.3 Statistical self-similarity

The concept of a self-similar measure can be enlarged to statistically self-similar random measures too. In this case the support of each realization is distributed on a statistical self-similar set, and weights of each subset are

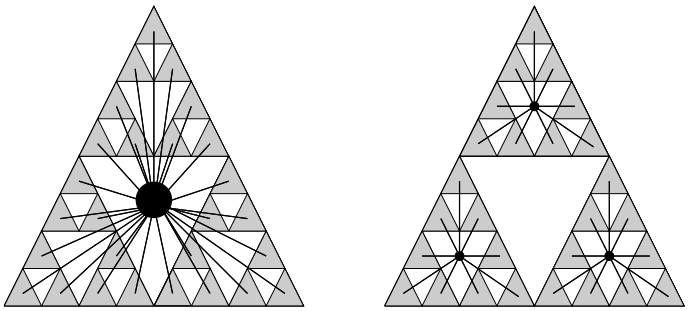

Figure 5: Connecting a population distributed on a Sierpinski triangle.

defined by some randomized procedure. In practice, a test for statistical self-similarity could be performed by testing the linearity of the partition sum on a log-log scale.

4 One-layer star networks

Suppose that an arbitrary population distribution is given and we want to determine the minimal cable length needed to interconnect all the inhabitants by a one-layer star network with n nodes. The problem is to find the positions and catchment areas for the nodes. Unfortunately, for general n no algorithm that would surely converge to the global solution is known, and we are forced to use some weaker methods.

4.1 Optimal way to place one node

Consider the problem of placing one node so that the following function is minimized:

$$f(\mathbf{r}) = \int_A |\mathbf{r} - \mathbf{x}|^\beta \mu(d\mathbf{x}), \quad \beta = 1, 2,$$

where $\mathbf{r}$ is the location of the node and μ is a population mass distribution. Using standard tools of the calculus of variations we find that the optimally placed node satisfies:

$$\beta = 1: \quad \int_{\mathbb{R}^2} \frac{\mathbf{r} - \mathbf{x}}{|\mathbf{r} - \mathbf{x}|} \mu(d\mathbf{x}) = 0$$

$$\beta = 2: \quad \int_{\mathbb{R}^2} (\mathbf{r} - \mathbf{x})\, \mu(d\mathbf{x}) = 0.$$

One should notice that the solution for $\beta = 1$ is not always unique. For example, if a mass is distributed into two points, the optimal placement of a node is anywhere in between them.

If μ is atomic, then the corresponding equations are

$$\beta = 1: \quad \sum_i \frac{p_i(\mathbf{r} - \mathbf{x}_i)}{|\mathbf{r} - \mathbf{x}_i|} = 0,$$

$$\beta = 2: \quad \sum_i p_i\,(\mathbf{r} - \mathbf{x}_i) = 0,$$

where p_i is the proportion of the population located at point $\mathbf{x}_i$. In the case $\beta = 1$, an optimal placement is found by a fixed point iteration

$$\mathbf{r} = \frac{\sum_i \frac{p_i \mathbf{x}_i}{|\mathbf{r} - \mathbf{x}_i|}}{\sum_i \frac{p_i}{|\mathbf{r} - \mathbf{x}_i|}}, \tag{7}$$

and in the case $\beta = 2$, simply by calculating the center of mass

$$\mathbf{r} = \frac{\sum_i p_i \mathbf{x}_i}{\sum_i p_i} = \sum_i p_i \mathbf{x}_i. \tag{8}$$

When performing the fixed point iteration, a good initial guess is the center of mass, that is, the solution for $\beta = 2$. In many cases the difference between solutions of these two problems is small, so that only a few iterative steps are needed. (In [App95], only $\beta = 2$ is used for simplicity.)

4.2 K–means algorithm

Our problem is to find the places of the nodes such that the total cable length of each person from the nearest node is minimal. The problem of finding the optimal placement for a large number of nodes for a given population is a hard one. There are several clustering methods to approximate the density centroids of a distribution, but the K-means algorithm [Sch92] is perhaps the best known; it was used also in [App95]. In the K-means algorithm, nodes are initially placed onto the map and the following calculations are repeated

until the situation does not get sufficiently better: for each data vector (i.e. here a person on the map) we search for the nearest node and the new nodes are calculated as a mean vector of all the nearest data vectors of the old node.

