No.432

Efficiencies of a sequential algorithm and an intuitive selection method in a planar location problem

by
Takeshi KOSHIZUKA

June 1990

EFFICIENCIES OF A SEQUENTIAL ALGORITHM AND AN INTUITIVE SELECTION METHOD IN A PLANAR LOCATION PROBLEM

National Contribution of Japan

Takeshi KOSHIZUKA
University of Tsukuba, Japan

Abstract: We consider a problem of determining the locations $x_1, x_2, ..., x_n$ of n public facilities in such a way that the total sum of distances F between all people and their nearest facilities be minimized. Using the geographical optimization method, we can get the locations $x_1, x_2, ..., x_n$ which gives the minimum value of F. On the other hand, we can obtain a sequential minimum location $x_i (i = k, ..., n)$ under the condition that the locations $x_1, x_2, ..., x_{i-1}$ are already determined. In the present paper, we discuss the efficiency of this sequential method relative to the minimum solution and also the efficiency of an intuitive selection method which is similar to the sequential method.

KEYWORDS
geographical optimization problem, Voronoi diagram, greedy algorithm, intuitive selection method, location of facilities

A SHORT RUNNING TITLE
Efficiencies of sequential and intuitive methods

INTRODUCTION

Consider a planar region where inhabitants are distributed and public facilities of one kind are located. If each inhabitant goes to his nearest facility, we can compute the total sum of distances between all the inhabitants and their nearest facilities. Adopting this total sum as the objective function to be minimized, we formulate a facility location problem.

Iri *et al.*(1984) got the 'optimal' numerical solution of this problem, using the geographical optimization method which is based on the fast Voronoi-diagram algorithm. Although geographical optimizations are big computational problems, this method has an outstanding ability to compute real world problems within a practical time. The global minimum of this problem is hopelessly difficult to get in the general case: if the inhabitant distribution is arbitrarily given. But starting from a reasonable initial condition, we can obtain a local minimum which is satisfactory in most practical situations. The quotation marks of 'optimal' indicate the circumstances mentioned above.

In real urban facility planning, however, all the locations of facilities are not determined at one time. Usually we must determine the locations of facilities one after another taking into consideration the inhabitants' distribution, the land availability and some other things. For this reason, comparing with the 'optimal' solution obtained by Iri *et al.*(1984), we discuss a sequential determining method to introduce a threshold value which estimates the efficiency of this sequential method and also discuss an intuitive selection method.

GEOGRAPHICAL OPTIMIZATION PROBLEM

Suppose there are n public facilities of a kind, for example post offices, at location points are $x_1, x_2, ..., x_n$ in a planar region D. The mean value F of the distances between all the people in D and their nearest facilities is given by

$$F(x_1, x_2, ..., x_n) = \frac{1}{P} \int_D \min_i \|x_i - x\| \; \mu(x) \; \mathrm{d}x, \tag{1}$$

where $\mu(x)$ represents the population density at a point x in D and P is the total population in D, namely

$$P = \int_D \mu(x) \; \mathrm{d}x.$$

In evaluating the right hand side of equation (1), it may be required that at every point we select the nearest among the n facilities. But with the help of the Voronoi diagram for the given location of facilities, we need not select at every point. Let $V(x_i)$ be the Voronoi region belonging to x_i, then the point x_i is the nearest facility for each point in $V(x_i)$. Thus we can rewrite equation (1) to

$$F(x_1, x_2, ..., x_n) = \frac{1}{P} \sum_{i=1}^{n} \int_{V(x_i)} \|x_i - x\| \; \mu(x) \; \mathrm{d}x. \tag{2}$$

If the density $\mu(x)$ is uniform in D, equation (2) is simplified as

$$F(x_1, x_2, ..., x_n) = \frac{1}{S} \sum_{i=1}^{n} \int_{V(x_i)} \|x_i - x\| \; \mathrm{d}x, \tag{3}$$

where S is the area of D. In this case, starting from a random distribution of facilities, the optimal solution forms almost the regular hexagonal system except for the boundary of D.

The distortion from regular hexagon reduces with increasing number of the facilities (Iri *et al.*, 1984).

A SEQUENTIAL METHOD

First we consider a greedy algorithm determining the location point x_i of ith facility in such a manner that the mean value of the distances F is minimized under the condition of predetermined $x_1, x_2, ..., x_{i-1}$. But this sequential optimal solution x_i cannot be obtained easily because of the nonconvexity of the function F which is evident from the fact that F has many local minima. In spite of this nonconvexity, the geographical optimization method succeeds based on the mobility of all the facilities.

For the above reason we suggest a simpler sequential method which locates the ith facility at the center of the largest empty circle in D under the condition that $x_1, x_2, ..., x_{i-1}$ are already distributed. This center is known to coincide with one of the Voronoi points which are the intersections of Voronoi edges or with one of the intersection points between Voronoi edges and the boundary of D when D is convex.

The center of the largest empty circle is not necessarily the minimum point of F for the i th facility location. Nevertheless, the center is expected to be near the minimum point or the second minimum or the third except on the boundary of D.

