

Tanaka, Ken-ichi; Furuta, Takehiro; Toriumi, Shigeki

Article

Railway flow interception location model: Model development and case study of Tokyo metropolitan railway network

Operations Research Perspectives

Provided in Cooperation with:

Elsevier

Suggested Citation: Tanaka, Ken-ichi; Furuta, Takehiro; Toriumi, Shigeki (2019) : Railway flow interception location model: Model development and case study of Tokyo metropolitan railway network, Operations Research Perspectives, ISSN 2214-7160, Elsevier, Amsterdam, Vol. 6, pp. 1-14, <https://doi.org/10.1016/j.orp.2018.11.001>

This Version is available at:

<https://hdl.handle.net/10419/246377>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Railway flow interception location model: Model development and case study of Tokyo metropolitan railway network

Ken-ichi Tanaka^{a,*}, Takehiro Furuta^b, Shigeki Toriumi^c

^a Keio University, Japan

^b Nara University of Education, Japan

^c Chuo University, Japan

^d Department of Administration Engineering, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, 223-8522, Japan

ARTICLE INFO

Keywords:

Discrete location problem
Flow interception location model
Integer programming
Railway passenger flow
Tokyo metropolitan railway network

ABSTRACT

The flow interception location model (FILM) focuses on vehicular traffic in a road network and locates a fixed number of facilities so as to maximize the total flow that can be serviced at facilities along preplanned routes, such as a daily commute to work. This paper develops a version of FILM, the railway flow interception location model (R-FILM), that explicitly focuses on railway passenger flows. For railway users, accessing a facility placed at an origin, destination, or transfer station (ODT station) is easier than visiting a facility at other stations included in the travel path. Accessing a facility at a non-ODT station involves the additional cost of disembarking the train to obtain a service and then boarding another train after consuming the service. R-FILM introduces this railway-specific structure to FILM by introducing two different coverages according to the types of station intercepted for each flow. Concretely, a given flow is called *strongly covered* when at least one facility is located at an ODT station. Similarly, a given flow is called *weakly covered* when no facility is located at an ODT station, but at least one facility is located among stations included in the travel path. We present an integer programming formulation of the proposed R-FILM. Using it, we conduct a large-scale case study of the Tokyo metropolitan railway network, which includes about 1500 railway stations. Input flow is constructed using census data for commuter traffic, with about 100,000 distinct flow paths. Optimal solutions of both models for single- and multi-facility problems are analyzed. In R-FILM solutions, large terminal stations tend to be selected more often than with FILM.

1. Introduction

In traditional network facility location models, a basic assumption is that demand for services is generated at nodes of the transportation network. This type of demand is called *point-based demand*. Consumers at the nodes are supposed to travel from the nodes to a facility to obtain a service. Typical examples are classical facility location models, such as the *p*-median model (Hakimi [14]) and the maximal covering location model (Church and ReVelle [8]). The flow interception location model (FILM) focuses on another type of demand, *flow-based demand*, that is defined in terms of flows of consumers traveling between their own origin and destination nodes. The objective of FILM is to locate *m* facilities in a way that maximizes the number of *covered* customers, where a given flow is defined to be covered (or captured or intercepted) when there exists at least one facility along their predetermined travel path. FILM was independently developed by Hodgson [15] and Berman et al. [4], and has been extensively studied in diverse areas such as

operations research, applied mathematics, transportation engineering, and geography. Due to the wide applicability of the model formulation, FILM and its variants have been applied to optimal locations of vehicle inspection stations (Hodgson et al. [17]), advertising billboards (Hodgson and Berman [16]), alternative-fuel stations (Kuby and Lim [20]), and rail park-and-ride facilities (Horner and Groves [19]). Note that the present paper uses the term “flow interception” instead of “flow capturing” because the former term was used in an early important review article (Hodgson et al. [3]) of which two of the three authors were the first advocates of FILM, namely Hodgson [15] and Berman (et al.) [4].

Most existing uses of FILM assume vehicular traffic flows over road networks. However, the FILM approach is also suitable for planning facility locations over a railway network. In a city with a well-developed railway network, such as Tokyo in Japan, the decision of where facilities should be located among stations is often a critically important problem for both public and private facilities. Thus far, however,

* Corresponding Author.

E-mail address: ken1tnk@ae.keio.ac.jp (K. Tanaka).

