Fuzzy Approach to Optimal Placement of Health Centers

Juan Carlos Figueroa Garcia, Carlos Franco, and Vladik Kreinovich

Abstract In countries with socialized medicine, it is important to decide how to distribute limited medical resources – in particular, where to place health centers. In this paper, we formulate and solve the corresponding constraint optimization problem. Once the locations are selected, it is necessary to decide which regions are served by each center. Traditionally, this decision is crisp, in the sense that each location is assigned to a single health center. We show that the medical service can be made more efficient if we allow fuzzy assignments, when some locations can be potentially served by two (or more) neighboring health centers.

1 Formulation of the Problem

Need for health centers. Many countries in the world have socialized medicine – in this sense, US is one of the few exceptions. In countries with socialized medicine, it is important to decide how to distribute the limited resources (and resources are always limited), so as to best serve the population.

In some case, all the patient needs is a regular general doctor. However, in many other cases, the patient also needs to undergo some tests – blood test, X-ray, etc., he/she may need to see a specialist, etc. From this viewpoint, it is more convenient

---

Juan Carlos Figueroa Garcia
Department of Industrial Engineering, Universidad Distrital, Bogotá D.C, Colombia
e-mail: filthed@gmail.com

Carlos Franco
School of Management, Universidad del Rosario, Bogotá, Colombia
e-mail: carlosa.franco@urosario.edu.co

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, El Paso, Texas 79968, USA
e-mail: vladik@utep.edu

for the patients if all the need medical professionals are placed at a single location. This is the main idea behind health centers.

Where to place health centers? Once we decided to place the doctors in health centers, the question is: where are the best locations for these centers? And, once we find these locations, what is the best way to assign each patient to one of these centers?

These are the problems that were raised in our previous paper [2]. These are the problems that we deal with in this paper as well.

2 Where to Place the Health Centers: General Analysis

What we know. Let X denote the area that we want to serve with health centers. Let $\rho(x)$ denote the population density at geographic location x, i.e., the number of people per unit area.

Once we know the population density, we can compute the overall number of people in any given area A as the integral $\int_A \rho(x)\,dx$. Let $P = \int_X \rho(x)\,dx$ denote the overall number of people in our area X.

What we want. Based on knowing the population density, we need to decide how many health centers to place at different locations. Let $h(x)$ denote the number of health centers per unit area in the vicinity of a geographical location x.

Once we determine this density $h(x)$, we can compute the overall number of health centers in any given area A as the integral $\int_A h(x)\,dx$.

Main limitation. Our resources are limited: we can only build so many health centers. Let N denote the overall number of health centers that we can build. This limitation means that

$$\int_X h(x)\,dx = N. \tag{1}$$

Objective function: informal description. In the ideal world, every patient should be immediately seen by a doctor. In reality, it takes some time for a patient to reach the nearest health center. The smaller this time, the better.

Thus, a reasonable objective function that gauges the quality of health center placements is the average time that it takes for a patient to reach the nearest health center.

Let us describe this objective function in precise terms.

Objective function: towards a formal description. The time $t(x)$ that it takes for a patient at location x to reach the nearest health center can be computed as

$$t(x) = \frac{d(x)}{v(x)}, \tag{2}$$

where:

- $d(x)$ is the distance from location x to the nearest health center, and
- $v(x)$ is the average transportation speed in the vicinity of the location x.

The speed $v(x)$ is usually smaller in the city center, slightly larger in the suburbs, and even larger outside the city limits.

The maximum distance $m(x)$ that it takes for points in the vicinity of a location x to reach a health center corresponds to the case when the location is at the edge of the zone allocated to this center, i.e., at the edge of a disk of radius $m(x)$ served by this center. In this circle, there is exactly one health center.

On the other hand, based on the density $h(x)$ of health centers, we can estimate the number of health centers in this disk area as $\int h(x)\,dx \approx h(x)\cdot(\pi\cdot m(x)^2)$. From the condition that this value is 1 (meaning that there is only one health center in this disk area), we conclude that $h(x)\cdot(\pi\cdot m(x)^2) = 1$, i.e., that

$$m(x) = \frac{1}{\sqrt{\pi\cdot h(x)}}. \tag{3}$$

What is the average distance $d(x)$ from a center of the disk of radius $m(x)$ to a point on this disk? For all the points at distance r from the center, this distance is r, and the area of the small vicinity of this disk is $2\pi\cdot t\,dr$. Thus, the average distance can be computed as

$$\frac{1}{\pi\cdot(m(x))^2}\cdot\int_0^{m(x)} r\cdot(2\pi\cdot r\,dr) = \frac{1}{\pi\cdot(m(x))^2}\cdot\frac{2}{3}\cdot\pi\cdot(m(x))^2 = \frac{2}{3}\cdot m(x). \tag{4}$$

