Journal of the Operations Research Society of Japan
Vol. 54, No. 4, December 2011, pp. 237–258

MAXIMUM FLOW-COVERING LOCATION AND SERVICE START TIME PROBLEM AND ITS APPLICATION TO TOKYO METROPOLITAN RAILWAY NETWORK

Ken-ichi Tanaka
The University of Electro-Communications

(Received August 3, 2011; Revised October 7, 2011)

Abstract This paper proposes the Maximum Flow-Covering Location and service Start Time Problem (MFCLSTP). The problem seeks to determine locations and service start time of p facilities which provide a service for a given duration, so as to maximally cover flows. Each flow is specified by a pair of origin-destination (OD) nodes and the departure time of the origin node. We assume that a given commuter flow is covered when commuters can stop at a facility, fully consume that facility's service, and arrive at the destination node by a given time. Two models are considered for MFCLSTP: MFCLSTP1, in which the service start time of each facility can be independently determined; and MFCLSTP2, in which all facilities have the same service start time. We provide integer programming formulations and propose heuristic solution algorithms. The proposed models are applied to a case study of the Tokyo metropolitan railway network using census data for commuter traffic. The solutions obtained by the heuristic algorithms for both models are compared. Solutions of MFCLSTP2 show that selected locations are spatially dispersed to cover different types of flows, whereas solutions of MFCLSTP1 closely locates some facilities having different start times in the central area of Tokyo.

Keywords: Facility planning, maximal covering objective, dynamic location model, service start time, railway network, commuter traffic flow

1. Introduction

Facility location decisions are an important element in strategic planning in both private and public organizations. Various models dealing with optimal location of facilities have been proposed in operations research and management science. Examples include the p-median problem and the maximal covering location problem. The p-median problem seeks the locations of p facilities among a given set of candidate locations, such that the total demand-weighted distance to the nearest facility is minimized, while the maximal covering location problem finds a given number of facilities so as to maximize the total demands that have a facility within a given distance threshold. There are a number of ways to classify location models, and Daskin [8] gives a taxonomy of the types of location models.

Many of the location models developed thus far have focused on static and deterministic problem formulations. However, facility location decisions are often long-term in nature, during which time the environment may change considerably. Focusing on this aspect, researchers have attempted to incorporate the temporal dimension into facility location problems. For instance, models have been formulated that consider the timing of the location and relocation of facilities based on the temporal variation of future demands. These dynamic location models describe time as a long-term planning horizon, and various important models have been proposed (e.g., see the review article of Owen and Daskin [20]). Another possible approach to dynamic location problems is to concentrate on the daily movement of people, and to consider the provision of services at facilities in spatio-temporal dimensions. The latter approach has yet to be seriously employed in the existing dynamic location literature.

In this paper, we propose a dynamic facility location model that seeks to determine the optimal provision of services on a daily basis. Consider, for example, how to provide after-work lectures to commuters who wish to study at graduate school. To fully capture the opportunity for attending school, the start time of lectures must be sufficiently late to allow many commuters access to the graduate school after work, but sufficiently early in order for commuters to get back home early enough. To deal with this type of problem, we assume that facilities provide fixed hours of service, and consider the problem of optimizing both the locations and service start times of the facilities so as to maximally cover flows. We call this the Maximum Flow-Covering Location and service Start Time Problem (MFCLSTP). Two different situations regarding the service start times of facilities are considered for MFCLSTP. The *independent* start time model (MFCLSTP1) assumes that the service start time of each facility can be determined independently, whereas the *common* start time model (MFCLSTP2) assumes that all facilities have the same service start time. The latter situation is analogous to a lecture being delivered at a number of locations simultaneously in real-time.

The remainder of this paper is organized as follows. In the next section, literature related to the MFCLSTP is reviewed, including the maximal covering location problem and its variants, dynamic facility location problems, topics concerned with space-time accessibility in time geography, and the single-facility model that determines the location and service start time of a facility in spatio-temporal dimensions. Then, in Section 3, we describe the general situation assumed in MFCLSTP, and formulate the proposed models as integer programming problems. As the size of the problems can easily become very large even for small-sized networks, obtaining exact optimal solutions for the proposed problems is not easy. Therefore, we propose a heuristic solution algorithm in Section 4, which is based on the well-known node-exchange algorithms for static location problems.

The latter part of this paper is devoted to a case study of the Tokyo metropolitan railway network. In modeling the case study, we assume commuters on their way home from work as potential demands for facility service. We obtain an OD matrix for the commuter traffic flow from census data, and construct dynamic flow data by introducing a departure time for the origin station of each flow. First, a case that locates a single facility at each station is examined in Section 5. The number of covered flows at each station for various service start times is evaluated, and the desirable site and start time for service provision in Tokyo metropolitan railway network is analyzed. Then, in Section 6, we explore multi-facility problems in which both the locations and service start times of several facilities are simultaneously determined for the target network. Applying the proposed heuristic solution algorithms, the solutions obtained for the independent and common start time models are compared. Solutions of the common start-time model show that selected locations are spatially dispersed to cover different types of flows, whereas solutions of the independent start time model closely locates facilities having different service start times in the central area of Tokyo. Finally, in Section 7, we conclude the paper and discuss future research directions.

In summary, the proposed model focuses on spatio-temporal movement of people and the provision of services on a daily basis, a topic not fully explored in the existing dynamic facility location literature. This modeling approach is the main contribution of this paper, by which several variants and extended models can also be constructed. Application of the model to the Tokyo metropolitan railway network is also a fundamental goal of this paper, from which interesting properties of the target network are discovered. As for the heuristic algorithm, simple local search method proposed in this paper can be extended by employing metaheuristic approaches.

2. Literature Review

We now review four classes of model that are related to the present study. The first class of models is the maximal covering location problem (MCLP) and it variants, which locate facilities in order to maximize the number of demands within a predefined distance of the nearest facility. In particular, the flow-capturing location problem (FCLP), an extension of MCLP, is closely related to the proposed problem, and MFCLSTP can be seen as a dynamic extended version of FCLP. The second class is dynamic facility location problems, which incorporate temporal aspects in facility location problems. For the third class, we briefly review accessibility models in time geography. These models explicitly measure accessibility in spatio-temporal dimensions, although the aim is not on optimizing facility service. Finally, we review the single facility model that was the first to investigate a similar situation to that considered here, but was limited to only the single facility case.

The objective of MFCLSTP is to maximize the number of potential demands for services (i.e., the covered flows). This type of objective first appeared in Church and ReVelle [5], and has been one of the most employed objectives in location literature. MCLP seeks the optimal locations of facilities such that the number of demands is maximized within a pre-specified distance (time) of the nearest facility. Extensions to the basic MCLP include models that consider a demand to be covered when two or more facilities are located within a distance threshold [9], and models that consider different levels of coverage [6]. Berman et al. [3] has recently reviewed current developments of a number of generalized coverage models. These extended MCLP approaches can also be applied when constructing extensions of MFCLSTP.

An important generalization of MCLP is the FCLP introduced by Hodgson [14], in which the demands for services are represented as flows traveling along paths between origin-destination pairs over a network. FCLP attempts to locate a given number of facilities on the network in order to maximize the total flows having a facility along their pre-planned route. The flow capturing approach is suitable for many services, such as automatic teller machines, convenience stores, advertising billboards, and vehicle inspection stations. Because of the wide applicability of FCLP, different modified versions have been proposed, including a model in which different coverage levels are introduced, according to the location captured [31]; and the multi-counting model [1], in which consumers can be captured more than once when there are several facilities on their route. The deviation situation assumed in Berman et al. [2] is particularly relevant to the present work. They considered that flows having a facility near the pre-planned route (in terms of the deviation distance) can access the facility, and hence those flows can be considered as captured flows. MFCLSTP assumes that commuter flows stop at a facility to consume service. Hence, deviation to access a facility is also allowed within the present framework. Details on FCLP variants can be found in a recent article by Zeng et al. [31]. Although, various generalized FCLP models have been proposed, all the existing models assume that flows are static and do not account for temporal factors. Thus, MFCLSTP can be considered as the first dynamic version of FCLP where each flow is identified not only by a pair of origin-destination nodes but also by the departure time of the origin node. In addition, MFCLSTP dynamically determines the service start times of facilities, as well as their locations.