The K-means algorithm produces a reasonably good approximation of the minimal squared length problem. However, the initial placement of the nodes may affect the result considerably. Usually several initial placements are calculated and K-means algorithms are performed and the best result is chosen.

Though the K-means algorithm is designed to solve the squared length problem, it can be used also in the minimal total length problem. As a first approximation, the node placements given by K-means can be thought as solutions for both problems. In many cases the difference between the minimal total length and the total length given by the node positions chosen by the basic K-means is only a few percents. The algorithm may be improved by solving the fixed point problem (7) among persons having the same nearest node at every step, and setting the new positions of the nodes according to this. In numerical studies, the best results seem to be found when the basic K-means is run for a while at the beginning and the improved version only after that.

4.2.1 Choosing initial values

We have considered the following three possibilities for choosing the initial placement of the nodes

- Uniform distribution
- Weighted distribution
- Neural network

In a uniform distribution each component is picked randomly and evenly distributed between the minimum and maximum value.

In a weighted distribution the distribution of population is taken into consideration. We may use either marginal distributions or the common population distribution. We can for example pick randomly one person from the population and place a node at that position, which weights the distribution of nodes according to the population distribution

A self-organizing neural network [Sch92] can also be used to determine the initial placement of nodes. This method places the nodes also according to the population distribution. The neural network uses a so called neighborhood matrix to organize the nodes. The learning process resembles the K-means algorithm but instead of updating each node with only the nearest persons we update a larger neighborhood of nodes with a nearest person. In the beginning of the learning the neighborhood radius is large (the diameter of the matrix can be used) and it shrinks during the learning process so that at the end of the process only individual nodes are updated.

From these initial placement methods we may conclude that the weighted and the neural network methods work considerably better than placing the nodes uniform-randomly onto the map. The neural network produces a few percent better results than the weighted initial placement method, but in some cases it performs worse than the simple weighted method, so we may conclude that there is no significant difference using a simple weighting method than a more complex weighting method.

4.3 Interconnecting inhabitants of Finland

The data about the population distribution of Finland was given at resolution 0.5×0.5 km^2, and no information about the distribution inside the quadrants was available. We have approximated that all the inhabitants in a quadrant live at the center of that. This approximation is the same as if we had first constructed a star network with nodes placed at the center of each quadrant and then tried to connect all these nodes to the upper layer.

First we consider how to optimally place an arbitrary number of nodes in Finland. Using the equations (7) and (8), the center of mass and the minimal cable length node were determined (see figure 6). In figure 7, the locations of nodes determined by the improved K-means algorithm are shown. As expected, their distributions follows nicely that of the population of Finland.

After determining the positions of the nodes, the corresponding cable lengths were calculated. We consider two different cable lengths: weighted cable length measuring the length of a link multiplied by the number of persons served by the link, and unweighted length counting only the length of a link without considering capacities. The exact definitions are given in section 5. The minimal total weighted cable length and the corresponding average cable length as functions of the number of nodes are both straight lines on a log-log scale with the negative reciprocals of the gradient 1.60 ± 0.05

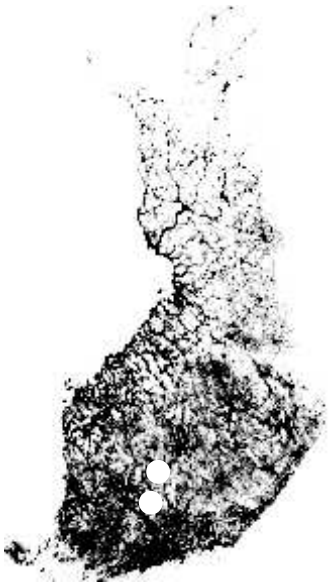

Figure 6: The Finnish population distribution with resolution 4×4 km^2, and the center of mass node (upper) and the minimal cable length node (lower).