<https://doi.org/10.1016/j.orp.2018.11.001>

Received 21 July 2018; Received in revised form 1 November 2018; Accepted 16 November 2018

Available online 17 November 2018

2214-7160/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

relatively few studies have used the FILM framework for railway problems. This paper develops a version of FILM, the *railway flow interception location model (R-FILM)*, in which characteristics of railway flows are explicitly incorporated. When trip takers move between two points within a city using railway networks, they board a train at the origin station (usually the nearest station to the trip origin), and (if necessary) transfer to different railway line(s), and finally disembark at the destination station (usually the nearest station to the trip destination). Therefore, for railway users, accessing a facility located at an origin, destination, or transfer (ODT) station is easier than visiting a facility at other stations on the travel path. Accessing a facility at a non-ODT station involves disembarking the train to obtain service and boarding another train to resume the trip after consuming the service, which is an obvious cost for train users. This structure is peculiar to railway applications and has not been explicitly explored in existing uses of FILM.

R-FILM introduces this railway-specific structure by introducing two different coverage types according to the type of station (ODT or non-ODT) that captures the service for each flow. Concretely, a given flow is called *strongly covered* when at least one facility is located at an ODT station. Similarly, a given flow is called *weakly covered* when no facility is located at an ODT station but at least one facility is located among the stations included in the travel path. This structure is formulated as an integer programming problem by attaching greater weight to strongly covered flows. As a natural extension of this formulation, we present R-FILM with an arbitrary number of coverage levels, and discuss a possible use of R-FILM with three coverage levels.

R-FILM has both public and private applications. One public application may be location of public libraries, in which accessibility to a library for railway flows is an important characteristic. It is convenient for railway users to have a library along their commuting route, especially at an ODT station. Thus, the location problem of maximizing the accessibility of public libraries to users at railway stations can be modeled by the proposed R-FILM. In private applications, selections of stations to open facilities have a strong impact on the attraction of customers to that facility. It is more likely for railway passengers to become an actual user of a facility when it is located at an ODT station than when it is only at other stations along the travel path. R-FILM can provide a basic framework for finding profit-maximizing stations to open private facilities. The construction of the flow interception model with railway-specific structure, as just described above, is one of the major contributions of this paper.

Another important contribution of the present paper is the application of the proposed R-FILM to locational analysis within the railway network of the Tokyo metropolitan area. In this area, nearly eight million commuting trips are taken daily. We construct flow data for FILM and R-FILM from census data covering commuter traffic in Tokyo metropolitan area (The Institute for Transport Policy Studies [34]). The census data, which is based on a large-scale questionnaire, provides the actual railway routes used from origin to destination stations for each trip. The corresponding flow volumes are available in the census data.

The number of stations is 1,470, each of which is taken as a candidate location for a facility. The number of distinct flow paths is 100,207. Together, these sizes make this one of the largest problem instances ever tackled by FILM or a variant. To our knowledge, it is by far the largest FILM instance using railway passenger flow.

The following sections proceed as follows. In Section 2, we review the literature on FILM from various viewpoints, including variants and extensions, types of facilities assumed, and applications for both road and railway flow data. Then, Section 3 describes considerations peculiar to railway applications and develops R-FILM with two coverage types. We also develop a variant of R-FILM with an arbitrary number of coverage levels and describe possible applications of the proposed framework to real-world situations. Section 4 explains the Tokyo metropolitan railway network, which is used in the case study, and analyzes spatial flow patterns involving large terminal stations. Then, in

Section 5, we investigate the performance of each station when a single facility is introduced at each station. Objective values of both the original FILM and the proposed R-FILM are compared. Section 6 explores multi-facility problems using both models, and compares similarities and differences between optimal solutions. Finally, in Section 7, we conclude the paper and discuss future research directions.

2. Literature review

The original FILM was independently developed by Hodgson [15] and Berman et al. [4], and many types of variants and extended models have been proposed since then. Extensions of the basic FILM include a model in which deviation from the predetermined trip path is allowed (Berman et al. [2]), a model in which hazard avoidance is maximized by detecting and removing hazardous vehicles as early in trips as possible (Hodgson et al. [17]), and a multi-counting model in which consumers can be covered multiple times when there are several facilities along their paths (Avebakh and Berman [1]). More details of the early development of FILM and variants can be found in the review article by Berman et al. [3].