In general, research on location theory has been devoted to static and deterministic models. These model formulations take constant quantities as inputs, such as demands and travel times, and output a single solution at a point in time. In reality, however, facilities (e.g., schools, distribution centers, plants, and retail outlets) typically exist for decades, during which time the environment in which they operate may change substantially. As a result, demands, travel times, land acquisition costs, and other inputs to classical static facility location models can be highly uncertain, and hence the need for dynamic facility location models arises. Current et al. [7] classifies dynamic models into two categories: *implicitly* and *explicitly* dynamic models. In implicitly dynamic models, facility location decisions are considered "static", in the sense that all facilities are opened at the same time and remain open over the planning horizon. However, they are dynamic models since they still recognize that problem parameters may vary over time. Examples of implicitly dynamic models include those developed by Mirchandani and Odoni [19], and Weaver and Church [30], who both consider problems where the demands and travel times change over time. Explicitly dynamic models assume a situation in which facilities will be opened (and possibly closed) over time. Typically, explicitly dynamic models extend known static models by introducing temporal subscripts to both the facility location and the assignment variables and constraints, thus adding a temporal dimension to these parameters. Explicitly dynamic model examples include those from Daskin et al. [10], Drezner and Wesolowsky [11], Schilling [21], and Gunawardane [12]. The latter two of these employ multi-objective approaches to explicitly dynamic problems. Several authors dealt with deployment problems of ambulances in a dynamic setting (e.g., [22]). The potential demand for emergency services and the travel times between points in a city can change throughout the day. By focusing on this issue, models that allow ambulance vehicles to change their locations, so as to cover potential demand at any point in time, have been explored. In this type of study, instead of modeling the movement of people to be serviced by a facility in spatio-temporal dimensions, the number of potential demands at each location and time is given as an input. For more detailed reviews of dynamic location problems, see Current et al. [7], Owen and Daskin [20], and Snyder [23].

Hägerstrand [13], who laid the foundation of time geography, stressed the temporal factor in spatial human activities. He introduced the space-time framework to evaluate individuals' accessibility to the environment, recognizing that activity participation has both spatial and temporal dimensions. Authors who have devised accessibility measures in spatial-temporal environments include Kwan [17] and Miller [18]. However, while efforts to develop space-time accessibility measures from an individual viewpoint have been actively made, attempts to optimize facility service from the viewpoint of a decision maker have not been seriously treated in this area.

The earlier model that first focused on a similar situation to the current study appeared in Tanaka [26]. This model assumed a one-dimensional linear city over which origins (workplaces) and destinations (homes) are continuously distributed, and the distribution of the departure time at an origin location was given by a continuous mathematical function. Tanaka [26] considered the location and service start time of a single facility that maximizes the number of commuters who can stop at the facility after work and can still get back home by a given time after consuming a service. The goal of the study was to describe the tradeoff structure of the service start time and to investigate the relation among the covered volume, optimal service provision and input parameters, such as the trip distributions and the duration of the facility service. The same problem was later examined for a two-dimensional circular city [27]. A variation of the model was analyzed by Honda [15], who assumed a point-based demand for facility service (home–facility–home movement), and introduced a joint distribution for possible departure and home times as input data. Honda's model used Web-based questionnaires to estimate these joint distributions, according to several classes of users attributes, and analyzed the desirable service provision for each attribute. However, all of these studies considered only the single facility case, and MFCLSTP can be regarded as the multi-facility, discrete version of these models.

3. Model Description and Formulation

To describe the movement of people and the provision of services in spatio-temporal dimensions, we consider the situation as shown in Figure 1 (a) in which the temporal dimension is introduced into a network. Each point in Figure 1 (a) corresponds to a specified location and time. A two-dimensional version of the model (one dimensional city, plus temporal dimension) is also presented in Figure 1 (b) to facilitate explanation. Let us consider the situation in which a decision maker is planning to introduce p facilities that provide fixed service hours c, as illustrated by the line segments in Figure 1 (a). We assume that the service provided at a facility must be fully consumed. Examples of this type of service are cinema, baseball games, and lectures, where partial consumption does not make sense.

We assume that after-work commuters access a facility, consume c hours of service, and desire to get back home sufficiently early. To describe this situation, the arrival time limit, t_{h}, is introduced and represents the latest allowable arrival time at a destination node (the commuter's home location). Thus, commuters can *access* a given facility's service when they can arrive at the facility before the service start time, consume c hours of service, and return home by t_{h}. In addition, a commuter flow is defined as being *covered* when there is at least one accessible facility among the p facilities.

Let us illustrate a covered customer in space-time dimensions by using Figure 1 (b). A commuter flow is identified by an origin node, i, a destination node, j, and a departure time from the origin node, t. We refer to this commuter flow by (i, j, t). Each facility's service is specified by a location, k, and a service start time, s, and thus can be referred to (k, s). Flow (i, j, t) in Figure 1 (b) is covered by facility service (k, s) since this flow can:

(i) arrive at location, k, before s; and

(ii) arrive at destination node, j, by t_{h} after fully consuming c hours of service at (k, s).

The first condition can be written as $t + u_{ik} \leq s$, while the second condition can be written as $s + c + u_{kj} \leq t_{\mathrm{h}}$. Here, u_{ij} is the travel time between node i and node j and we assume that t and s take only discrete values.

Covered flows can be considered as potential demands in the sense that, if they wish to consume a facility's service, they can access the service at a facility. The notion of coverage here is similar to that of the classical maximal covering location problem [5], in which customers are considered covered when they have at least one facility within an acceptable distance.

The decision maker wants to determine the provision of the p facilities' services such that the number of covered flows is maximized, the problem we have termed MFCLSTP. In MFCLSTP, two different situations can be considered. The first model (MFCLSTP1) assumes that the service start time of each facility can be determined independently. Conversely, the second model (MFCLSTP2) supposes that all facilities have the same service start time. To maximize opportunities for consuming service, service start times must be sufficiently late to allow many commuters access to a facility after work, but early enough for commuters to get back home by t_{h}. This tradeoff in the service start time of facilities has not been sufficiently focused on in the existing facility location literature.

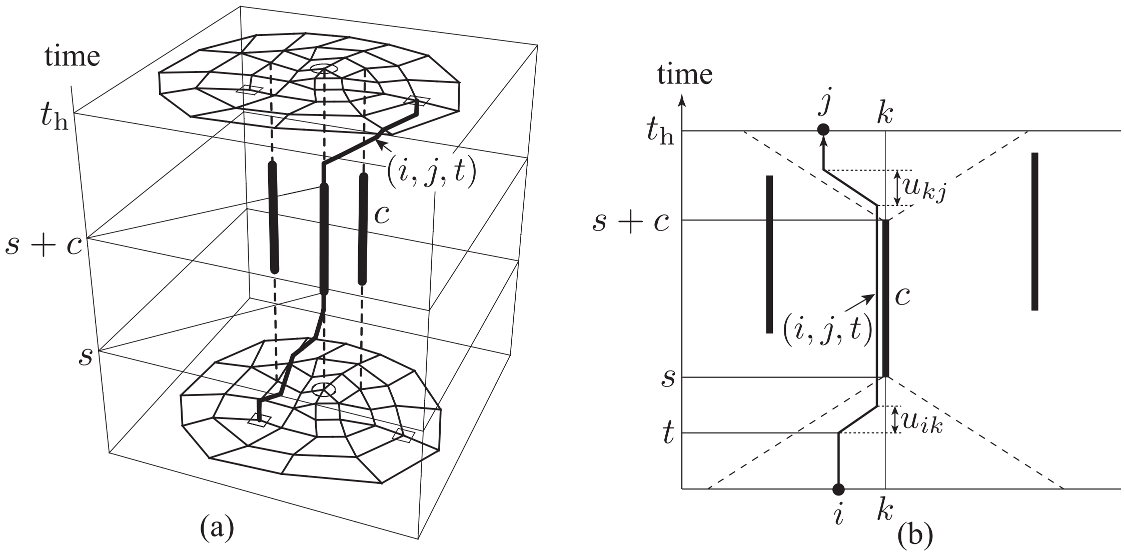


Figure 1: Covered flow in the spatio-temporal dimensions: (a) general network case, (b) linear city case

To formulate MFCLPSTP as an integer programming problem, we introduce the following notation.

Sets

N: set of nodes (used for both origin and destination nodes)
T: set of departure times for an origin node
K: set of potential facility locations
S: set of potential service start times

Parameters

p: number of facilities to be located
c: duration of facility service
f_{ijt}: volume of flow (i, j, t)
u_{ij}: travel time between node i and node j
t_{h}: the arrival time limit which describes the latest allowable arrival time to a destination node in order for a commuter flow to be covered

Moreover, we introduce the coverage index a_{ks}^{ijt} that indicates whether flow (i, j, t) can access facility service (k, s):

$$a_{ks}^{ijt} = \begin{cases} 1 & \text{if flow } (i, j, t) \text{ can access facility service } (k, s), \\ 0 & \text{otherwise.} \end{cases}$$

To calculate a_{ks}^{ijt}, all combinations of (i, j, t) and (k, s) must be correctly assigned 0 or 1.

