The Eighth International Symposium on Operations Research and Its Applications (ISORA'09)
Zhangjiajie, China, September 20–22, 2009

Optimal Location of Facilities with Limited Capacity

Takamori Ukai

Mathematical Sciences and Information Engineering Research Center,
Nanzan University, Seto, 489-0863, Japan.

Abstract At the facility which has limited capacity, users can not always be provided service, but be passed around. In the situation that the total number of servers is constant, more facilities reduce servers of individual facility. Therefore though the distance from users to facility decreases, the loss probability increases. We consider the number of facilities and their location which minimizes total distance, taking account of the loss probability which depends on their capacity and demands.

Keywords Facility Capacity, Average Distance, Loss Probability, Facility Location

1 Introduction

In this paper, we deal with a service system in which demands must travel to the facilities to receive service. More than one servers are assigned to each facility and attend customers. Customers select a facility in their own preferred order. If all servers assigned to the facility are simultaneously busy, customers cannot be served as soon as they arrive there, and they either enter in queue or shift to next preferred facility. In emergency medical service system, time to start treatment is significant for its effectiveness. Thus, calls for service are assigned to as near facilitiy as possible in which there exists at least one available server.

Progress of medical technology and aging population needs more cost for medical treatment. Against that, it is difficult to immediately increase medical expense and a number of medical service workers. Although it is desired that all emergency demands can always be responded, to provide such a infrastructure is impossible from view of both cost and manpower. Therefore, it is wished to manage limited resources effectively.

From basic results of queuing theory, more servers decrease the loss probability drastically. Under constraints that a number of server is constant, servers assigned to each facility are inverse proportion to facilites. That means that services can be provided effectively at individual facilities by decreasing facilities and utilizing the resources intensively. At the same time, the situation that all servers in one facility are unavailable is hard to occur. On the other hand, less facilities force users to travel more distance to facility. Because it takes more times to move to facility, it may occur to influence following treatment.

Carter, Chaiken and Ignall[1] analyzed a case of two service units and two fixed home location, assuming a simplified cooperation between the two service units. They incorporated the probabilistic and interdistrict behaviors of system into their model. Larson[2, 3] utilized the hypercube queing model and extended number of service unit for up to more than ten. He solved accompanied steady-state equations numerically on a computer. These models are useful to obtain perfomance measures of system. However, there is no numerical or analytical insight about appropriate number and/or size of facilities.

In this paper, we assume continuous region with a certain number of servers, and formulate stochastic model in which demands distribute uniformly randomly on the region-wide. We investigate basic characteristics of optimal dividing of limited resources and location of facilities while taking account of both efficiency of medical service which depends on resources assigned to facilities and burden of distance that users should move to facility.

2 State Probability of Number of Users

2.1 Total Number of Facility Users in the City

To analyze focusing on both efficiency of services and movement of users, let us consider the total number of users in a limited region setting their location aside.

Consider a city region where M medical servers exist. There are h facilities, that is hospitals, and each of them has servers of equal number. Let m be number of servers per facility, that is $M = mh$. In addition, we assume λ as the rate of demand occurrence per unit time in the whole city, that is arrival rate from the point of view of hospitals, and assume μ as the service rate per server per unit time.

Then, suppose vector $\boldsymbol{x}$ which has the dimension of mh corresponding to the number of servers,

$$\boldsymbol{x} = (x_1, \ldots, x_{mh})^T, x_i \in \{0, 1\}. \tag{1}$$

x_i represents the state of server i. It is one if the server is busy, and zero otherwise. Let $\lambda_i(\boldsymbol{x})$ be arrival rate to server i at state $\boldsymbol{x}$. We assume

$$\sum_{i=1}^{N} \lambda_i(\boldsymbol{x}) = \lambda. \tag{2}$$

For every state $\boldsymbol{x}$, we obtain following equations:

$$\sum_i \{\mu(1 - x_i) + \lambda_i(\boldsymbol{x}[i])x_i\} p(\boldsymbol{x}[i]) = \sum_i \{\mu x_i + \lambda_i(\boldsymbol{x})(1 - x_i)\} p(\boldsymbol{x}) \tag{3}$$

where $\boldsymbol{x}[i]$ is the vector which differs from $\boldsymbol{x}$ in i-th element and $p(\boldsymbol{x})$ is the probability of state $\boldsymbol{x}$.