Substituting the expression (3) into this formula, we conclude that

$$d(x) = \frac{2}{3\cdot\sqrt{\pi}}\cdot\frac{1}{\sqrt{h(x)}}. \tag{5}$$

Thus, from the formula (2), we conclude that

$$t(x) = \frac{d(x)}{v(x)} = \frac{2}{3\cdot\sqrt{\pi}}\cdot\frac{1}{\sqrt{h(x)}\cdot v(x)}. \tag{6}$$

This is the time that it takes for each patient to reach the health center. The average time that it takes all the patients to reach the health center can be then computed as

$$\frac{1}{P}\cdot\int_X \rho(x)\cdot t(x)\,dx = \frac{2}{3\cdot\sqrt{\pi}\cdot P}\cdot\int_X \frac{\rho(x)}{\sqrt{h(x)}\cdot v(x)}. \tag{7}$$

Now, we are ready to formulate the problem in precise terms.

Exact formulation of the problem. We know the functions $\rho(x)$ and $v(x)$. Based on this knowledge, we need to find the function $h(x)$ that minimizes the objective function (7) under the constraint (1).

Solving the corresponding problem. Multiplying all the value of the objective function by the same constant does not change which value is larger and which is smaller. Thus, minimizing the function (7) is equivalent to minimizing a simpler expression

$$\int_X \frac{\rho(x)}{\sqrt{h(x)}\cdot v(x)}. \quad (8)$$

To solve the problem of minimizing the expression (8) under the constraint (1), we can use the Lagrange multiplies method, according to which the above constraint optimization problem is equivalent, for an appropriate λ, to the unconstrained optimization problem of minimizing the expression

$$\int_X \frac{\rho(x)}{\sqrt{h(x)}\cdot v(x)} + \lambda\cdot\left(\int_X h(x)\,dx - N\right). \quad (9)$$

Differentiating this expression with respect to the unknown $h(x)$ and equating the derivative to 0, we conclude that

$$-\frac{1}{2}\cdot\frac{\rho(x)}{(h(x))^{2/3}\cdot v(x)} + \lambda = 0, \quad (10)$$

i.e., that

$$h(x) = c\cdot\left(\frac{\rho(x)}{v(x)}\right)^{2/3}, \quad (11)$$

for some constant c. This constant can be found if we substitute the expression (11) into the constraint (1). Then, we get the following solution.

Optimal solution.

$$h(x) = N\cdot\frac{\left(\frac{\rho(x)}{v(x)}\right)^{2/3}}{\int_X \left(\frac{\rho(y)}{v(y)}\right)^{2/3} dy}. \quad (12)$$

Discussion. The density of health centers is proportional to the population density raised to the power 2/3. Thus, in the regions with higher population density $\rho(x)$, we place more health centers – but the number of health centers grows slower than the population density.

How many doctors are needed in each health center. The number of medical personnel $M(x)$ needed for each health center is proportional to the number of people $N(x)$ served by each center:

$$M(x) = m_0\cdot N(x). \quad (13)$$

The coefficient m_0 can be obtained if we know the overall number M of medical professionals, since for the whole population, the formula (13) implies that $M = m_0\cdot P$ and thus, that $m_0 = \frac{M}{P}$ and hence,

$$M(x) = \frac{M}{P} \cdot N(x). \tag{14}$$

The number of people $N(x)$ served by a health center can be obtained by multiplying the population density $\rho(x)$ in the vicinity of a given location x by the area $1/h(x)$ covered by the center; thus,

$$M(x) = \frac{M}{P} \cdot \frac{\rho(x)}{h(x)}. \tag{15}$$

In view of the formula (12), we get

$$M(x) = \frac{M}{P \cdot N} \cdot \left(\int_X \left(\frac{\rho(y)}{v(y)} \right)^{2/3} dy \right) \cdot (\rho(x))^{1/3} \cdot (v(x))^{2/3}. \tag{16}$$

3 Where to Actually Place the Health Centers?

Formulation of the problem. The above formulas describe how many health centers to place in the vicinity of each location x. But where exactly should we place them?