Finally, two further binary variables are introduced:

$$x_{ks} = \begin{cases} 1 & \text{if a facility is to be located at node } k \text{ and starts its service at time } s, \\ 0 & \text{otherwise,} \end{cases}$$

$$y_{ijt} = \begin{cases} 1 & \text{if commuter flow } (i, j, t) \text{ is covered,} \\ 0 & \text{otherwise.} \end{cases}$$

Using the above definitions, MFCLSTP1, with independent service start time for each facility, is formulated as

MFCLSTP1

$$\begin{aligned}
\text{maximize} \quad & \sum_{i \in N} \sum_{j \in N} \sum_{t \in T} f_{ijt} y_{ijt}, && (3.1)\\
\text{subject to} \quad & \sum_{k \in K} \sum_{s \in S} x_{ks} = p, && (3.2)\\
& y_{ijt} \leq \sum_{k \in K} \sum_{s \in S} a_{ks}^{ijt} x_{ks} \quad \forall i \in N, \forall j \in N, \forall t \in T, && (3.3)\\
& x_{ks} \in \{0,1\} \quad \forall k \in K, \forall s \in S, && (3.4)\\
& y_{ijt} \in \{0,1\} \quad \forall i \in N, \forall j \in N, \forall t \in T. && (3.5)
\end{aligned}$$

The objective function (3.1) is the total number of covered flows, that is, the number of commuters with access to at least one of the p facilities. It should be noted that maximizing this function is equivalent to minimizing the number of commuters who cannot access a service at any of the facilities. Constraint (3.2) stipulates that exactly p facility services are provided. Notice that this constraint does not prohibit placing multiple facilities at the same node, and two or more co-located facilities with different service start times may cover different flows. Constraints (3.3) require that at least one facility service, (k, s), be accessible by commuter flow, (i, j, t), for this flow to be covered. Finally, constraints (3.4) and (3.5) are the standard binary constraints on the decision variables.

In certain situations, it is difficult or may be impossible to independently decide service start time of each facility. For example, delivery of a lecture at a number of locations simultaneously in real time requires all facilities to start their service at the same time. In other situations, it may be desirable to start services at the same time because doing so makes the operating cost of facilities lower than in the independent case. To formulate MFCLSTP2, the variables z_s that determine service start times are introduced:

$$z_s = \begin{cases} 1 & \text{if all facilities start service at time } s, \\ 0 & \text{otherwise.} \end{cases}$$

With this definition of z_s, MFCLSTP2 can be formulated similarly to MFCLSTP1 with the addition of two extra constraints:

MFCLSTP2

$$\begin{aligned}
\text{maximize} \quad & (3.1), \\
\text{subject to} \quad & (3.2), (3.3), (3.4), (3.5), \\
& p\, z_s = \sum_{k \in K} x_{ks} \quad \forall s \in S, && (3.6)\\
& z_s \in \{0,1\} \quad \forall s \in S. && (3.7)
\end{aligned}$$

Combining constraints (3.6) and (3.2) implies that all facilities start their service at the same time, and constraints (3.7) denote standard binary constraints placed on the decision variables.

4. Heuristic Algorithm

The number of variables and parameters in MFCLSTP can become large even for a network of moderate size. Therefore, we develop a heuristic algorithm for MFCLSTP, that iteratively updates each facility's service. The basic idea of the proposed heuristic is that if spatial and temporal variables are separated, we can update one set of variables, while the other is fixed. This separation of variables reduces computational burden to a great extent. The algorithms outlined in this section are implemented under a multi-start local search framework, and are applied to the location analysis for the example of the Tokyo metropolitan railway network in the next section.

The proposed algorithm can be seen as an extension of the well-known node-exchange algorithm that was originally developed by Teitz and Bart [28] for solving the p-median problem. This procedure has been successfully applied to several static facility location models and several modified versions have been proposed ([4]). Here, to reduce the computational load, the number of locations selected as potential exchange candidates is restricted to the Q nearest locations to the current facility location, as shown in Figure 2.

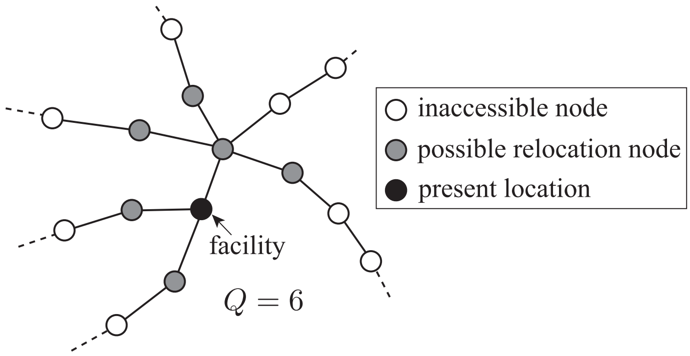


Figure 2: Relocation candidates of a facility during the location improvement procedure for a facility

Local search algorithm for MFCLSTP1

In MFCLSTP1, facility service decisions are described by $2p$ variables (p facility locations and p service start times). We separate the p spatial variables and p temporal variables, and attempt to improve the objective value by varying one set of variables, while the other set is fixed. The proposed algorithm for MFCLSTP1 is described as follows.

Step 1: *Initialization*
Randomly select p facility locations and p service start times.

Step 2: *Update procedure for facility locations*

Step 2-1: For facilities $l = 1, \cdots, p$, by fixing the locations of the remaining $p - 1$ facilities, perform the following procedure:

(1) Evaluate the objective value obtained from relocating the lth facility to each of the candidate locations selected in the neighborhood of the present location of the lth facility;

(2) If the best relocation of lth facility among Q nodes improves the objective value, relocate the lth facility to the best node.

Step 2-2: If at least one relocation is performed among the p facilities in Step 2-1, go to Step 3; otherwise go to Step 2-3.

Step 2-3: If no exchange was made in the last execution of Step 3-1, output the current solution and terminate the algorithm; otherwise go to Step 3.

Step 3: *Update procedure for service start times*

Step 3-1: For facilities $l = 1, \cdots, p$, perform the following procedure while fixing the service start times of remaining $p-1$ facilities:

(1) Evaluate the objective value obtained from changing the service start time of the lth facility to each element in S;

(2) If the best service start time for the lth facility improves the objective value, select the best start time for the lth facility.

Step 3-2: If at least one change is performed among the p facilities in Step 3-1, go to Step 2; otherwise go to Step 3-3.

Step 3-3: If no exchange was made in the last execution of Step 2-1, output the current solution and terminate the algorithm; otherwise go to Step 2.

Termination of the algorithm thus occurs when both the facility locations and service start times remain unaltered in the update procedures.

Local search algorithm for MFCLSTP2

Similarly, the algorithm for MFCLSTP2 is given as follows.

Step 1: *Initialization*

Randomly select p facility locations and one service start time.

Step 2: *Update procedure for facility locations*

Step 2-1: For facilities $l = 1, \cdots, p$, by fixing the locations of the remaining $p-1$ facilities, perform the following procedure:

(1) Evaluate the objective value obtained from relocating the lth facility to each of the candidate locations selected in the neighborhood of the present location of the lth facility;

(2) If the best relocation of lth facility among Q nodes improves the objective value, relocate the lth facility to the best node.

Step 2-2: If at least one relocation is performed among the p facilities in Step 2-1, go to Step 3; otherwise output the current solution and terminate the algorithm.

Step 3: *Update procedure for service start time*

Step 3-1: Evaluate the objective value obtained from changing the service start time of all facilities to each element in S. If the objective value is improved, change the service start time of all facilities to that producing the best value.

Step 3-2: If a change has occurred, go to Step 2; otherwise go to Step 3-3.

Step 3-3: If no exchange occurred in the last execution of Step 2-1, output the current solution and terminate the algorithm; otherwise go to Step 2.

Owing to the common service start time for the facilities in MFCLSTP2, Step 2 and Step 3 are simpler than those for MFCLSTP1. Note that a single execution of Step 3 returns the optimal service start time for fixed facility locations. Therefore, when Step 2 does not relocate any of the facilities, further execution of Step 3 cannot improve the objective value, and the algorithm can be terminated in Step 2-2.