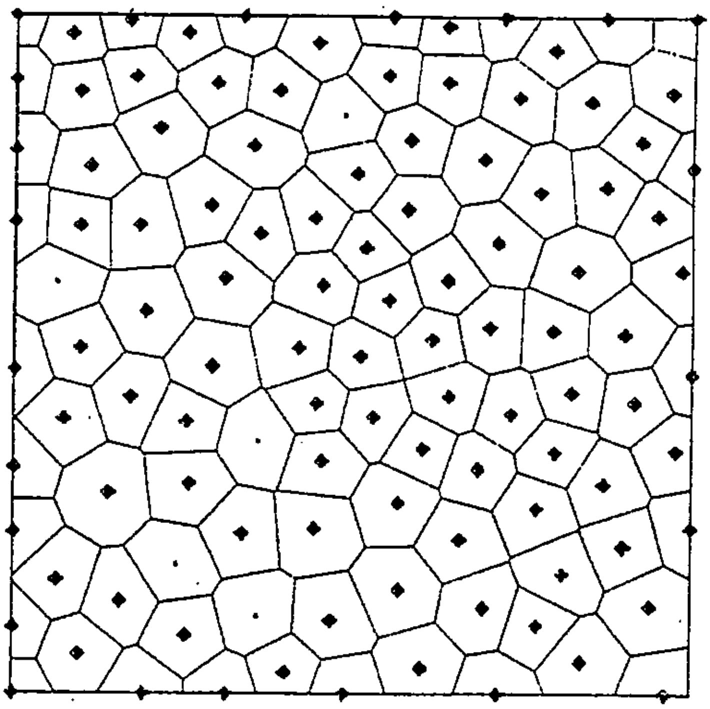

Fig.1. Point pattern with its Voronoi diagram by the sequential method
• : first 5 facilities, ◆ : the others

For simplicity in the case of a uniform population density, Fig.1 shows one of the numerical experiments of this method in the region D of a 10×10 square where 95 facilities are sequentially located from a random distribution of 5 facilities. In these experiments the facilities on the boundary of D are counted as 1/2 for each because of the facility density in D. By the distance density algorithm applied to Voronoi diagram (Koshizuka and Ohsawa, 1986), we can calculate the mean value F to find that $F \approx 0.3967$ in the case of Fig.1.

As discussed previously, the solution of the geographical optimization method is almost the regular hexagonal system. Hence instead of the numerical optimal value we found from calculation of the regular system that

$$F = \frac{\sqrt{2\sqrt{3}}}{18\sqrt{\rho}}(2 + 3\log\sqrt{3}) \approx \frac{0.3772}{\sqrt{\rho}}, \tag{4}$$

if the facility density is ρ. As a result, we get numerically the efficiency ψ of this sequential method

$$\psi \approx 0.3772/0.3967 \approx 0.95, \tag{5}$$

because the density of Fig.1 is equal to 1. In several cases starting from different random 5 facility distributions, the efficiencies ψ have stable values ranging from 0.94 to 0.95.

Let a set of n facility locations be $W = \{x_1, x_2, ..., x_n\}$ and a subset of W be V, namely $V \subseteq W$. The effect of putting another facility at a location x, under the condition that the locations $x_1, x_2, ..., x_n$ are already determined, is given by $F(W) - F(W \cup x)$ which means the reduction of the mean distance between all the people and their nearest facilities. For $W \supseteq V$, the effect of x under the condition of W is less than that of x under the condition of V:

$$F(W) - F(W \cup x) \ \leq \ F(V) - F(V \cup x),$$

because we consider the nearest facility distance at every point. The objective function F, therefore, is supermodular: namely $-F$ is submodular. In general, Nemhauser *et al.*(1978) proved that for submodular functions the efficiency of any greedy algorithm satisfies some inequality. The sequential method discussed in this section, though not greedy algorithm strictly, may have much higher efficiencies than the lower bound of the inequality.

AN ESTIMATION OF EFFICIENCIES

Distance Density Function

Figure 1 indicates that the pattern of the facility points is rather regular. This regularity causes the sequential method the high efficiency of (5). But strictly, the point pattern of Fig.1 is neither completely regular pattern nor random pattern, standing at 'intermediate' between these two. Thus to discuss the efficiencies of the sequential method, we will formulate a model which explains the 'intermediate' pattern.

From now on, assuming a uniform population distribution we will start from equation (3). Here we consider the distance r between a point and its nearest facility. Let $G(r)$ be the distribution of the population ratio whose distance to its nearest facility is less than r, then we can get

$$G(r) = \frac{1}{S} \sum_{i=1}^{n} \int_{V(x_i),\ \|x - x_i\| < r} \mathrm{d}x. \tag{6}$$