The flow-intercepting modeling approach continues to attract researchers' attention. Kuby and Lim [20] modified the basic FILM and applied it to locating hydrogen-refueling stations. They introduce a driving range parameter and formulate the problem of maximizing the number of trips that can potentially be refueled with p facilities, assuming that two or more stations have to be present for a single flow to be refueled. Zeng et al. [36] proposed an important generalized model, what they call a GFIM (generalized flow-interception model), that considers the location at which each flow is covered. They reviewed the existing FILMs and showed that most of them can be described as special cases of the proposed general model. Tanaka and Furuta [32] extended the basic FILM to allow a decision maker to select the size of facilities among different-size alternatives, where larger (but more costly) facilities can induce larger deviation from the shortest path and thereby attract more demand. Chung and Kwon [7] introduced multi-period planning model in which decisions about facility locations are made over several years. Focusing on the location of weigh-in-motion systems, tollbooths, and security checkpoints, Marković et al. [25] considered a situation where targeted users behave non-cooperatively by changing their travel paths to avoid fixed facilities. Zeng et al. [36] and Marković et al. [25] have reviewed the recent developments.

The FILM approach is also suitable for planning facility locations in cities that have a railway network. When providing facility services in such cities, choosing the railway stations for facilities is often critically important. Thus far, however, the number of FILMs focusing explicitly on railway problems is low. The first model involving railway problems seems to have been the model by Horner and Groves [19], which considers optimal locations of rail park-and-ride facilities. The purpose of the model is to identify desirable locations of facilities so as to intercept vehicles as early in their journeys as possible, thereby decreasing vehicular traffic on congested roadways. Tanaka [27] focuses on the situation where commuter flows on the way back home access a facility to obtain a service for a fixed duration. The problem is to determine both the location of facilities and their service start times so as to maximally cover flows that can consume service at one of the facilities after work and be home by a given time. The problem was applied to the Tokyo metropolitan railway network using commuter flow data. Two variants of this model have been proposed in Tanaka and Furuta [31], and Tanaka and Toriumi [33].

A few attempts have been made to incorporate railway-specific factors into the FILM framework. Tanaka and Furuta [28] considered the location of large and small facilities on railway stations. In that paper, a large facility can intercept flows whose travel path includes the station at which the facility is located, whereas a small facility can intercept flows only if it located at either the origin or destination station. They applied the model to locational analysis of the Yamanote

Line, which is operated by East Japan Railway Company and a busy circular line that services the center of Tokyo. Tanaka and Furuta [30] focused on the situation where railway flows can only be covered at stations where the train stops, with some stations passed by some trains. They set the coverage coefficient to reflect real stopping patterns of trains for each flow, and conducted locational analysis of Keio Railway network. Tanaka and Furuta [29] focused on railway passenger flows on a single railway line, and introduced two different levels of accessibility to facilities: located at either end station or located along the path. They introduced two types of coverage and assume that a given flow is fully covered when there is a facility at the origin or destination station and partially covered when the facility is at a station along the path. The R-FILM proposed in this paper can be regarded as an extended version of this model by formulating the model over a general complex railway network. To formulate a model over an arbitrary network, R-FILM focuses on transfers to other railway lines and assumes that facilities located at transfer stations (in addition to origin and destination stations) are also easily accessible. To extend the model further, we also propose a version of R-FILM that allows an arbitrary number of coverage types and illustrate some possible applications to real-world scenarios.

The case studies in all three papers [28–30] used simple networks such as a path (single railway line), a cycle, and a tree network. In addition, the networks are small, having only dozens of stations. This paper applies the model to a much larger railway network, that of the Tokyo metropolitan area in Japan. In this network, 136 different railway lines are operated by 37 railway companies. The number of distinct flow paths is high (100,207), and the number of stations is 1,470, each of which is considered a candidate location for a facility. This analysis is the first large-scale case study in the FILM framework focusing on the behavior of railway users. To our knowledge, the existing case studies of FILMs applied to large-scale real data were vehicle traffic flow data of Edmonton, Canada, for the morning peak period in 1989 and the afternoon peak period in 2001. The number of different flow paths in 1989 (resp., 2001) data was 23,958 (resp., 16,488). See Hodgson et al. [18] and Zeng et al. [36] for applications of the Edmonton data to flow interception problems. Another real-world example is by Kuby et al. [21] that applies the flow-refueling location model for optimization of hydrogen stations in Florida. In a more recent study, Chung and Kwon [7] proposed a multi-period planning of flow-refueling location problem and applied it to real traffic flow data from the Korean Expressway network in 2011. They combined the flows from all paths between the same origin/destination pair and assume drivers always choose the shortest path. The total number of different flow paths was 104,652, which is near the size of the present case study. When dealing with large-scale data, some preprocessing (e.g., data aggregation) is often required. However, as we will show in the numerical example section, we can solve the R-FILM instances of the Tokyo metropolitan area using a mathematical optimization solver. The fact that optimal solutions are obtained within a moderate amount of time with no data modification indicates the usefulness of the proposed framework.