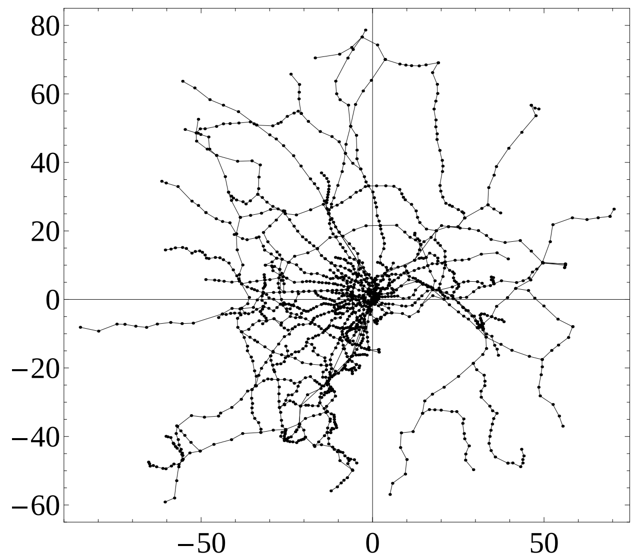


Figure 3: Tokyo metropolitan railway network used in the analysis

5. Analysis of Covered Flows for Each Station in the Tokyo Metropolitan Railway Network

First, single facility MFCLSTP (i.e., when $p = 1$) is applied to our example case study of the Tokyo metropolitan railway network by using commuter traffic flow data extracted from census data. The purpose of this section is to analyze the number of covered flows that can be observed and at what time service should start for each station. We start by explaining how to construct data used in the analysis. Then, the single facility case is analyzed by placing one facility at each location and varying the service start times. Cases of multiple-facilities are investigated in the subsequent section.

5.1. Data description

We construct the railway network covered by census data for the Tokyo metropolitan area [16]. The target network is shown in Figure 3 (Tokyo station is at the origin and each unit represents 1 km). The network is composed of 1,804 stations along 125 lines, from a total of 1,815 stations along 128 lines in the census data; here we exclude three Shinkansen (bullet train) lines. This number of 1,804 includes stations possessing the same name on different lines; for example, both JR Chuo and Yamanote lines have a station at Shinjuku. In the census data, these two Shinjuku stations are treated separately. However, for the purposes of the present analysis, transfer stations bearing the same name and stations so close to each other are treated as a single station. This aggregation of stations is conducted by referring to the Station Database [24], which contains information on groups of stations between which transfers are possible. For example, although there are nine Shinjuku stations and eight Tokyo stations, we consider them as single “Shinjuku station” and “Tokyo station”. Among the targeted 1,804 stations, 663 stations have at least one transfer station, and these can be aggregated into a total of 252 stations. Thus, the number of target stations becomes 1,393 ($= 1{,}804 - 663 + 252$) after aggregation, and we use all 1,393 stations as candidate locations for facilities.

The model and corresponding heuristic algorithm use the times required to travel between station pairs (the travel time matrix) as an input. Construction of the travel time matrix that is to be used in the following analysis is next explained. The cost (time) between adjacent stations is calculated as the length of the link (Euclidean distance) between the stations divided by the speed of the trains. We use an identical speed of 50 km/h for all trains when calculating link costs. Then, the travel time matrix is obtained by calculating the shortest paths for all station pairs by using the Dijkstra's algorithm.

By using the census data, we construct an origin-destination (OD) traffic flow matrix, in which each entry represents the number of commuters traveling from an origin station, i, to a destination station, j. In total, 113,591 out of 1,940,449 (1,393 by 1,393) entries in the OD matrix have non-zero flow volumes, and the total number of flows is 8,529,216. To generate the input data for MFCLSTP, this static flow data is divided and assigned to departure times t such that dynamic flow data, f_{ijt}, is created for each flow, (i, j, t).

5.2. Covered flows for each station

We now analyze the case in which only one facility is located at a station within the network for various service start times. To conduct the analysis, values for parameters are set to $c = 3$ h, $t_{\mathrm{h}} = 23{:}00$. Moreover, we assume that departure from the origin station for each commuter flow, (i, j), occurs at a uniform rate between 17:00 and 21:00 at 10 min intervals; specifically, for each (i, j), the flow volume is multiplied by 0.04 and assigned to 25 equally spaced departure times. Under these assumptions, the latest service start time is 20:00, since a service starting after 20:00 has a finish time later than 23:00, and commuters cannot reach their destination station by t_{h}. Service start times are chosen at 19 candidates times every 10 min from 17:00 to 20:00, i.e., 17:00, 17:10, ..., 19:50, 20:00.

Before proceeding to the numerical results, let us calculate the objective value when we can locate a facility at all stations; this value is an upper bound for the objective value when constraint (3.2) is relaxed. The resulting value provides useful information for evaluating the quality of a given solution (the services for the p facilities). This upper bound for the objective value can be computed by summing up the covered flows for all (i, j, t) when each of the 1,393 stations has a facility providing a service from 20:00 to 23:00. In this situation, all flows are at the destination node at exactly 23:00, the arrival time limit t_{h}. For each origin-destination node pair (i, j), commuter flows departing from their origin station before $t = t_{\mathrm{h}} - c - u_{ij}$ are covered, since the commuters can be at the facility before the service start time. For the case presented here, 5,441,794 (63.80%) flows among a total 8,529,216 flows are covered. The remaining flows cannot be covered by any solution under this parameter setting. Hence, in the following, the flows covered by a given solution are evaluated as a percentage of the upper bound of this objective value, 5,441,794.

We built a program in the C++ programming language that computes the covered flows. By using the program, covered flows are calculated for each station at the 19 service start times. Figures 4 (a) to (e) show the percentage of covered flows at each station by the area of circle centered at the station for five service start times: 17:30, 18:00, 18:30, 19:00, and 19:30. In general, stations located in the central part of the Tokyo metropolitan area can cover a large percentage of commuters, and stations located in suburban areas have smaller coverage. We see that the service start time strongly affects the number of covered flows. When service starts at 17:30, many commuters are unable to arrive before the service start time, whereas starting service at 19:30 makes many commuters unable to reach their destination station by t_{h}. Therefore, both cases are unattractive when planning service at a facility. In this example, starting service at 19:00 captures a large number of commuters.

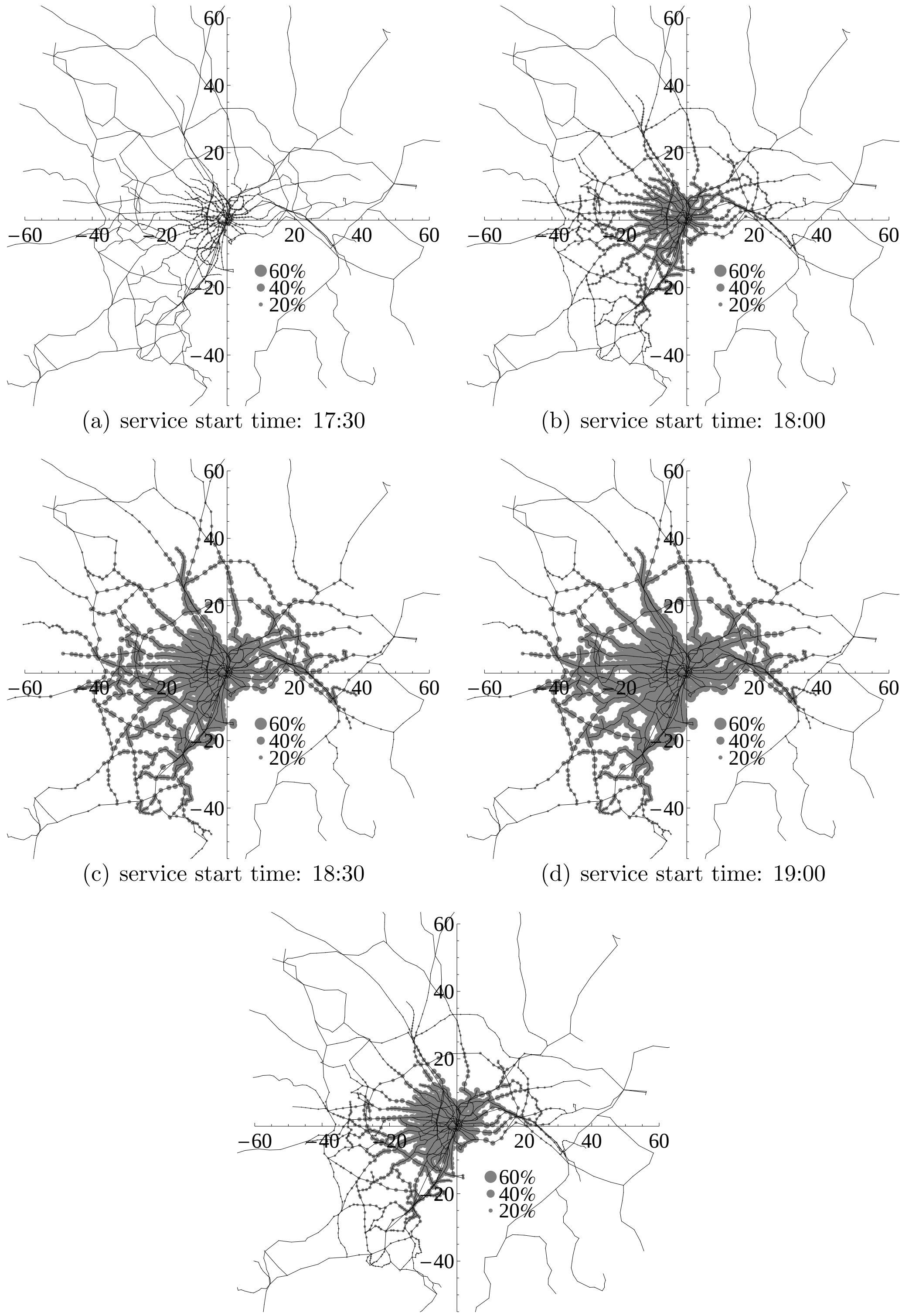
To study the effect of service start time on the objective value in more detail, we concentrate on three stations on the JR Chuo line: Shinjuku, Musashisakai, and Hachioji. Figure 5 shows the objective values for the three stations at each service start time. Shinjuku, which is located in the center of Tokyo, has the highest values, followed by Musashisakai, and finally Hachioji, which is located in a suburban area. Furthermore, the form of the graphs is different, especially between Shinjuku and Hachioji. At early start times, the covered flow for the facility at Hachioji is small, while the facility at Shinjuku can capture larger volumes of flows. This difference is due to workplaces being densely located around Shinjuku, allowing commuters access to the facility at Shinjuku earlier. In contrast, the population during the day near Hachioji is not so large, and few commuters can be in time for early facility services, especially those commuters working in the central area of Tokyo. Another interesting characteristic in Figure 5 is that the optimal service start time that maximizes the covered flow is 20 min later for Shinjuku than for Hachioji. This late start time is a result of the high accessibility of Shinjuku. After consuming a service at Shinjuku, a large number of destination stations can be reached within a short time, permitting commuters to return home by t_{h}. Therefore, a later start for the facility's service is advantageous, since a large number of commuters who are able to arrive at Shinjuku before the service start time can be captured.