Summing up equation (3) of set $X_k = \{\boldsymbol{x} | \boldsymbol{x}^2 = k\}$ that is the states of which number of users is k, we obtain

$$(k+1)\mu p(X_{k+1}) + \lambda p(X_{k-1}) = k\mu p(X_k) + \lambda p(X_k) \tag{4}$$

where $\sum_{\boldsymbol{x} \in X_k} p(\boldsymbol{x}) = p(X_k)$. Above equation shows that state of total users in the whole city is expressed queuing system of which type is $M/M/mh$.

Assuming the whole city as one facility, we can consider it as the system which has mh servers. Thus, if the total number of servers in the city is constant, the loss probability that they can't respond at the facilities in that region is constant.

2.2 Number of Individual Facility Users

In this section, we consider the number of users in individual facility. Supporse vector $\boldsymbol{u}$ which has the dimension of h corresponding to the number of facilities,

$$\boldsymbol{u} = (u_1, \ldots, u_h)^T, u_j = \sum_{i=(j-1)m+1}^{jm} x_i. \tag{5}$$

Namely, $\boldsymbol{u}$ shows the state that the numbers of users is $u_1, \ldots, u_h$. Let $X_{\boldsymbol{u}} = \{\boldsymbol{x} | u_j = \sum_{i=(j-1)m+1}^{jm} x_i (j = 1, \ldots h)\}$ be the set of states of which the numbers of users in individual facilities is $\boldsymbol{u}$, and $\lambda'_j(\boldsymbol{u})$ be the arrival rate to facility j at the state $\boldsymbol{u}$. We assume

$$\sum_{j=1}^{h} \lambda'_j(\boldsymbol{u}) = \lambda. \tag{6}$$

Then, summing up equation (3) of $\boldsymbol{x} \in X_{\boldsymbol{u}}$,

$$\sum_j \{(u_j + 1)\mu p(\boldsymbol{u}_j^+) + \lambda'_j(\boldsymbol{u}_j^-)p(\boldsymbol{u}_j^-)\} = (n\mu + \lambda)p(\boldsymbol{u}). \tag{7}$$

where $\boldsymbol{u}_j^+$ is the states of which j-th element is larger by one than $\boldsymbol{u}$ and $\boldsymbol{u}_j^-$ is smaller.

Above equations are obtained $\forall \boldsymbol{u}$. The possible number of states $\boldsymbol{u}$ is $(m+1)^h$. Under the condition $\sum_{\boldsymbol{u}} p(\boldsymbol{u}) = 1$, solving simultaneous equations, we obtain the values of $p(\boldsymbol{u})$.

3 Numelical Example in Linear City

3.1 Case of Equal Interval Between Facilities

Assume a city region as a line segment of which length is $L = 10$. Let the number of facility be $h = 5$, servers per facility be $m = 4$, arrival rate be $\lambda = 15$, and service rate per server be $\mu = 5$. Demand uses facilities in the near order. Set the origin at the left edge of the city region. The position of facility $1, \ldots, 5$ is $1, 3, 5, 7, 9$, respectively.

Figure 1 illustrates using rate of individual facilities to position where demand occur and Figure 2 shows accepted rate of k-th nearest facility, that is demand isn't provided service until $k-1$-th nearest facility. Demands receive service at high rate in the near order. There exists demand which can't receive service at constant rate. In addition, accepted rate at facility isn't constant, and it decreases progressibly as facility is farther.

Conditional average distance from demand to facility where service is provided is linear function. As shown in figure 3, it takes local minimum at positions of facility and local maximum at the middle point between facilities next to each other.

3.2 Optimal Location

As we described in section 2.1, the probability that demands are provided service in the city is constant, if it is constant that the total number of servers. This indicates that only distance to facility where demands are accepted is important for evaluating medical service in the whole city at this situation. In this section, we consider the location of