In each Voronoi region $V(x_i)$, if $S_i(r)$ denotes the area within r from the facility point x_i, equation (6) is rewritten in the form

$$G(r) = \frac{1}{S} \sum_{i=1}^{n} S_i(r). \tag{7}$$

By differentiating $G(r)$ with respect to r, we get the distance density function $g(r) = G'(r) = 1/S \sum_{i=1}^{n} S_i'(r)$. The differential $S_i'(r)$ means apparently the length $L_i(r)$ of the intersection of the Voronoi region $V(x_i)$ and the circle of radius r from the point x_i (see Fig.2). The substitution of $S'(r)$ for $L_i(r)$ yields

$$g(r) = \frac{1}{S} \sum_{i=1}^{n} L_i(r). \tag{8}$$

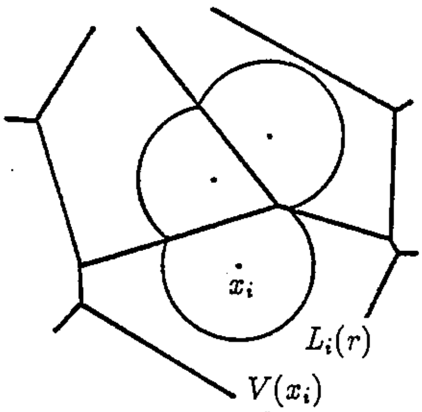

Fig.2. Voronoi region $V(x_i)$ and the length of perimeter $L_i(r)$

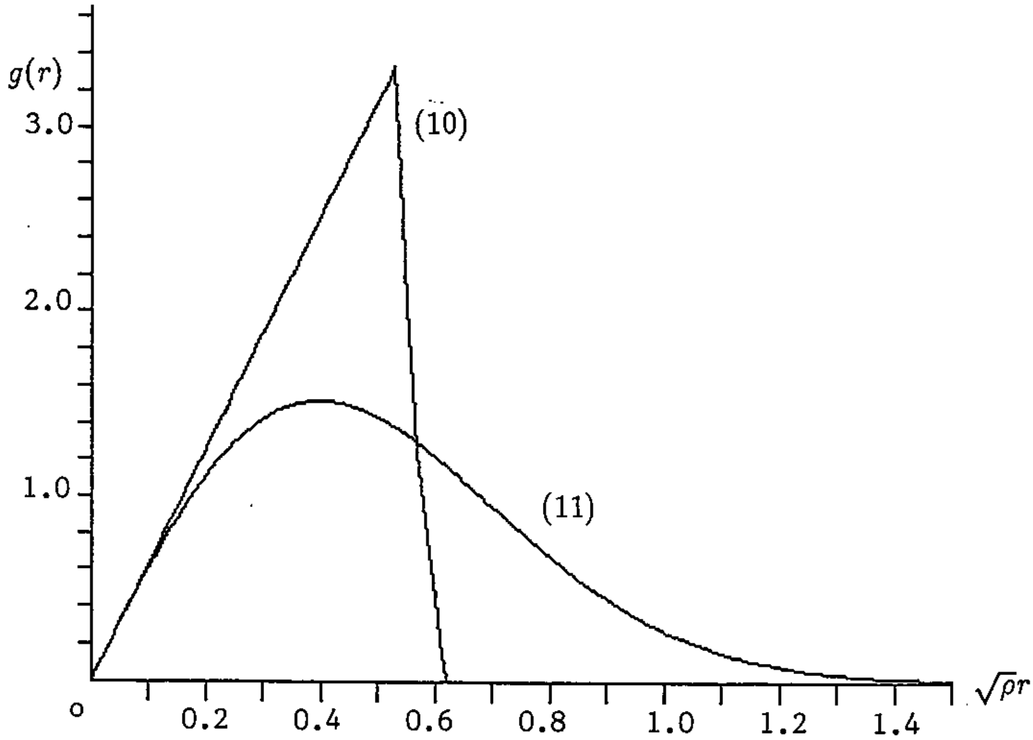

Fig.3. Equation (10): hexagonal and equation (11): random

Then the mean value F of the distances of equation (3) can be calculated from the definition

$$F = \int_0^\infty r\, g(r)\, \mathrm{d}r \tag{9}$$

based on the property of equation (8) and the algorithm (Koshizuka and Ohsawa, 1986) which is formed by computing the length $L_i(r)$ in each region $V(x_i)$. In the case of completely hexagonal system, if ρ denotes the density of facilities n/S in D, we get the distance distribution density $g_h(r)$ from (8) as follows:

$$g_h(r) = \begin{cases} 2\rho\pi r & \text{if } 0 < r \leq \sqrt{6\sqrt{3}}/(6\sqrt{\rho}), \\ 2\rho\pi r - 12\rho r \arccos\{\sqrt{6\sqrt{3}}/(6r\sqrt{\rho})\} & \text{if } \sqrt{6\sqrt{3}}/(6\sqrt{\rho}) < r < \sqrt{2\sqrt{3}}/(3\sqrt{\rho}). \end{cases} \tag{10}$$

On the other hand, in the case of uniformly random distribution, if the facility density is the same ρ as at (10), the density $g_r(r)$ is shown by

$$g_r(r) = 2\rho\pi r e^{-\rho\pi r^2} \tag{11}$$

which was derived by Clark and Evans(1954). Figure 3 shows the graphs of equations (10) and (11), suggesting $g_h'(0) = g_r'(0)$ which is proved from (10) and (11).

In Fig.3 the graph of equation (10) is linear with r at the interval $0 < r < \sqrt{6\sqrt{3}}/(6\sqrt{\rho})$. This fact means that no circles of radius r from the facilities intersect one another at the interval, in other words each circle is completely contained in each Voronoi region. In contrast to the regular pattern, the graph of equation (11) shows that the circles can intersect at the neighborhood of $r = 0$ when the facilities are distributed uniformly at random.