We next analyze the distribution of the maximum value across the network, and also when the optimal service start time occurs for a facility at each station. Figure 6 denotes the maximum objective value at each station by the area of its associated circle, and those colored black correspond to the 20 stations with the highest objective values. In addition, Figure 7 illustrates the start time at which the objective value is maximized. In Table 1, the names are listed of the 40 stations having the highest percentage of coverage and their optimal service start time. From the table, the optimal solution of the MFCLSTP when $p = 1$ is to start a service at 19:10 at a facility located in Shinjuku. The second best location is Shibuya. Both Shinjuku and Shibuya are vast terminal stations having numerous railway lines such that a considerable number of stations are accessible within a reasonable amount of time, and both have large daytime populations in and around them. Furthermore, many of the highest performing stations are those nearby Shinjuku and Shibuya. Tokyo, which has the second largest number of transfer stations after Shinjuku, also performs well. As can be seen from Figure 6 and Table 1, stations located between Shinjuku and Tokyo, and their environs, attain large objective values. Figure 7 and Table 1 indicate that in all cases the optimal service start time for the 40 stations is later than for only having smaller maximum objective values: 19:00 or 19:10. As already discussed for Shinjuku, stations with high accessibility overall can be reached from a large number of stations in a relatively short time. Thus, starting the facility's service at these stations later is advantageous in maximizing the number of commuters able to access the facility before the service start time.

6. Multi-facility Analysis on the Tokyo Metropolitan Railway Network

Having analyzed the single facility case, we now turn our attention to locating two or more facilities. When multiple facilities are located, depending on whether the service start times are independent or equal for all facilities, we can consider the two models outlined in Section 3: MFCLSTP1 and MFCLSTP2. To obtain solutions for these two cases, we apply the heuristic algorithms introduced in Section 4 to MFCLSTP1 and MFCLSTP2 to cases where the numbers of facilities to be located were $p = 2$ to 7. With the exception of p, all parameter values are set to those for the single-facility case.

All 1,393 stations are considered as candidate locations for facilities, and we implement the heuristic algorithms as a multi-start local search (MLS), with the number of candidate relocation nodes set as $Q = 20$. For each trial of the local search algorithm, the initial



(a) service start time: 17:30

(b) service start time: 18:00

(c) service start time: 18:30

(d) service start time: 19:00

(e) service start time: 19:30

Figure 4: Covered flows for each station at five service start times

solution is created randomly by selecting p locations and p service start times for MFCLSTP1 (and a single start time for MFCLSTP2). For each problem instance, the best solution obtained among the local optimal solutions starting from 500 initial solutions is chosen as an output by the algorithm. The proposed method was implemented in Microsoft Visual C++ 2008 Professional Edition and experiments were performed on a PC with an Intel

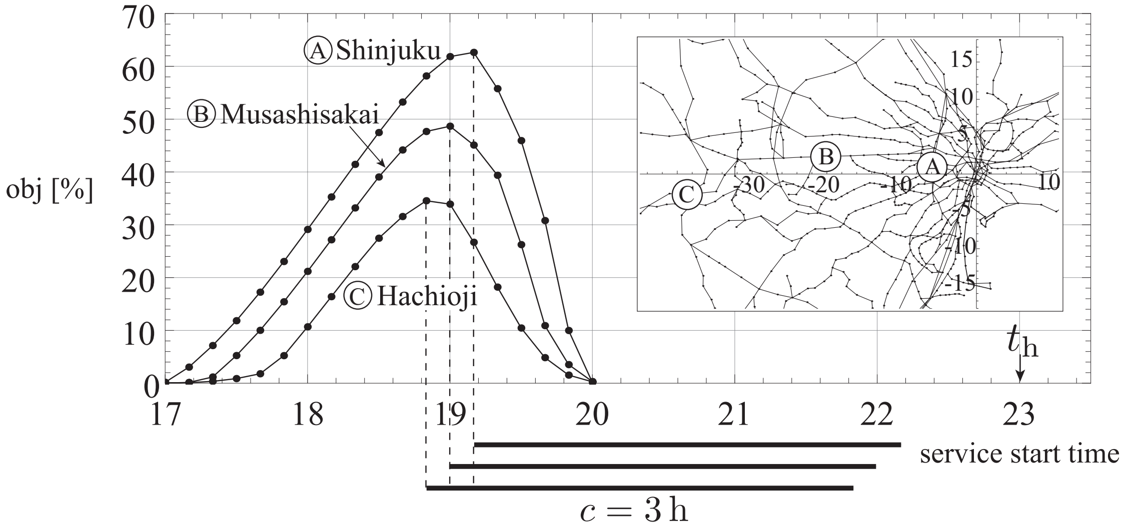


Figure 5: Covered flows over the range of service start times for three stations

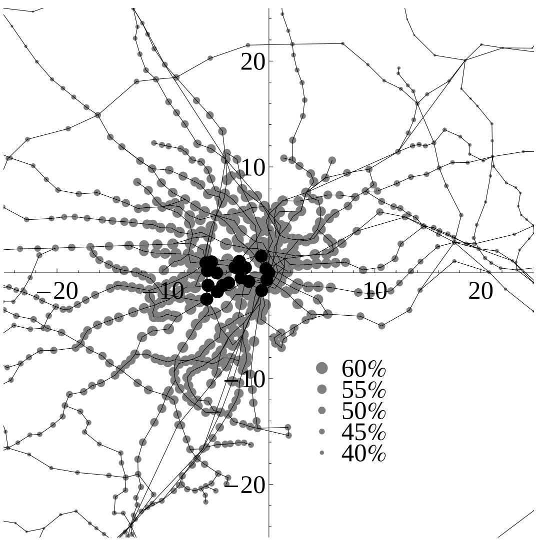


Figure 6: Maximum covered flows for each station

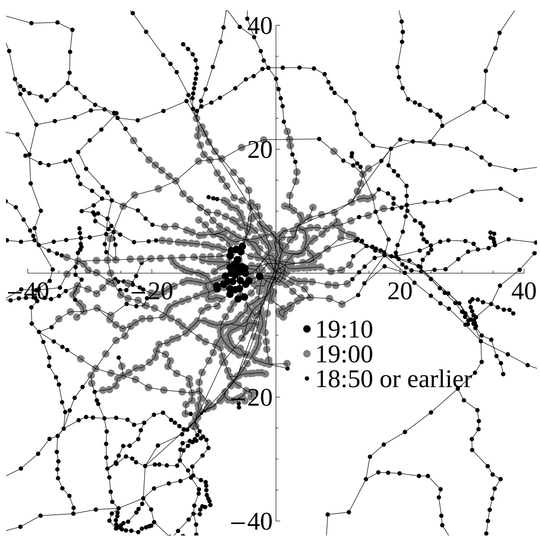


Figure 7: Optimal service start time for each station

Table 2: Computing time for MLS (s)

	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$
MFCLSTP1	219	416	641	886	1,214	2,025
MFCLSTP2	137	264	429	585	780	1,292

Core i7-2620M processor and 8 GB of RAM. Table 2 shows the computational time in seconds required for each problem instance (the total time for obtaining 500 local optimal solutions), and can be seen to be within acceptable limits. Less computational time is required for MFCLSTP2 than MFCLSTP1, since Step 3 must find only one service start time.