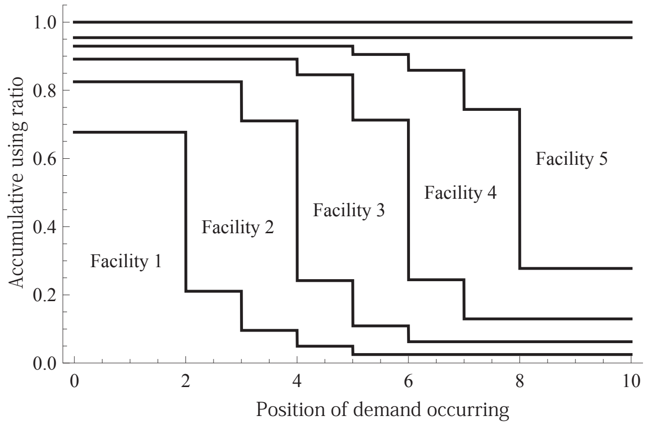

Figure 1: Using rate of individual facilities

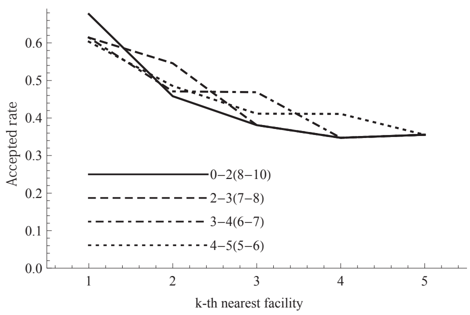

Figure 2: Accepted rate of k-th nearest facility

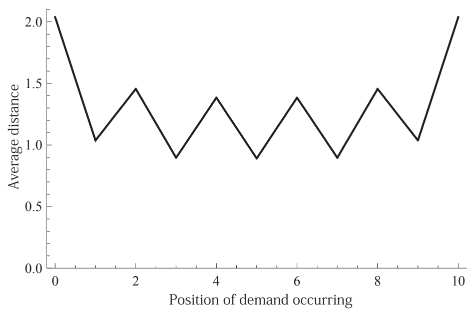

Figure 3: Average distance to accepted facility

facilities which minimizes average distance to acceptable facility, assuming a line segment as a city region

Using rate of individual facility depends on position of facilities and average distance is obtained by solving simultaneous equations (7). Hence, to obtain the location of facilities minizing average distance is so hard. Then, we obtain the optimum location by following simple iteration.

1: Set initial location.
2: Divide city region by order of closeness to facilities.
3: Calculate use rate of individual facility at each subregion.
4: Calculate the optimum location which minimizes total distance weighted by use rate of individual facility at each subregion.
5: Repeat from 2: to 4: certain times.

Results shown below are the locations of which total distance is the minimum obtained in iteration.

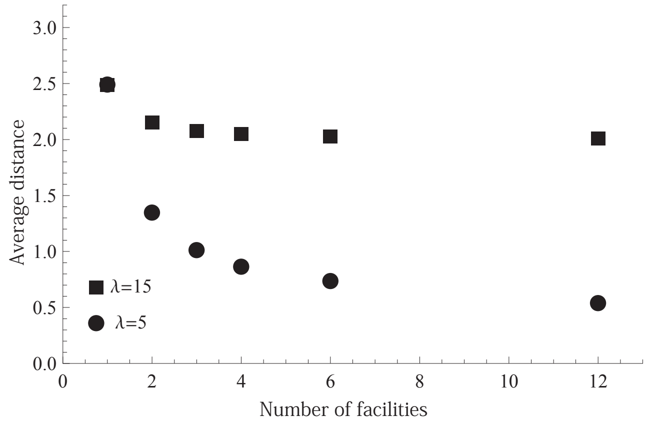

Figure 4: Minimum average distance and number of facilities

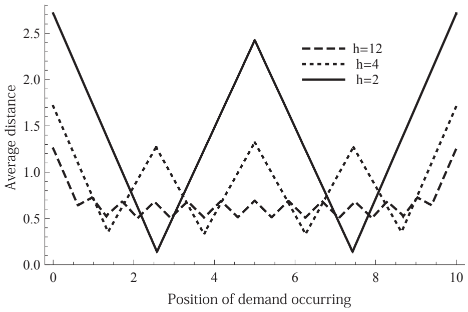

Figure 5: Average distance to accepting facility($\lambda = 5$)

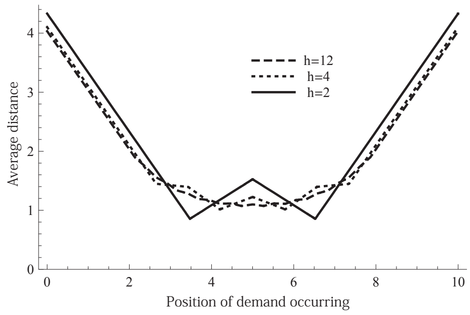

Figure 6: Average distance to accepting facility($\lambda = 15$)

As same as previous section, let a city region be a line segment of which length is $L = 10$, and let total number of servers be $M = 12$. We assume servrice rate is $\mu = 1$, arrival rate is $\lambda = 5, 15$. Number of facilities and average distance at optimal location are plotted in figure 4 . Average distance decreases monotonically as inversely proportinal to the number of facilities. In the case of same number of facilities, average distance at $\lambda = 15$ is always larger than that at $\lambda = 5$. This is explained as follows. Larger arrival rate λ leads the higher probability that demands can't use near facility if other conditions are same. Since demands are forced to use farther facility, average distance is long at large λ.