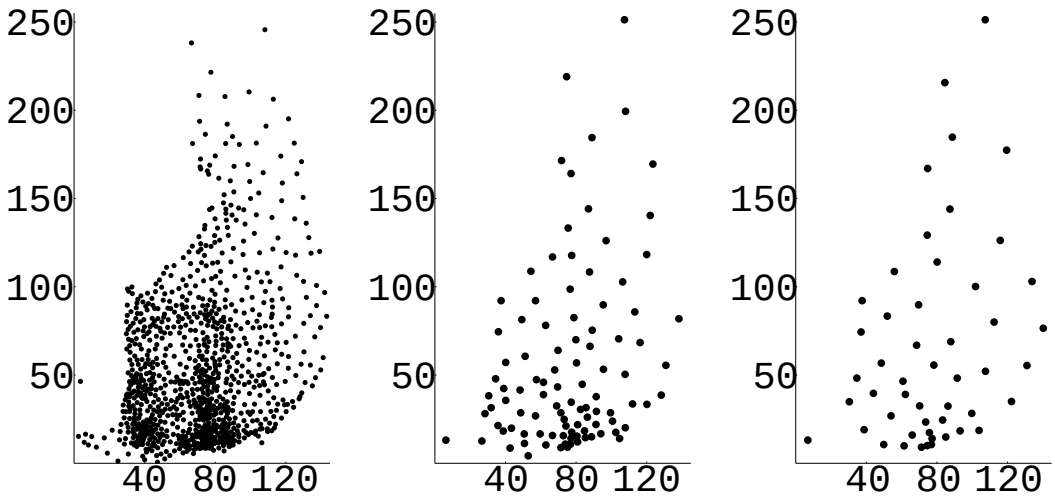

Figure 7: Placements of n nodes with $n = 1000,\ 100,\ 50$.

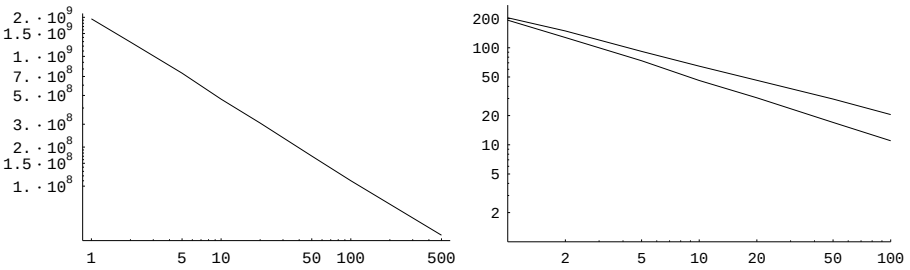

Figure 8: On the left, total weighted cable length and, on the right, the corresponding average weighted (lower) and unweighted (upper) cable lengths, all plotted against number of nodes.

(figure 8) – approximately the same as the box counting dimension of the population distribution of Finland! The average unweighted cable length for the same locations of the nodes scales too but with a different scaling parameter, the negative reciprocal of the gradient is 2.0 ± 0.05 in this case.

5 Hierarchical star networks

The next step in [App95] is to consider hierarchical star networks which are built from bottom upwards, i.e., the number and locations of the nodes on the lower layer are chosen without consideration of the upper layers of the network. With this strategy, we need to handle only one layer at time. Thus our problem is to determine the optimal number and positions of nodes on a upper layer when the location of nodes on the previous layer are given. Our reasoning and the final formula are a slightly different from those in [App95].

Consider two consecutive layers with the positions of nodes $\{t_i,\ i = 1, \ldots, N_t\}$ (lower) and $\{s_i,\ i = 1, \ldots, N_s\}$ (upper). Denote by $p(t_i)$ and $p(s_j)$ the fraction of the population served by nodes t_i and s_j, respectively. Suppose that the number of nodes on the lower layer, N_t, is given and, furthermore, that nodes t_i are distributed in a statistically self-similar way. Let L_{N_s} be the minimal weighted cable length needed to interconnect a lower layer to a upper layer via N_s nodes, i.e.,

$$L_{N_s} = \sum_{i=1}^{N_s} \sum_{t_j \in V(s_i)} p(t_j) d(s_i, t_j), \tag{9}$$

where s_i's are optimally placed, $d(s_i, t_j)$ is the length of the link between nodes s_i and t_j and $V(s_i)$ is the set of nodes whose nearest node on the upper level is s_i, i.e., the Voronoi cell of node s_i. The corresponding unweighted average cable length ℓ_{N_s} with the same s_i's as above is

$$\ell_{N_s} = \frac{1}{N_t} \sum_{i=1}^{N_s} \sum_{t_j \in V(s_i)} d(s_i, t_j). \tag{10}$$

By the numerical studies of the population distributions of Great Britain, the United States and Finland, we may assume that these both obey power laws[1]:

$$L_{N_s} \sim L_1 N_s^{-1/D(0)},$$

where $D(0)$ is the box-counting dimension of the distribution of the lower layer nodes, and

$$\ell_{N_s} \sim \ell_1 N_s^{-1/D^*},$$

where D^* can be approximated from the log-log plot.