Table 1: The 40 stations with highest covered flows

rank	station name	start time	obj [%]	rank	station name	start time	obj [%]
1	Shinjuku	19:10	62.640	21	Hibiya	19:00	61.858
2	Shibuya	19:10	62.521	22	Nijubashimae	19:00	61.856
3	Omote-sando	19:10	62.518	23	Nihombashi	19:00	61.819
4	Harajuku	19:10	62.514	24	Ginza	19:00	61.787
5	Yoyogi	19:10	62.336	25	Takaracho	19:00	61.786
6	Shinjuku-sanchome	19:10	62.237	26	Ochanomizu	19:00	61.785
7	Gaienmae	19:10	62.207	27	Tameike-sanno	19:00	61.784
8	Akasaka-mitsuke	19:10	62.135	28	Shinjuku-gyoemmae	19:10	61.783
9	Tokyo	19:00	62.049	29	Kudanshita	19:00	61.773
10	Yotsuya	19:00	62.027	30	Shin-nihombashi	19:00	61.732
11	Otemachi	19:00	61.997	31	Ogawamachi	19:00	61.728
12	Ichigaya	19:00	61.976	32	Uchisaiwaicho	19:00	61.721
13	Aoyama-itchome	19:00	61.966	33	Nogizaka	19:00	61.718
14	Kojimachi	19:00	61.960	34	Shinanomachi	19:00	61.709
15	Jimbocho	19:00	61.944	35	Tochomae	19:10	61.704
16	Shimbashi	19:00	61.934	36	Takadanobaba	19:10	61.696
17	Sendagaya	19:10	61.929	37	Takebashi	19:00	61.686
18	Hanzomon	19:00	61.877	38	Sakuradamon	19:00	61.658
19	Yurakucho	19:00	61.875	39	Yotsuya-sanchome	19:00	61.653
20	Kokkai-gijidomae	19:00	61.860	40	Kanda	19:00	61.641

Results for common service start time model (MFCLSTP2)

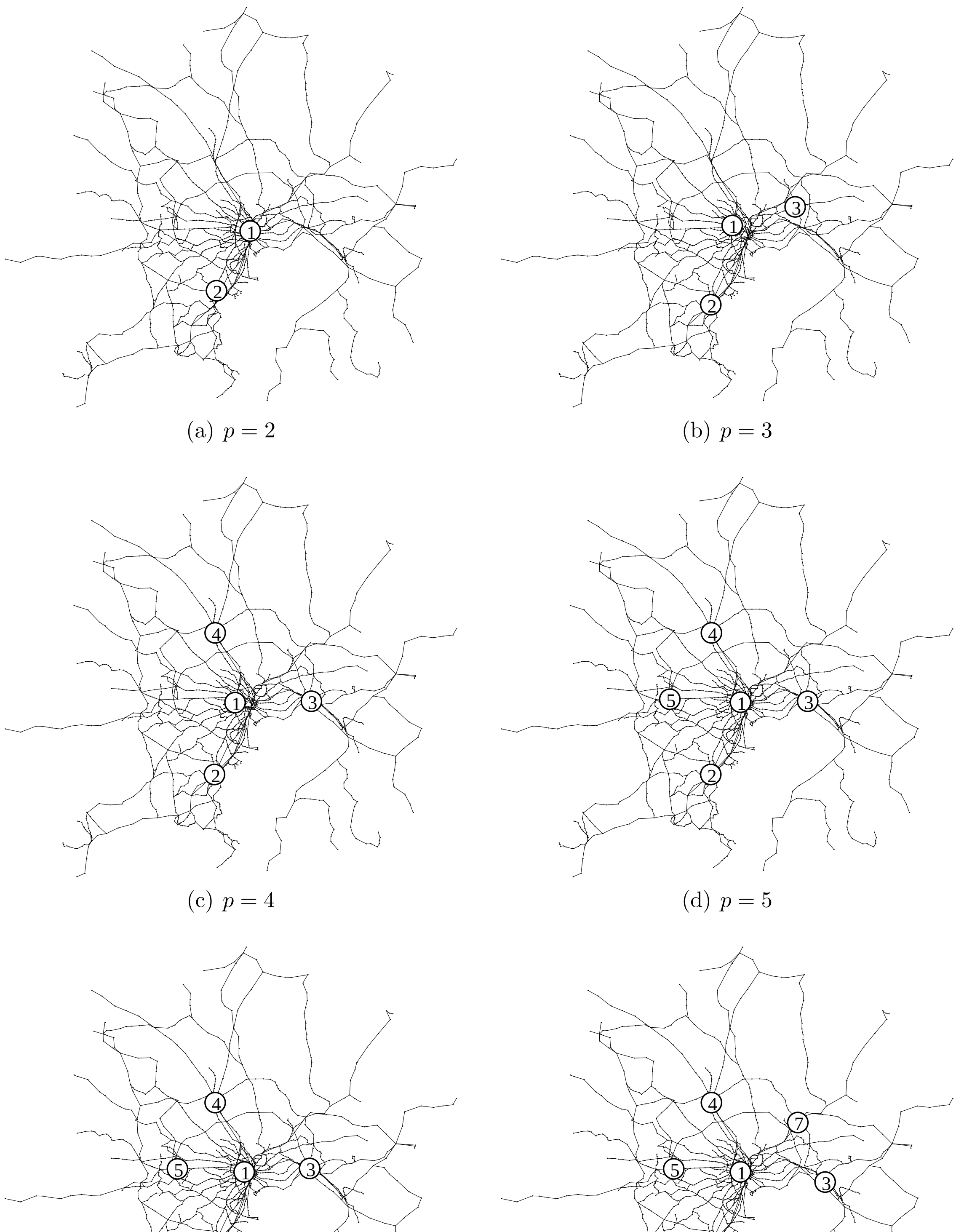
Let us begin by analyzing the results obtained for MFCLSTP2, the simpler of the two models. Table 3 summarizes the solutions and associated objective values obtained for $p = 2$ to $p = 7$. Moreover, in Figure 8, the locations selected for each case are shown, where the number in each circle corresponds to the number in the leftmost column of Table 3.

Facilities are spatially dispersed across the railway network so as to capture different flows. All solutions contain one facility in the central area of Tokyo, with other facilities located at a distance from the central area. As the number of facilities increases, a star-shaped pattern for facility locations is formed, with points in various directions from the center. For $p = 2$, Kikuna is selected in addition to the central location, Jimbocho. The facility at Kikuna can provide a service for those working near Yokohama and having a home in the southern half of the network. Yokohama is selected in all solutions except when $p = 2$.

Note that, as the number of facilities grows, the selected service start time tends to become later. This can be interpreted as follows. Recall that a late service start time increases the accessibility to commuters, while the number of destination stations reachable from the facility by t_{h} decreases. However, when a number of facilities provide services at spatially dispersed locations, a wide area of destination stations reachable by t_{h} are easily covered. This allows facilities to start services later.

Results for independent service start time model (MFCLSTP1)

We next analyze the results obtained for MFCLSTP1, in which service start times are determined independently. Table 4 summarizes the best solutions obtained and correspond-



(a) $p = 2$ (b) $p = 3$

(c) $p = 4$ (d) $p = 5$

(e) $p = 6$ (f) $p = 7$

Figure 8: Comparison of the best solutions obtained for MFCLSTP2

ing objective values for $p = 2$ to $p = 7$, and in Figure 9 the selected locations for each case are shown.