Modeling from the Regular to the Random by a Threshold Value

To formulate the 'intermediate' patterns such as shown in Fig.1, we introduce a threshold value r_0 in such a way that at $0 < r \leq r_0$ no circles of radius r intersect and from r_0 some of the circles begin to intersect. The efficiency of the 'intermediate' patterns depends on this threshold value r_0 which controls the regularity of the 'intermediate' patterns. Here at $r_0 < r$ we must discuss the intersecting manner of the circles which gives us the function $g(r)$.

For completely regular patterns we can formulate the intersecting manner such as equation (10). But for the 'intermediate' patterns we cannot calculate the density $g(r)$ exactly because the patterns include some randomness in the intersecting manner. For this reason we will extract a property of the intersecting manner of the random distribution.

Equation (11) is derived from

$$G(r) = 1 - e^{-\rho\pi r^2} \tag{12}$$

which is obtained from the Poisson distribution in a planar region. By the transposition of (12) we get $e^{-\rho\pi r^2} = 1 - G(r)$ and substituting this into equation (11), we have

$$G'(r) = 2\rho\pi r\{1 - G(r)\}. \tag{13}$$

Since $G'(r) = \sum_{i=1}^{n} L_i(r)/S$ and $\rho = n/S$, equation (13) is rewritten in the form

$$\frac{\sum_{i=1}^{n} L_i(r)}{2n\pi r} = 1 - G(r) \tag{14}$$

which means that the ratio between the length $\sum_{i=1}^{n} L_i(r)$ of the n r-radius circles within each Voronoi region and the total length $2n\pi r$ of the n circles is equal to the ratio of the area not covered by the n circular disks.

Let us consider that differential equation (13) indicates the intersecting manner of the distributions and the variable ρ increases with the regularity of the distributions, starting from the facility density of the random distribution. Thus at $0 < r \leq r_0$

$$G(r) = \rho\pi r^2 \tag{15}$$

is obtained because no circles intersect. At the interval $r_0 < r$ the circles intersect depending on differential equation (13) in which the variable ρ is replaced by ρ^*, because ρ shows the real facility density in equation (15).

As a result, solving differential equation (13) we get

$$G(r) = 1 + Ce^{-\rho^*\pi r^2}, \tag{16}$$

where C is an integration constant. Substituting into (16) the initial condition that $G(r_0) = \rho\pi r_0^2$ from (15), we have

$$G(r) = 1 - (1 - \rho\pi r_0^2)\, e^{-\rho^*\pi(r^2 - r_0^2)}. \tag{17}$$

And from the condition that the differential $G'(r)(= \sum_{i=1}^{n} L_i(r))$ is continuous at $r = r_0$ because each length $L_i(r)$ is continuous at $r = r_0$, we find

$$\rho^* = \frac{\rho}{1 - \rho\pi r_0^2}.$$

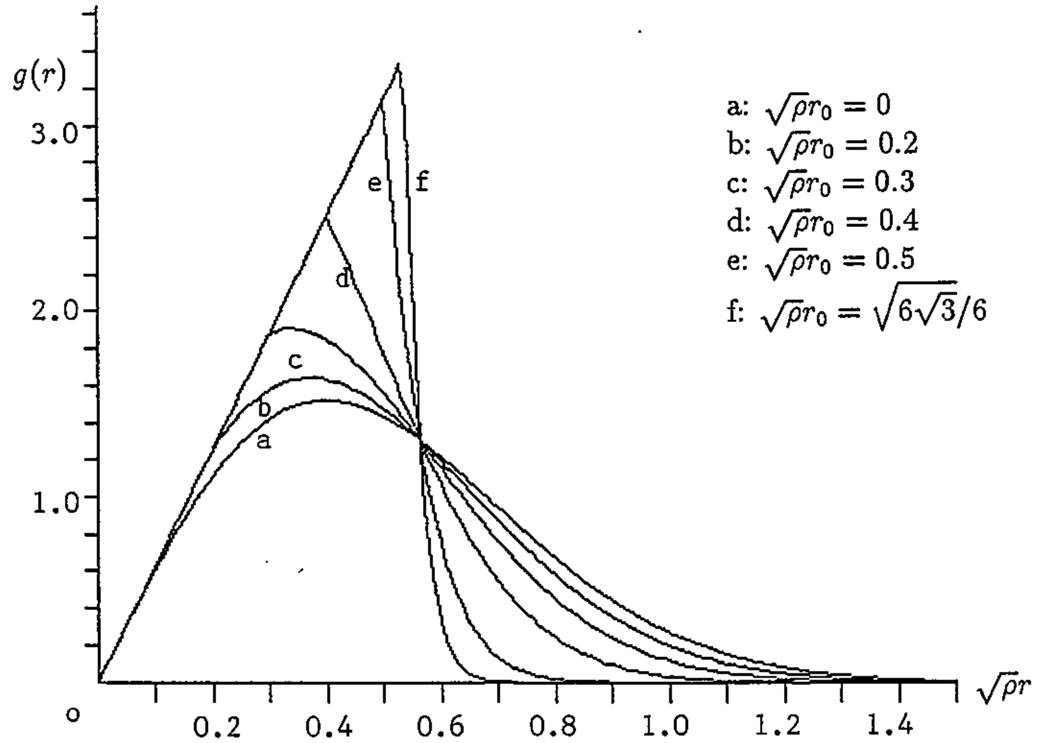

Fig.4. Demonstration from the random to the regular by a threshold value r_0 : equation (18) with r_0