Figure 5 and 6 display average distance from position where demands occur to facilities where service is provided when the location of facilities is optimum. As we described in previous section, facilities are located at points where it takes local minimum in each function. While arrival rate λ is small, optimal location is distributed in whole region. In contrast, facilities are located intensively around center of the region for large λ. When λ is small, service rate has enough available capacity. The probability that demands are accepted at near facility is high. So, distance from demands to their nearest facility has great effect to average distance that is objective function. As a result, the optimal location is similar to typical p-median location. Conversely, when λ is large, service rate has not enough available capacity. The probability that demands are accepted at near facility is low. The situation that demands are passed around occurs highly. Then, high rate of using second, third nearest facility leads intensively located facilities.

From the view point of facilities, most of arrival demands are those from region where individual facility is the nearest in the case of small arrival rate. Then, optimum is the position minimizing total distance from their own region. For large arrival rate, users of facility exist distributed to the whole city region. Hence, each facility has the optimal position around center of the region, the location of facilities is intensive.

4 Aporoximation of Using Probability at Infinity Region

As shown above, the facility which accepts demands occuring in the city is determined by solving equation (7). This equation has $(m+1)^h$ variables. Thus, it is so hard to solve this equation for relative small value of m, h. To obtain average distance of demand to accepted facility, necessary thing is only the probability that demands uses each facility. The probability of state $\boldsymbol{u}$ is not required. In this section, we consider infinity line with uniformly distributed demands and facilities. We derive an approximation of the probability that demands are accepted by the k-th nearest facility.

4.1 Total Number of Users in Arvitrary Facilities

At first, we consider the total number of users in arbitrary facilities to obtain the probability that demands use k-th nearest facility. Let U_i^k be the set of state $\boldsymbol{u}$ that number of users of a facility i is k:

$$U_i^k = \{\boldsymbol{u} | u_i = k\}. \tag{8}$$

Then, summing up equation (7) all over U_i^k, we obtain

$$\sum_{\boldsymbol{u}\in U_i^{k+1}} (k+1)\mu p(\boldsymbol{u}) + \sum_{\boldsymbol{u}\in U_i^{k-1}} \lambda_i'(\boldsymbol{u}) p(\boldsymbol{u}) = \sum_{\boldsymbol{u}\in U_i^{k}} (k\mu + \lambda_i'(\boldsymbol{u})) p(\boldsymbol{u}). \tag{9}$$

From $\sum_{\boldsymbol{u}\in U_i^k} k\mu p(\boldsymbol{u}) = \sum_{\boldsymbol{u}\in U_i^{k-1}} \lambda_i'(\boldsymbol{u}) p(\boldsymbol{u})$, we obtain recurrence equation

$$\sum_{\boldsymbol{u}\in U_i^k} p(\boldsymbol{u}) = \frac{\sum_{\boldsymbol{u}\in U_i^{k-1}} \lambda_i'(\boldsymbol{u}) p(\boldsymbol{u})}{k\mu}. \tag{10}$$

Furthermore, consider set of arbitrary facilities $I = \{i_1, \ldots, i_\alpha\} (i_j \in \{1, \ldots, h\}, \alpha \leq h)$. Let U_I^k be the set of state $\boldsymbol{u}$ that total number of users of faciilty i the element of I is k. The probability is expressed as following recurrence equation:

$$\sum_{\boldsymbol{u}\in U_I^k} p(\boldsymbol{u}) = \frac{\sum_{\boldsymbol{u}\in U_I^{k-1}} \Lambda_I(\boldsymbol{u}) p(\boldsymbol{u})}{k\mu}, \tag{11}$$

where $\Lambda_I(\boldsymbol{u}) = \sum_{i\in I} \lambda_i'(\boldsymbol{u})$.