Table 3: The best solutions obtained and corresponding objective values for MFCLSTP2

	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$
1	Jimbocho	Takadanobaba	Shinjuku	Ichigaya	Ichigaya	Ichigaya
2	Kikuna	Yokohama	Yokohama	Yokohama	Yokohama	Yokohama
3		Higashi-matsudo	Daijingushita	Daijingushita	Funabashi	Makuhari
4			Omiya	Omiya	Omiya	Omiya
5				Nishi-kokubunji	Kokubunji	Kokubunji
6					Ebina	Ebina
7						Minami-kashiwa
start time	19:10	19:20	19:20	19:20	19:30	19:30
obj [%]	69.054%	72.821%	76.092%	78.855%	80.742%	82.124%

The location patterns obtained for MFCLSTP1 are different from those for MFCLSTP2. For MFCLSTP1, we observe facilities located close to each other at the center of Tokyo. For example, for the case $p = 2$, Shinjuku and Aoyama-itchome are selected, both of which are in the central area of Tokyo. Conversely, the service start time for the facilities at the two stations differs considerably: 19:30 for Shinjuku and 19:00 for Aoyama-itchome. When focusing on capturing flows only "spatially", two nearby locations are not simultaneously selected, since they capture similar flows. However, when temporal factors are also accounted for, facilities at close locations can capture different types of flows when their service start times are widely separated. Note that in MFCLSTP1, two facilities at close locations can capture over 3% more flows than when compared with the two facilities in MFCLSTP2. Furthermore, the objective value for MFCLSTP1 in $p = 2$ case is only about 0.5% less than that for MFCLSTP2 in $p = 3$ case. Therefore, when considering the cost of each facility, this result may lead a decision maker to prefer two facilities having separated service start times above three facilities with the same service start time.

For $p = 3$, two stations in central Tokyo (Shinjuku and Ochanomizu) plus Yokohama are selected. We again observe that the service start times of the two nearby facilities are separated by 30 min. Yokohama is the largest station in the southern half of the network and captures commuter flows that cannot easily access either of the other two facilities. When $p = 4$, the pattern is similar to that for $p = 3$; three stations located in central Tokyo (Shinjuku, Sugamo, and Iidabashi) plus Yokohama. All three facilities in the central area start their services at different times to capture different flows. Solutions, such as those shown in Figure 9, with clustered facilities having different service start times can be obtained only when we describe the movement of people in spatio-temporal dimensions and consider the service hours for the facilities, in addition to their locations, which cannot be analyzed by using classical static location models.

For the cases of $p = 5$, $p = 6$, and $p = 7$, two or three stations from the central area of Tokyo plus Yokohama (or near Yokohama for $p = 7$) are selected. Other facilities are chosen in regions to the west and north; facility locations in the solutions for MFCLSTP2, such as Omiya, Funabashi, and Nishikokubunji, are also selected in MFCLSTP1. These large stations are located at a distance away from the center and capture flows that may not be able to access facilities in the central area.

Comparison between objective values of MFCLSTP1 and MFCLSTP2

Finally, we compare between the objective values obtained for MFCLSTP1 and MFCLSTP2. Figure 10 shows trends in the percentage of covered flows obtained for the solutions of the two models as the number of facilities increases from $p = 2$ to $p = 7$. Note

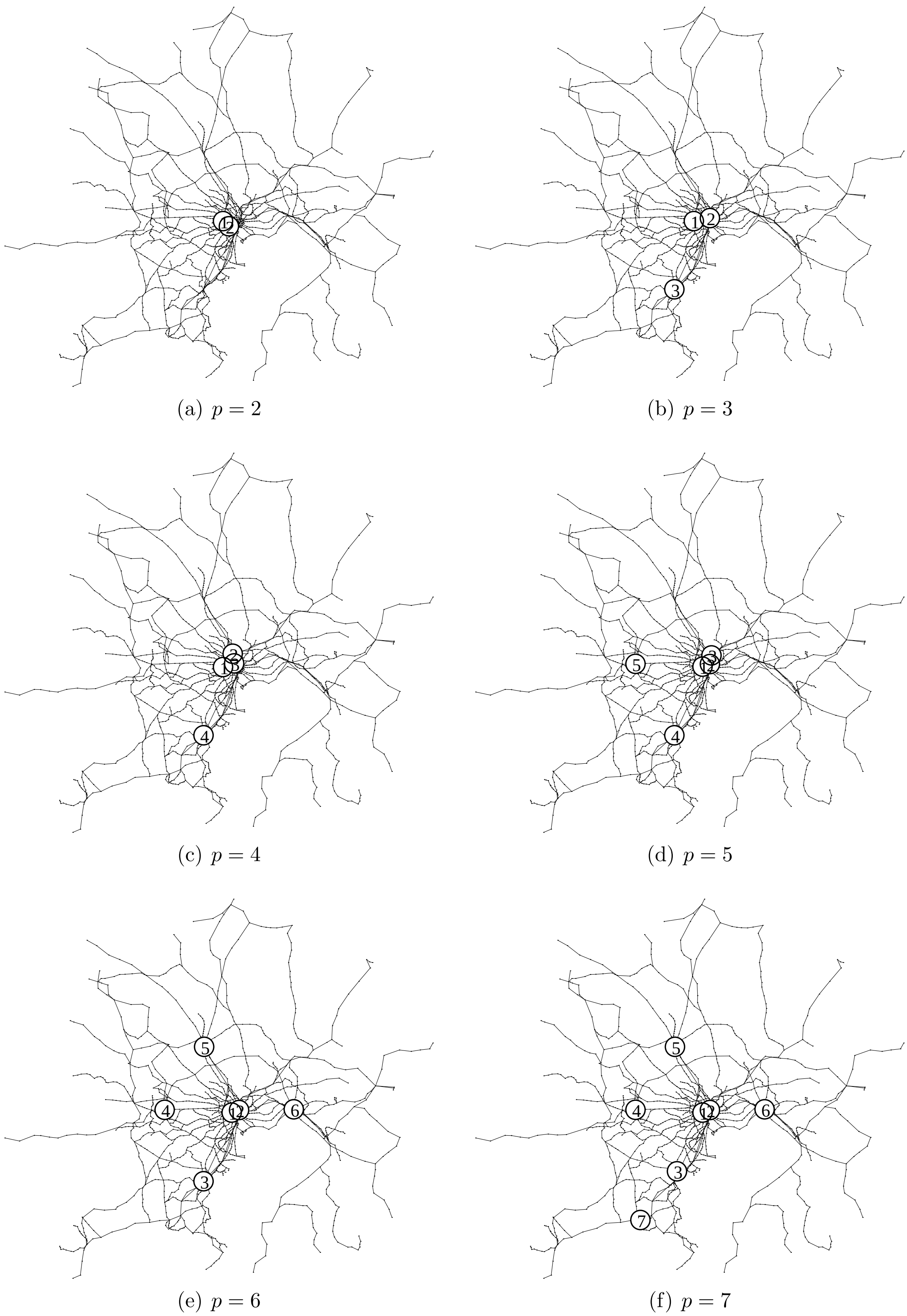


Figure 9: Comparison of the best solutions obtained for MFCLSTP1

that both curves generally exhibit a decreasing marginal coverage with the addition of each facility. This tendency is often encountered in the maximal covering location problem and

Table 4: The best solutions obtained and corresponding objective values for MFCLSTP1

	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$
1	Shinjuku 19:30	Shinjuku 19:30	Shinjuku 19:40	Ichigaya 19:40	Ichigaya 19:40	Ichigaya 19:40
2	Aoyama-itchome 19:00	Ochanomizu 19:00	Sugamo 19:00	Ochanomizu 19:20	Ochanomizu 19:20	Ochanomizu 19:20
3		Yokohama 19:20	Iidabashi 19:20	Nippori 18:50	Yokohama 19:20	Myorenji 19:30
4			Yokohama 19:20	Yokohama 19:20	Nishi-kokubunji 19:30	Nishi-kokubunji 19:30
5				Nishi-kokubunji 19:30	Omiya 19:10	Omiya 19:10
6					Funabashi 19:10	Funabashi 19:10
7						Fujisawa 19:10
obj [%]	72.318%	76.975%	79.715%	81.428%	82.885%	84.083%

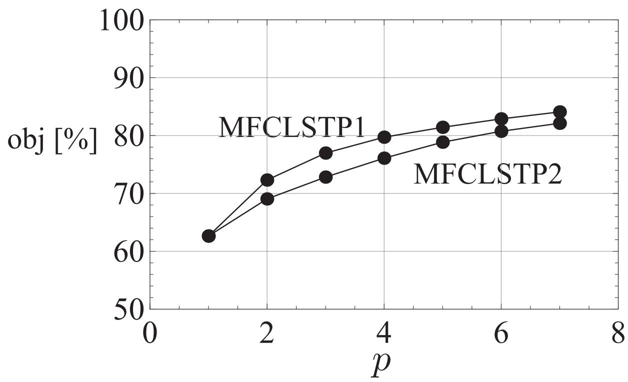


Figure 10: Covered flows for the number of facilities

its variants. The effect of independently determining service start times can be seen from the vertical separation of the two curves. For small p, a difference greater than 3% can be observed. This difference may provide important information when choosing between a common or independent timing strategy.