Another important related problem is the design of rapid transit networks. This is a hard problem involving several players, multiple objectives, sizeable costs, and a high level of uncertainty (Laporte and Mesa [23]). One of the important objectives in the design of rapid transit networks is to maximize trip coverage for a planned network. For example, Laporte et al. [24] considered the construction problem of a rapid transit alignment with the objective of maximizing the total origin–destination demand (station-to-station ridership) coverage. In another example, Körner et al. [22] formulate a facility location problem in a mixed planar-network space. They assume that traveling along a given network is faster than traveling within the plane (assuming a fixed rate of planar travel according to Euclidean distance). The objective is to find entry and exit points for the network that maximize the number of covered pairs, where a pair of points in the

plane is called covered if the travel time using the network is not more than the travel time using planar travel. Maximization of trip coverage has been used as a basic design criterion in many contexts (e.g., García-Archilla et al. [11]; Gutiérrez-Jarpa et al. [13]; Gutiérrez-Jarpa et al. [12]). For further information, see the related book chapter of Laporte and Mesa [23]. Although maximizing trip coverage is considered in the design of rapid transit networks, the notion of trip coverage is different from that considered in the present study. The above class of problems focuses on designing the shapes of networks to capture as many potential trips as possible. For our current problem, the objective is to derive the optimal subset of stations at which to open a fixed number of facilities such that it maximizes the flow coverage, where the shape of the railway network and flow data over it are fixed in advance.

The proposed model can also be regarded as a FILM version of the maximal covering location model with variable levels of coverage, which is an extension of the maximal covering location problem that was first proposed in Church and Roberts [9]. For example, the model can be applied to locational analysis of supermarkets. Demands occurring within a distance (equivalently, time) of T_1 from the shop can be considered as strongly covered, while those occurring within distance T_1 to T_2 (where $T_2 > T_1$) can be regarded as weakly covered, with lower visit frequency for the supermarket. Covering models with more than one coverage level are also treated in Berman and Krass [6] for discrete models and in Drezner et al. [10] for continuous models, both of which assume that the strength of coverage decreases with increased distance to the nearest facility. More information on gradual coverage models for point-based demands can be found in Berman et al. [5].

The proposed model can be regarded as a FILM version of a multiple coverage model by considering flows having facilities at ODT stations to be strongly covered and flows having a facility at some station along the flow but not at an ODT station to be weakly covered. The pickup location model proposed by Zeng et al. [35] introduced location-dependent coverage levels. They considered three different scenarios in which the coverage level is strongest at (i) the origin node, (ii) the destination node, and (iii) a particular node midway along the path, with coverage decreasing as the distance from the specified node increases. All three models assume that the desirability of having a facility along the path changes continuously along the path. Focusing on railway applications, the proposed R-FILM evaluates accessibility based on station type, thereby dealing with discontinuous change of coverage levels along the flow path, a structure that has not been explicitly considered before.

3. Model description and formulation

Firstly, this section reviews the original FILM. Then, we introduce the proposed R-FILM and formulate it as an integer programming problem.

3.1. FILM as originally proposed

Hodgson [15] and Berman et al. [4] independently developed the original flow interception location problem. The aim in this problem is to maximize the number of consumers who encounter at least one facility along their predetermined trips. Let us consider a network with node set K . We denote by Q the set of nonzero flow paths between origin–destination pairs of the network. The volume of flow $q \in Q$ is given by f_q . The objective is to locate m facilities across the nodes in the network so as to maximize the total flow covered. To formulate the FILM as an integer programming problem, the following two sets of decision variables are introduced.

$$x_k = \begin{cases} 1 & \text{if a facility is located at node } k \in K \\ 0 & \text{otherwise} \end{cases}$$

$$y_q = \begin{cases} 1 & \text{if flow } q \in Q \text{ is covered} \\ 0 & \text{otherwise} \end{cases}$$

The problem formulation is given as follows.