4.2 Approximation of Using Probability on Infinity Region

Suppose that demands exist uniformly on infinity region and facilities are located to equal interval. Let rate of demand ocurring per length be λ and density of facilities be μ. $I(t)$ denotes the set of facilities which are first, second, $\ldots$, t-th nearest from arbitrary position on the region. $R_s(\boldsymbol{u})$ denotes the set of subregion on which demands use $I(t)$ at state $\boldsymbol{u}$ among all regions of which s-th nearest is $I(t)$ and $L(r)$ denotes the largeness of subregion r. Since occurring rate λ is constant, arrival rate of demands which occur on subregion r is expressed as $\lambda L(r)$. Demands on $r \in R_s(\boldsymbol{u})$ are passed around to $I(t)$, if the first, second, $\ldots$, $s-1$-th nearest facility are not available. Let the set of such states $\boldsymbol{u}$ be Ω_{s-1}^r. We obtain righ-hand of equation (11) as

$$\sum_{\boldsymbol{u} \in U_{I(t)}^{k-1}} \Lambda_{I(t)}(\boldsymbol{u}) p(\boldsymbol{u}) = \sum_{s=1}^{\infty} \sum_{r \in R_s} \lambda L(r) p(U_{I(t)}, \Omega_{s-1}^r). \tag{12}$$

Here, we assume

$$p(U_{I(t)}, \Omega_{s-1}^r) \simeq p(U_{I(t)}^{k-1}) \cdot p(\Omega_{s-1}^r).$$

Then, $p(\Omega_s^r)$ is independent from region r. From the assumption that facilities are located at equal interval, total measurement of the subregions of which the s-th nearest are a certain facility is constant. Considering this, we obtain from equation (11) as

$$\begin{aligned} p(U_{I(t)}^k) &= \frac{a_t}{k} p(U_{I(t)}^{k-1}) = \frac{a_t^k}{k!} p(U_{I(t)}^0) \\ a_t &= \frac{\lambda}{h\mu} \left(t + \sum_{s=1}^{\infty} p(\Omega_s) \right) \end{aligned} \tag{13}$$

In above equation, $k = tm$ derive the probability $p(\Omega_t)$ that all facilities belonging to $I(t)$ are unavailable.

$$p(\Omega_t) = \frac{a_t^{tm}/(tm)!}{\sum_{k=0}^{tm} a_t^k / k!} \tag{14}$$

4.3 Numerical Example

Average distance and density of facilities on infinity line are plotted in figure 7 Density of servers is 2.4, rate between arrival and service is $\lambda/\mu = 12, 18$.

Average distance decreases as inversely propotional to density of facilities for small ratio λ/μ. In contrast, it takes local minimum at certain density of facilities for large λ/μ. Same as numerical example on the city of line segment in section 3, the probability that near facility is unavailable is low in the case of enough service rate, and facilities are located distributedly. On insuffcent service rate, the probability that demands are accepted is low. Hence, increasing the scale of facility, they decrease the probability.

5 Summary and Conclusion

In this paper, we considered number and location of facility, based on distance from demands to facility which provides service, taking into account the available probability of server at the situation that resources are distributed to each facility.

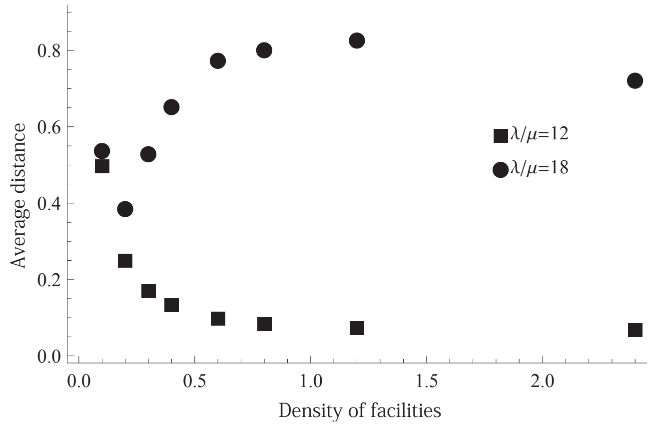

Figure 7: Relation between density of facilities and average distance on infinity line

In the example of line segment city in section 3, we find that average distance to accepted facility monotonically decreases as number of facilities increase. Optimal location of facilities is distributed if servers have enough availability. This means that it is better to construct as many facilities as possible, however, to build desired number is impossible in realistic situation. Though many facilities decrease average distance, difference In section 4, we shows that certain density of facilities minimizes average distance depending on a relation between ability of server and occurring rate, assuming infinite region.

Acknowledges

This research was partly supported by Japan Society for the Promotion of Science Grant-in-Aid for Young Scientists(B), 21760492, 2009.

References

[1] G. M. Carter, J. M. Chaiken and E. Ignall, Response areas for two emergency units, *Operations Research*, 20 (1972), 571–594.

[2] R. C. Larson, A hypercube queueing model for facility location and redistricting urban emergency services, *Computer & Operations Research*, 1 (1975), 67–95.

[3] R. C. Larson, Approximating the performancs of urban emergency service systems, *Operations Resarch*, 23 (1975), 845–868.