7. Summary and Future Work

This paper focused on facilities providing fixed hours of service and proposed Maximum Flow-Covering Location and service Start Time Problem (MFCLSTP). MFCLSTP seeks to find both locations and service start times, so as to maximize the number of covered flows. Each flow is specified by an origin-destination node pair and the departure time at the origin node, thus modeling flows in spatio-temporal dimensions. In determining the service start time of each facility, two models were considered: MFCLSTP1, in which service start time of each facility can be independently determined, and MFCLSTP2, in which all services start at the same time. Integer programming formulations of these models are presented as natural extensions of the static facility location problem.

We proposed heuristic algorithms for the two models, which are generalized versions of the Teitz and Bart vertex-exchange algorithm for the p-median problem. The models were applied to a case study of the Tokyo metropolitan railway network composed of 1,804 stations on 125 lines. We built an OD matrix for the commuter traffic flow from census data, and generated input data for MFCLSTP, by dividing this static flow data between each departure time to create dynamic data for each flow. Solutions obtained by the heuristic algorithms can be summarized as follows. When all facilities have to start their service at the same time, the facilities are spatially dispersed and exhibit a star-shaped pattern in order to capture flows that move out from the center in various directions. When the start time for each facility can be determined independently, clustered facilities having different start times appear in central Tokyo. A dynamic facility location model of this type has not been fully developed in previous literature, and the results from the case study revealed characteristics that cannot be obtained using existing location models.

There are a number of important directions for future research. The definition of coverage may change according to the types of services considered. We have considered a service where full consumption is desirable or required, such as cinema, baseball games, and lectures. However, many types of services may be consumed where people stop at a facility for a fixed length of time (e.g., 2 h for dinner at a restaurant). This type of extension can be included in the model by changing the definition of the coverage index, but may require development of different solution algorithms.

A case where we might improve an existing service that is already provided in spatio-temporal dimensions is also of interest. Changing the present location of facilities or service start times requires costs in terms of time and money, but can result in capturing larger demands. This modeling approach is found in classical static location models (e.g., Wang et al. [29]), and therefore this topic can be considered as a spatio-temporal extension of the existing framework. Many variants of the maximal covering location problem have been proposed. Similarly, various generalized models of MFCLSTP can also be constructed. For example, coverage can be redefined when there are two or more accessible facilities for a given flow.

An important research direction is to analyze the problem under more realistic settings. Taguchi [25] constructed a space-time network for modeling the detailed movement of trains in the Tokyo metropolitan area and thereby analyzed interesting real-world problems. By using this approach, detailed analysis can be conducted to evaluate existing facility services or service alternatives. In the present analysis, we build a (static) travel time matrix and use this matrix as an input for the proposed heuristic, allowing us to develop a simple algorithm. Developing more complex solution algorithms is required for solving MFCLSTP when using the approach employed by Taguchi [25].

Finally, the solution algorithms proposed in this paper are straightforward and easy to implement. Extensions under a metaheuristic framework are practically important. Evaluating the quality of solutions obtained for several algorithms is also important.

Acknowledgements

The author would like to thank the two anonymous referees for their very detailed comments and suggestions, which greatly improved the final version of the paper. The author is also very grateful to Dr. Takehiro Furuta for valuable discussions. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), 22710139, 2010-2011.

References

[1] I. Averbakh and O. Berman: Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. *European Journal of Operational Research*, **91** (1996), 495–506.

[2] O. Berman, D. Bertsimas, and R.C. Larson: Locating discretionary service facilities, II: Maximizing market size, minimizing inconvenience. *Operations Research*, **43** (1995), 623–632.

[3] O. Berman, Z. Drezner, and D. Krass: Generalized coverage: New developments in covering location models. *Computers and Operations Research*, **37** (2010), 1675–1687.

[4] J. Brimberg and J.M. Hodgson: Heuristics for location models. In H.A. Eiselt and V. Marianov (eds.): *Foundations of Location Analysis* (Springer, New York, 2011), 335–356.

[5] R. Church and C. ReVelle: The maximal covering location problem. *Papers of the Regional Science Association*, **32** (1974), 101–118.

[6] R. Church and K.L. Roberts: Generalized coverage models and public facility location. *Papers of the Regional Science Association*, **53** (1984), 117–135.

[7] J. Current, S. Ratick, and C. ReVelle: Dynamic facility location when the total number of facilities is uncertain: A decision analysis approach. *European Journal of Operational Research*, **110** (1997), 597–609.

[8] M.S. Daskin: *Network and Discrete Location* (John Wiley & Sons, New York, 1995).

[9] M.S. Daskin, K. Hogan, and C. ReVelle: Integration of multiple, excess, backup, and expected covering models. *Environment and Planning B*, **15** (1988), 15–35.

[10] M.S. Daskin, W.J. Hopp, and B. Medina: Forecast horizons and dynamic facility location planning. *Annals of Operations Research*, **40** (1992), 125–152.

[11] Z. Drezner and G.O. Wesolowsky: Facility location when demand is time dependent. *Naval Research Logistics*, **38** (1991), 763–777.

[12] G. Gunawardane: Dynamic versions of set covering type public facility location problems. *European Journal of Operational Research*, **10** (1982), 190–195.

[13] T. Hägerstrand: What about people in regional science?. *Papers of Regional Science Association*, **24** (1970), 7–21.

[14] M.J. Hodgson: A flow-capturing location-allocation model. *Geographical Analysis*, **22** (1990), 270–279.

[15] K. Honda: Maximum covering location and service start time problem considering users' temporal constraints of daily life —Analyses using questionnaires on personal time schedule—. *Master's Thesis, School of Science for Open and Environmental Systems, Keio University* (2010) (in Japanese).

[16] Institution for Transport Policy Studies: *Census Data for Commuter Traffic in Tokyo Area in 2000* (2002) (in Japanese).

[17] M. Kwan: Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis*, **31** (1999), 187–212.

[18] H.J. Miller: Measuring space-time accessibility benefits within transportation networks: Basic theory and computational procedures. *Geographical Analysis*, **31** (1999), 187–212.

[19] P.B. Mirchandani and A.R. Odoni: Locations of medians on stochastic networks. *Transportation Science*, **13** (1979), 85–97.

[20] S.H. Owen and M.S. Daskin: Strategic facility location: A review. *European Journal of Operational Research*, **110** (1998), 423–447.

[21] D.A. Schilling: Dynamic location modelling for public-sector facilities: A multicriteria approach. *Decision Sciences*, **11** (1980), 714–724.

[22] V. Schmid and K.F. Doerner: Ambulance location and relocation problems with time-dependent travel times. *European Journal of Operational Research*, **207** (2010), 1293–1303.

[23] L.V. Snyder: Facility location under uncertainty: A review. *IIE Transactions*, **38** (2006), 537–554.

[24] Station Database, http://www.ekidata.jp/ (in Japanese).

[25] A. Taguchi: Time dependent traffic assignment model for commuter traffic in Tokyo metropolitan railway network. *Transactions of the Operations Research Society of Japan*, **48** (2005), 85–108 (in Japanese).

[26] K. Tanaka: Dynamic flow-capturing location problem. *Papers of the Annual Conference of the City Planning Institute of Japan*, **41** (2006), 265–270 (in Japanese).

[27] K. Tanaka: Optimal location and opening hours of a single facility which maximally cover flows in a circular city. *FORMA*, **23** (2008), 43–50.

[28] M.B. Teitz and P. Bart: Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations Research*, **16** (1968), 955–961.

[29] Q. Wang, R. Batta, J. Bhadury, and C.M. Rump: Budget constrained location problem with opening and closing of facilities. *Computers and Operations Research*, **30** (2003), 2047–2069.

[30] J.R. Weaver and R.L. Church: Computational procedures for location problems on stochastic networks. *Transportation Science*, **17** (1983), 168–180.

[31] W. Zeng, I. Castillo, and M.J. Hodgson: A generalized model for locating facilities on a network with flow-based demand. *Networks and Spatial Economics*, **10** (2010), 579–611.

Ken-ichi Tanaka
The University of Electro-Communications
1-5-1 Chofugaoka, Chofu
Tokyo 182-8585, Japan
E-mail: `ken1tnk@uec.ac.jp`