Consequently, introducing the threshold value r_0 for the 'intermediate' distributions, by differentiating (15) and (17) with respect to r, we can derive the distance density function $g(r)$ as follows:

$$g(r) = \begin{cases} 2\rho\pi r & \text{if } 0 < r \leq r_0 \\ 2\rho\pi r e^{-\rho^*\pi(r^2-r_0^2)} & \text{if } r_0 < r, \end{cases} \tag{18}$$

where

$$\rho^* = \frac{\rho}{1-\rho\pi r_0^2}.$$

If $r_0 = 0$, the density function (18) agrees with that of the random pattern (11) and the distribution increases its regularity with progression of r_0 as shown in Fig.4. If $r_0 = \sqrt{6\sqrt{3}}/(6\sqrt{\rho})$, Fig.5 shows that equation (18) is almost the same as equation (10) of the completely regular hexagonal system. It is indicated, therefore, that equation (18) is a model, parametrized by the threshold value r_0, which demonstrates the point distribution patterns continuously from the random to the regular.

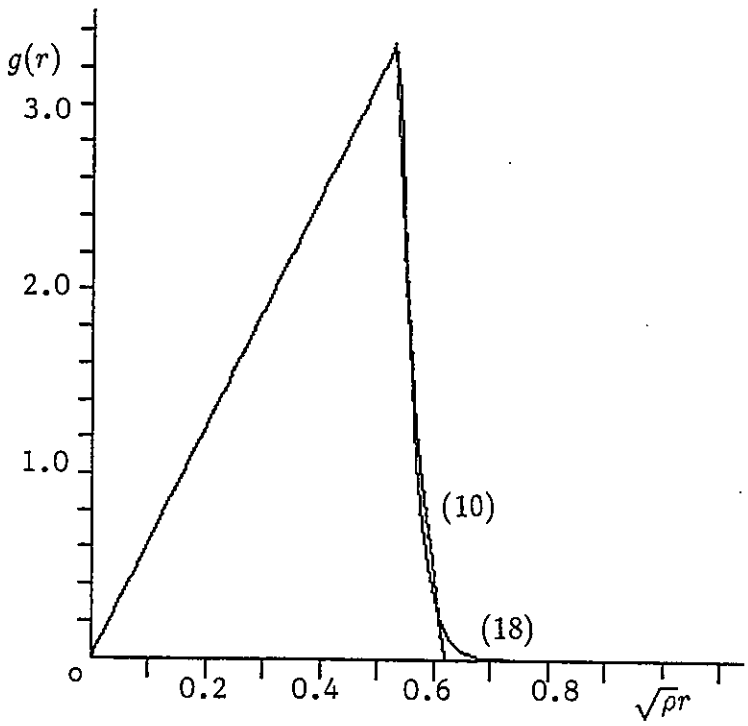

Fig.5. Equation(10): hexagonal and equation (18):$r_0 = \sqrt{6\sqrt{3}}/(6\sqrt{\rho})$

An Estimation of Efficiencies

From equations (9) and (18), the mean value F is given by

$$F = \frac{2}{3}\rho\pi r_0^3 + \frac{\rho}{2\rho^*\sqrt{\rho^*}} e^{\rho^*\pi r_0} \{1 - \frac{4}{\sqrt{\pi}} I(\sqrt{\rho^*\pi} r_0)\}, \tag{19}$$

where

$$I(t) = \int_0^t x^2 e^{-x^2} dx.$$

When $r_0 = 0$, the value of (19) equals $1/(2\sqrt{\rho})$ which is derived from equation (11) because $\rho^* = \rho$ and $I(0) = 0$. On the other hand when $r_0 = \sqrt{6\sqrt{3}}/(6\sqrt{\rho})$, the value of (19) is numerically given by $F \approx 0.3773/\sqrt{\rho}$ which coincides with $0.3772/\sqrt{\rho}$ of the completely regular hexagonal system in the order of $1/\sqrt{\rho} \times 10^{-3}$. Thus using equations (4) and (19), we get an estimate for the efficiency of the patterns as follows:

$$\hat{\psi} = \frac{\frac{\sqrt{2\sqrt{3}}}{18}(2 + 3\log\sqrt{3})}{\frac{2}{3\sqrt{\pi}}R_0^3 + \frac{(1-R_0^2)^{3/2}}{2}\, e^{\frac{R_0^2}{1-R_0^2}}\, \{1 - \frac{4}{\sqrt{\pi}}I(\sqrt{\frac{R_0^2}{1-R_0^2}})\}}, \tag{20}$$

where $R_0 = \sqrt{\rho\pi}r_0$. The equation above means that the efficiency $\hat{\psi}$ depends only on R_0 $(= \sqrt{\rho\pi}r_0)$. Hence we can estimate the efficiencies of the distributions only by R_0(namely ρ and r_0). Figure 6 demonstrates equation (20) to show that the efficiency $\hat{\psi}$ increases to 1 as the value $\sqrt{\rho}r_0$ increases to 0.5373 which is approximately the threshold value $\sqrt{6\sqrt{3}}/6$ of the completely hexagonal system.