We need to find 2-D locations $p_1, \ldots, p_N$ so that the average distance from each place x to this location be the smallest possible. For each location x, let us denote, by $i(x)$, the number of the health center that will be associated with this location. Then, the distance from each location x to the corresponding health center is $d(x, p_{i(x)})$, and the corresponding travel time is equal to the ratio $\dfrac{d(x, p_{i(x)})}{v(x)}$. The average distance can be therefore computed as

$$\int \rho(x) \cdot \frac{d(x, p_{i(x)})}{v(x)} \, dx, \tag{17}$$

or, if we take into account the discrete character of the information, as the sum

$$\sum_x \rho(x) \cdot \frac{d(x, p_{i(x)})}{v(x)}. \tag{18}$$

We need to find the values $p_1, \ldots, p_N$ and the value $i(x)$ (for all $x \in X$) that minimize the expression (18).

Towards an algorithm. It is difficult to immediately minimize the objective function (18) with respect to all the unknowns, so a natural idea is to minimize it iteratively. Namely, we start with some location of the centers. Then:

- First, we fix the locations p_i of the health centers and find the corresponding assignments $i(x)$. For each spatial location x, minimizing the expression (18)

with respect to $i(x)$ means minimizing the distance $d(x,i(x))$, i.e., finding the health center which is the closest to this location.
- Then, we fix the assignments $i(x)$ and find the locations p_i that, for these assignments, minimize the expression (19). For each health center i, this is equivalent to finding the new location p_i for which the average distance

$$\sum_{x:i(x)=i} \rho(x) \cdot \frac{d(x,p_i)}{v(x)} \tag{19}$$

is the smallest possible. This can be done, e.g., by gradient descent.

Then, we repeat the procedure again and again until the process converges, i.e., until the distance between the locations of each health center on two consequent iteration is smaller than some pre-determined small value ε.

Comment. This is similar to the standard algorithm for computing fuzzy clusters (see, e.g., [1]), where we iteratively:

- first, assign each point to clusters depending on this point's distance to the cluster centers, and
- then find the new centers which are, on average, closest to all the points assigned to the corresponding cluster.

Resulting algorithm. We first randomly place the centers in accordance with the center density $h(x)$ (that we computed in the previous section). Then, we iteratively do the following:

- First, for each spatial location x, we find the closest health center; we will denote the index of this health center by $i(x)$.
- Then, for each i from 1 to N, we find a new location p_i for which the average distance (19) is the smallest possible. This is done, e.g., by gradient descent.

Then, we repeat the procedure again and again until the process converges, i.e., until the distance between the locations p_i and p'_i of each health center on two consequent iteration is smaller than some pre-determined small value ε, i.e., until $d(p_i, p'_i) \le \varepsilon$ for all i.

4 Need for a Fuzzy Approach

Discussion. In the above description, we assumed that each location is assigned to exactly one health center. This assignment was based on the simplified assumption that the travel time (and waiting time) depends only on the distance from the location to the health center (and on the allowed traffic speed in the vicinity of this location).

In reality, as everyone who lives in a big city knows, travel time can change drastically. Sometimes there are traffic jams, sometimes there are accidents. Also, we only took into account travel time, but – again as everyone who ever went to a doctor knows – there is also waiting time.

Idea. From this viewpoint, if we have a patient who is slightly closer to one health venter than to the other, it does not make sense to assign this patient always to the nearest health center: maybe there is a long waiting time in this nearest health center, but there is no waiting time in the other health center – as a result of which the patient will be served faster if he or she goes to this second health center this time.

In other words, instead of assigning each patient to a single health center, it is beneficial to make a "fuzzy" allocation: namely, to allow the patient to go to any health center in the nearest vicinity – namely, to the one for which the travel time + waiting time will be the smallest. There are many apps already for predicting travel time, there are similar apps for predicting the waiting time, and nowadays, when most medical records are electronic, it is not a problem to access the records from each of the health centers.

Acknowledgements This work was supported in part by the US National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).

References

1. J. C. Bezdek, *A Primer on Cluster Analysis: 4 Basic Methods that (Usually) Work*, First Edition Design Publishing, 2017.
2. C. Franco, E. Ramiro López-Santana, and J. C. Figueroa-García, "A mathematical model under uncertainty for optimizing medicine logistics in hospitals", *Proceedings of the 5th Workshop on Engineering Applications WEA'2018*, Medelln, Colombia, October 17–19, 2018, Part II.