$$\text{maximize} \quad \sum_{q \in Q} f_q y_q \quad (1)$$

$$\text{subject to} \quad \sum_{k \in K} x_k = m \quad (2)$$

$$y_q \leq \sum_{k \in K} \alpha_{qk} x_k \quad \forall q \in Q \quad (3)$$

$$x_k \in \{0, 1\} \quad \forall k \in K \quad (4)$$

$$y_q \in \{0, 1\} \quad \forall q \in Q \quad (5)$$

In this formulation, the parameters α_{qk} are the following.

$$\alpha_{qk} = \begin{cases} 1 & \text{if node } k \in K \text{ is included in the travel path of flow } q \in Q \\ 0 & \text{otherwise} \end{cases}$$

The objective function (1) seeks to maximize the number of customers traveling on the network who have at least one facility on their path. Constraint (2) stipulates that the number of facilities to be located is m . Constraints (3) state that flow q cannot be covered unless at least one facility is located in the nodes contained in q . Note that the right-hand side of inequality (3) represents the number of facilities that can cover flow q ; this value should be more than one in order for flow q to be covered ($y_q = 1$). Constraints (4) and (5) are standard binary constraints. The flow coverage variable y_q can be relaxed to $0 \leq y_q \leq 1$ instead of $y_q \in \{0, 1\}$ because of the form of constraints (3) and objective function (1).

3.2. R-FILM

We develop R-FILM by incorporating features specific to railways. To travel between two different locations within a city using a railway network, a traveler first boards a train at an origin station, and (if necessary) makes a number of transfers to different railway lines, and finally disembarks at a destination station. Because of this, train users can more easily access facilities at ODT stations than at non-ODT stations, as shown in Fig. 1 (a) and (b). This railway-specific structure cannot be described using the original FILM, which assumes that a given flow is either covered or not covered at all.

To describe the above-mentioned difference in accessibility to a facility, we introduce two types of coverage, with the type according to whether the flow is covered by an ODT station. To explain the difference between the two types, we use the network shown in Fig. 1, in which there are three different railway lines, namely lines A, B, and C. The flow path shown in Fig. 1 runs from station A1 to station A3 along line A, then from B3 to B5 along line B, and finally from C3 to C4 along line C. Because this flow changes train at stations A3 and B5, these two stations are transfer stations for this flow. In Fig. 1 (a), a facility

(denoted F) is located at B5, which is one of the two transfer stations for the flow. Fig. 1 (b) shows another situation in which a facility is located at B4, now requiring the commuter to leave the train to obtain the service. To reflect this difference in accessibility to a facility between ODT and non-ODT stations, we define the flow in Fig. 1 (a) to be strongly covered whereas that in Fig. 1 (b) is only weakly covered.

Note that the notion of transfer stations is flow specific: a single station (e.g., A3 or B5) can be either a transfer station or a non-transfer station depending on the flow being studied. For example, let us focus on a flow whose origin station is A1 and destination station is A5. This flow can go directly from A1 to A5 using line A, making station A3 a non-transfer station for this flow, unlike the flow case shown in Fig. 1. Also, as we will explain in the next section, connecting stations such as A3 and B3 are considered as one station and treated as a single candidate for facility location.

To formulate R-FILM with two types of coverage, we introduce the following additional notation.

θ_s : weight given to strongly covered flows

θ_w : weight given to weakly covered flows ($\theta_w < \theta_s$)

α_{qk}^s : binary coverage coefficient, 1 if station $k \in K$ is any of origin, destination, or transfer station for the flow $q \in Q$

α_{qk}^w : binary coverage coefficient, 1 if station $k \in K$ is on the travel path of flow $q \in Q$ but not located at an origin, destination, or transfer station.

To differentiate the two types of station, we introduce the corresponding coverage variables.

y_q^s : binary variable, takes 1 if flow $q \in Q$ is strongly covered, and 0 otherwise

y_q^w : binary variable, takes 1 if flow $q \in Q$ is weakly covered, and 0 otherwise

Using the above notation, the proposed R-FILM can be formulated as the following integer programming problem.