In Fig.1 the circles begin to intersect at the nearest pair of facilities. If the distance between the nearest pair is set to the threshold value r_0, we find $r_0 = 0.395$ for Fig.1. From this value and $\rho = 1$, we have the value R_0 and by insertion of this to (20), we can estimate for the efficiency of Fig.1 as

$$\hat{\psi} \approx 0.92.$$

This estimate is a little smaller than the true value of (5), because from equation (19) we estimated the mean value as$F \approx 0.4094$ which is slightly larger than the true value 0.3967. Figure 7 is the comparison between the estimate and the true distance density function. Graph (9) shows the exact form of the function $g(r)$ for Fig.1 computed by the algorithm (Koshizuka and Ohsawa, 1986) based on equation (8). Graph (18) shows the function (18) into which $\rho = 1$ and $r_0 = .395$ are substituted.

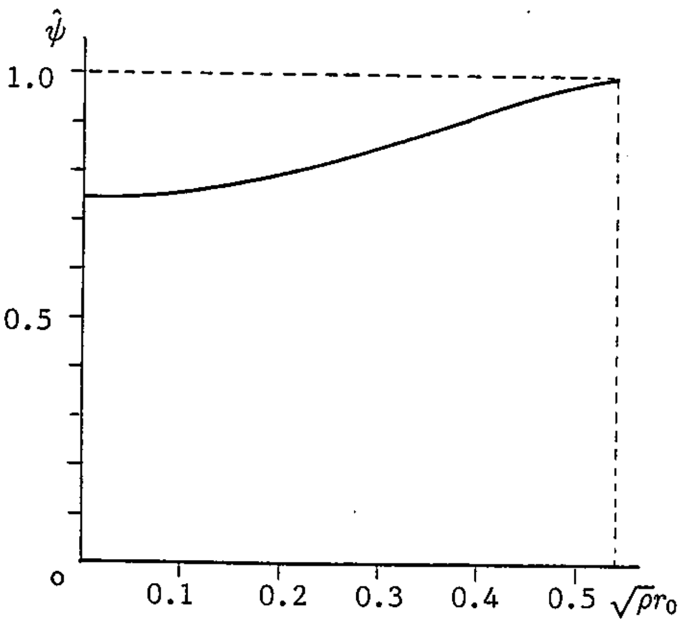

Fig.6. Efficiency with r_0: equation (20)

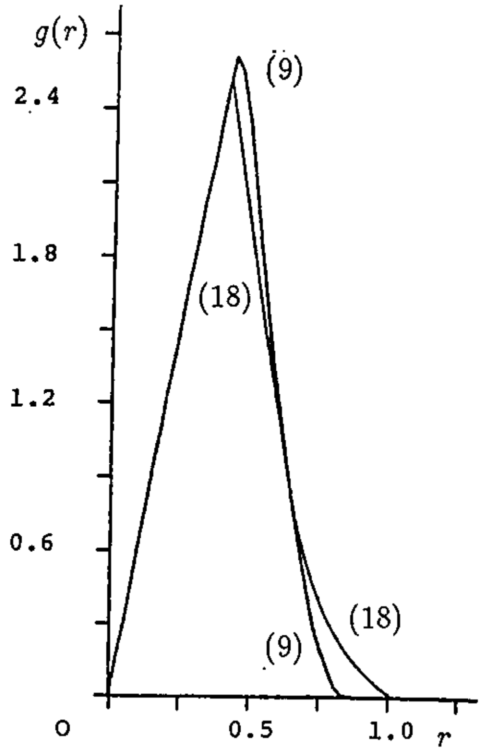

Fig.7. Equations (9) and (18)

In our model (18), the threshold value r_0 coincides with the point at which the density begins to decrease rapidly with progression of r when $\sqrt{\rho}r_0$ is somewhat large (see Fig.4). But this fact suggests that in the sequentially determined patterns as Fig.1, the rapid decrease point is slightly larger than the distance between the nearest pair of facilities (namely the threshold value in the estimation). For this reason, if the nearest facility distance is employed as the threshold value, the estimation of efficiencies is probably a little smaller than the true.

AN INTUITIVE SELECTION METHOD

In the sequential method discussed previously, we carried out computations to find the center of the largest empty circle at each stage. We can also dot the ith facility intuitively without computations, looking over the points $x_1, x_2, ..., x_{i-1}$ in the region D. In this intuitive selection method, we determine the ith point in such a way that the mean value F may be minimized, but by freehand dotting not by computations.

With its Voronoi diagram, Fig.8 shows one of the point distributions (100 points in a 10×10 square) made by the author's freehand dotting which was done intuitively and not so carefully. In the same way as at (5), we computed the mean value $F \approx 0.4038$ to get the efficiency

$$\psi \approx 0.3776/0.4038 \approx 0.93. \tag{21}$$

Thus the efficiency of this intuitive method is as high as the sequential method. This fact suggests that our eyes have a great ability in the planar problem to locate the point intuitively near the center of the largest empty circle. Even if the intuitively selected point is not near the center of the largest empty circle, this point might be near that of the second or the third.