Suppose that the cost of the network satisfies following assumptions:

- Cost of a network is divided into costs of nodes and costs of links.
- Cost of a node is a fixed cost plus a cost proportional to the number of people served:

$$C_{\text{node}} = f(s_i) = a_0 + a_1 p(s_i).$$

- Cost of a link is the product of the length and a function depending on its capacity, which again depends on the population served by it:

$$C_{\text{link}}^{i,j} = d(s_i, t_j) g(p(t_j)) = d(s_i, t_j)(b_0 + b_1 p(t_j)).$$

Using the above assumptions and definitions (9) and (10), the cost of the nodes on the upper layer is

$$C_{\text{nodes}} = \sum_{i=1}^{N_s} f(s_i) = \sum_{i=1}^{N_s} (a_0 + a_1 p(s_i)) = a_0 \ N_s + a_1$$

[1]The validity of the latter claim is tested only with the population of Finland

and the cost of the links is

$$\begin{aligned} C_{\text{links}} &= \sum_{i=1}^{N_s} \sum_{t_j \in V(s_i)} d(s_i, t_j) g(p(t_j)) \\ &= b_0 \sum_{i=1}^{N_s} \sum_{t_j \in V(s_i)} d(s_i, t_j) + b_1 \sum_{i=1}^{N_s} \sum_{t_j \in V(s_i)} p(t_j) d(s_i, t_j) \\ &= b_0 N_t \ell_{N_s} + b_1 L_{N_s}. \end{aligned}$$

By the scaling assumptions the total cost is

$$C = C_{\text{nodes}} + C_{\text{links}} \approx a_0 \ N_s + a_1 + b_0 N_t \ell_1 N_s^{-1/D^*} + b_1 L_1 N_s^{-1/D(0)} \qquad (11)$$

and the optimal number of nodes on the upper level is found by minimizing the previous equation with respect to N_s.

6 Concluding remarks

We have shown by numerical studies that the population distribution of Finland has a scale-invariant structure and that the solution of the cable length problem obeys a power law whose exponent depends on the box-counting dimension of the population of Finland. This adds evidence to Appleby's results which indicate that this kind of relationship may be quite a general one. Furthermore, the utilization of self-similarity in dimensioning problems seems to be very promising. If one was able to construct a large toolbox of easily applicable thumb rules for dimensioning hierarchical networks, it would have great potential in higher level network planing problems.

Only an introductory treatment of the subject was presented in this report. Possible future's tasks could be, e.g,

- to study effects of limited node capacities
- to test numerically the accuracy of the cost estimate (11)
- to consider more complex cost functions
- to apply similar analysis to other network topologies than hierarchical star network.

References

[Adj97] I. Adjali. Multifractal analysis of population distributions. In *Fractals in engineering, June 25–26, 1997, Arcachon, France*. INRIA, 1997.

[App94] S. Appleby. Fractal telecommunications networks. *British Telecom Technol. J.*, 12(2):19–29, 1994.

[App95] S. Appleby. Estimating the cost of a telecommunications network using the fractal structure of the human population distribution. *IEE proceedings: Communications*, 142(3):172–178, 1995.

[App96] S. Appleby. Multifractal characterization of the distribution pattern of the human population. *Geographic analysis*, 28(2):147–160, 1996.

[Fal90] K. Falconer. *Fractal geometry.* John Wiley & Sons, Chichester, 1990.

[Man88] B. B. Mandelbrot. An introduction to multifractal distributions. In *Fluctuations and pattern formation*. Kluwer Academic, Dordrecht, 1988.

[Rie95] R Riedi. An improved multifractal formalism and self-affine measures. *J. Math. Anal. Appl.*, 189:462–490, 1995.

[Sch92] R. Schalkoff. *Pattern recognition: statistical, structural and neural approaches*. John Wiley & Sons, New York, 1992.