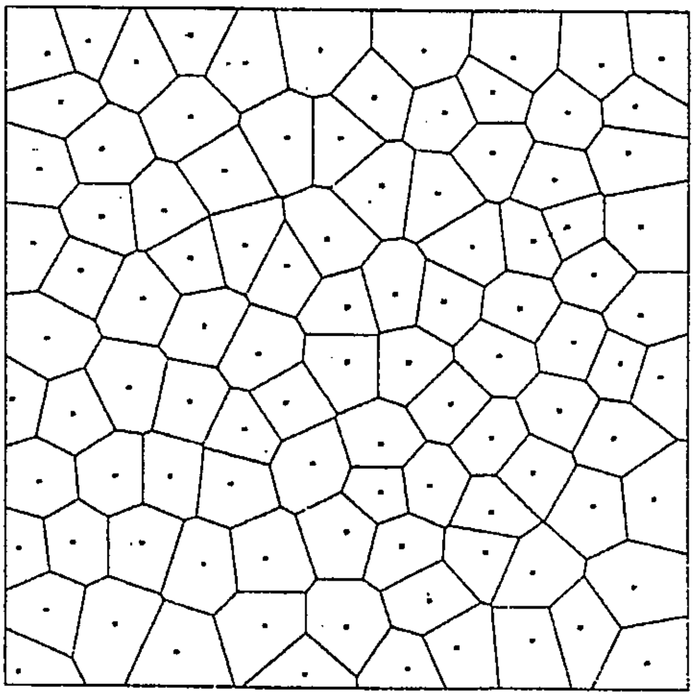

Fig.8. Freehand dotting pattern with its Voronoi diagram

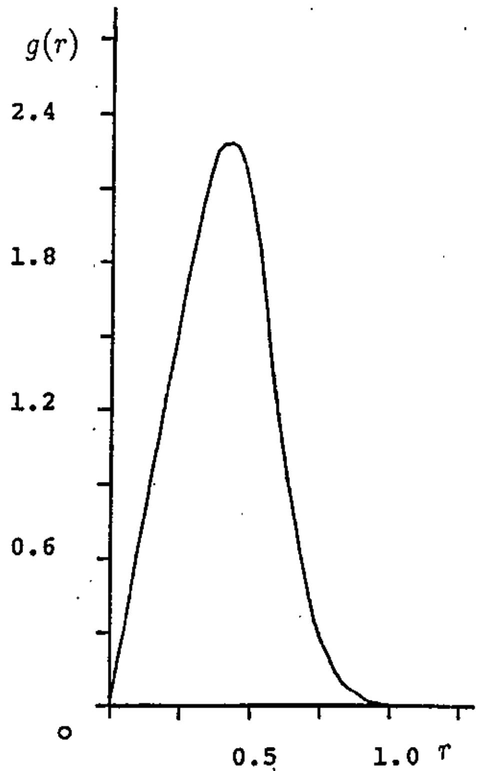

Fig.9. Density function

At a glance the pattern of Fig.8 is the same 'intermediate' pattern as that of Fig.1. But showing the distance density function of Fig.8, Fig.9 has a difference from the graph (9) of

Fig.7 which is the density function of Fig.1. In Fig.9, we have a smooth curve near the maximum point in the contrast to the graph (9) of Fig.7.

We sought for a similar pattern as Fig.8 to find an arabesque printed in a cloth wrapper, what we call 'furoshiki' in Japanese, as shown in Fig.10. Considering the tips of this arabesque as the points discussed previously, we can calculate the efficiency and the distance density function which indicate the characteristics of this arabesque as compared with the facility distributions. The density function $g(r)$ based on equation (8) is obtained as Fig.11, which gives us the value of efficiency $\psi \approx 0.92$. Figure 11 indicates that the arabesque pattern is similar to the intuitive selection pattern of Fig.9 in distance density function.

From ancient times, we Japanese have preferred the 'intermediate' patterns than regular patterns. The 'intermediate' patterns have two important features. One is high uniformity: efficiency almost the same as regular patterns, which was revealed for the first time in the present paper. The other is a sequential strong identity of pattern which suggests that if we eliminate one point from a 'intermediate' pattern or add one to that, we cannot recognize the change of pattern as a whole. On the other hand, if we get rid of one point from a regular pattern, we can find easily a defect of the regular pattern. I think we have found a beauty in the 'intermediate' patterns because of these two remarkable features. We do not prefer the perfect. We like the presence of blanks which can be filled in the future.

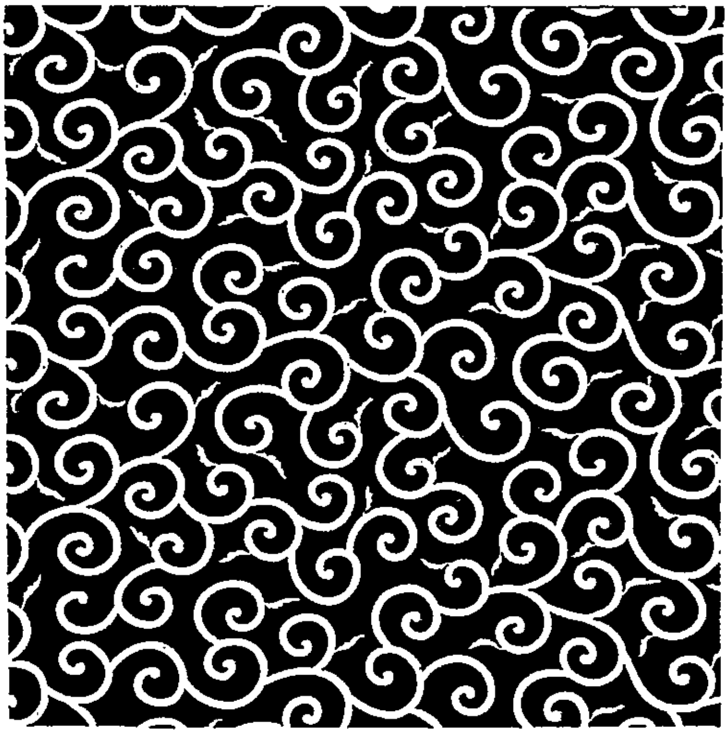

Fig.10. An arabesque printed in a cloth wrapper

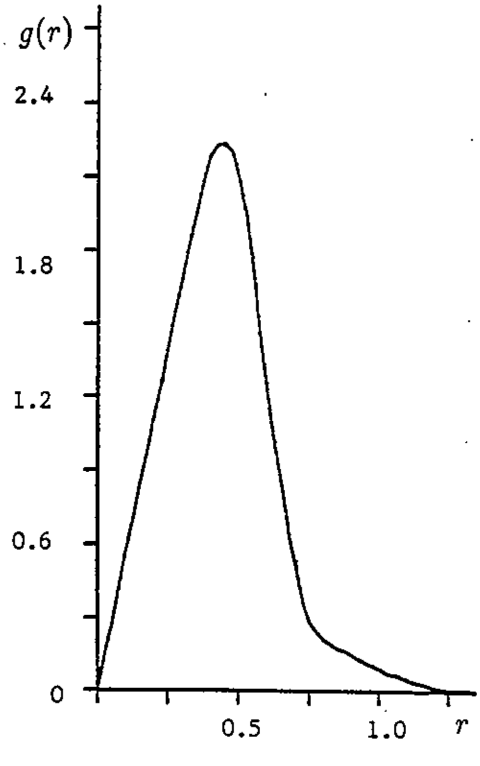

Fig.11. Density function of the arabesque

CONCLUDING REMARKS

Considering a uniform population density, we discussed the efficiency of sequential methods compared with the completely hexagonal system which is almost the same as the 'optimal' solution in finite regions. Locating a facility at the center of the largest empty circle at each stage, the sequential method provides the 'intermediate' pattern between the random and the regular. Hence this pattern causes a high efficiency relative to the 'optimal', we introduced a threshold value to formulate a model which explains the point patterns continuously from the random to the regular. Using this threshold value, we were able to estimate the efficiency of the sequential method.

On the other hand, we discussed an intuitive selection method to find a pattern with almost as high efficiency as the sequential method. These high efficiencies of the two methods might be given because this optimization problem is formulated in a planar region. In general, a facility location x_i and another facility location x_j are not independent in such a manner that the mean value of the distance between all the people and their nearest facilities be minimized. However, if the distance $d(x_i, x_j)$ between x_i and x_j is fairly large with respect to the facility density ρ: if $\sqrt{\rho}d(x_i, x_j)$ is fairly large; the locations x_i and x_j must be almost independent, namely if x_i was determined previously in the sequence and x_j is now seeking for its location, x_i has not so strong influence on the location of x_j in the manner mentioned above.

It should be noted, therefore, that this almost independence in a planar region gives us the high efficiencies of the sequential method and the intuitive selection method.

ACKNOWLEDGEMENT

The auther would like to thank Prof. K. Murota of the unversity of Tokyo, a coauther of (Iri *et al.*,1984), for valuable comments on an early draft of the present paper. He also thanks Dr. Y. Ohsawa of Kumamoto University for useful information concerning the subject.

REFERENCES

Clark, P.J. and F.C. Evans (1954). Distance to Nearest Neighbour as a Measure of Spatial Relationship in Population. *Ecology*, 35, 445-453.

Iri, M., K. Murota and T. Ohya (1984). A Fast Voronoi Diagram Algorithm with Applications to Geographical Optimization Problems. *Proceedings of the 11th IFIP Conference on System Modelling and Optimization*, 273-288.

Koshizuka, T. (1988). Relations between point patterns and the areas of the regions covered by circles of radius r from points (in Japanese). *Proceedings of the 1988 Spring Conference of the Operations Research Society of Japan*, 145-146.

Koshizuka, T. (1989). Efficiencies of an intuitive selection method in planar location problems (in Japanese). *Proceedings of the 1989 Spring Conference of the Operations Research Society of Japan*, 185-186.

Koshizuka, T. and Y. Ohsawa (1986). Derivation of the Probability Density Function of Distance and its Applications to Urban Facility Planning. *Science on Form: Proceedings of the First International Symposium for Science on Form*, 491-498.

Nemhauser, G.L., L.A. Wolsey and M.L. Fisher (1978). An Analysis of Approximations for Maximizing Submodular Set Functions-I. *Mathematical Programming*, 14, 265-294.