Railway flow interception location model

$$\text{maximize} \quad \theta_s \sum_{q \in Q} f_q y_q^s + \theta_w \sum_{q \in Q} f_q y_q^w \quad (6)$$

$$\text{subject to} \quad \sum_{k \in K} x_k = m \quad (7)$$

$$y_q^s \leq \sum_{k \in K} \alpha_{qk}^s x_k \quad \forall q \in Q \quad (8)$$

$$y_q^w \leq \sum_{k \in K} \alpha_{qk}^w x_k \quad \forall q \in Q \quad (9)$$

$$y_q^s + y_q^w \leq 1 \quad \forall q \in Q \quad (10)$$

$$x_k \in \{0, 1\} \quad \forall k \in K \quad (11)$$

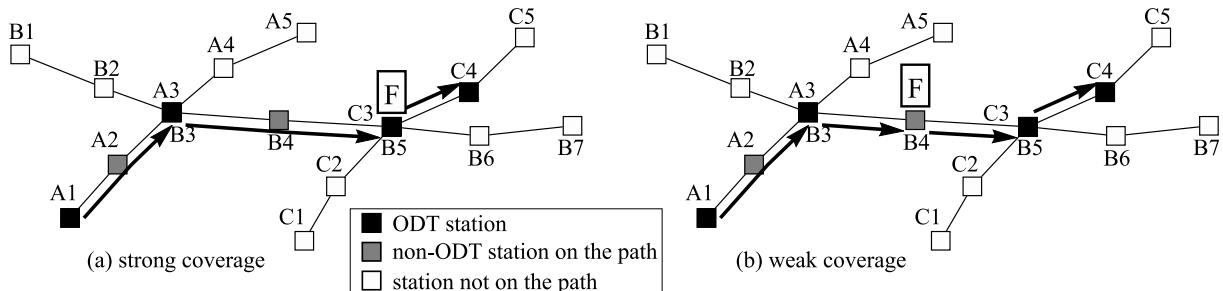


Fig. 1. Different coverages based on type of station covering a flow.

$$y_q^S, y_q^W \in \{0, 1\} \quad \forall q \in Q \quad (12)$$

The objective is to maximize the total covered flow volume, weighted by the desirability levels θ_S and θ_W . Constraint (7) ensures that the number of facilities located equals m . Constraints (8) and (9) define the conditions in which flow q is strongly covered and weakly covered, respectively. Constraints (10) stipulate that flow q is either strongly covered or weakly covered or not covered at all. To satisfy this constraint, when both variables y_q^S and y_q^W cover flow q , we set $y_q^S = 1$ and $y_q^W = 0$ due to the greater contribution to the objective value. Constraints (11) and (12) are standard binary constraints. Similar to the original FILM, flow coverage variables y_q^S and y_q^W can be relaxed to lie between zero and one because of the form of the constraints (8), and (9) and objective function (6). It is important to point out that the proposed R-FILM is an extension of the original FILM; when $\theta_S = \theta_W$, the problem is reduced to the original FILM.

It should be noted that the structure considered here can also be described using GFIM proposed by Zeng et al. [36], which is a location-allocation extension of the original FILM. However, the GFIM formulation is not suitable for describing the proposed variable coverage structure. For each flow q , we have to assign θ_S or θ_W for all of the stations on the travel path according to the type of station. This approach unnecessarily increases the number of variables, thereby incurring computational burden. The proposed formulation employs only two types of variables for each flow, which is more suitable for our application.

As a natural extension of the proposed model, R-FILM with an arbitrary number of coverage types can be considered. Let us denote by T the set of all types of coverages, and denote by θ_t the weight of each coverage type, where we assume that $\theta_1 > \theta_2 > \dots > \theta_{|T|}$. Then, we can formulate the problem as follows.

Railway flow interception location model with multiple coverages

$$\text{maximize} \quad \sum_{t \in T} \theta_t \sum_{q \in Q} f_q y_{qt} \quad (13)$$

$$\text{subject to} \quad \sum_{k \in K} x_k = m \quad (14)$$

$$y_{qt} \leq \sum_{k \in K} \alpha_{qkt} x_k \quad \forall q \in Q, \forall t \in T \quad (15)$$

$$\sum_{t \in T} y_{qt} \leq 1 \quad \forall q \in Q \quad (16)$$

$$x_k \in \{0, 1\} \quad \forall k \in K \quad (17)$$

$$y_{qt} \in \{0, 1\} \quad \forall q \in Q, \forall t \in T \quad (18)$$

Here, the coverage coefficient α_{qkt} takes one if station $k \in K$ is a candidate location for flow $q \in Q$ that can provide cover at level $t \in T$, and 0 otherwise. Also, coverage variables are extended to y_{qt} for each $q \in Q$ and $t \in T$. The next section applies R-FILM with two types of coverage to the case study of the Tokyo metropolitan railway network.

Before that, we consider how to employ the three coverage-level model to real-world scenarios. One candidate is to further classify stations that can provide strong coverage into two types: origin and destination stations for one group, and transfer stations for the other group. The basic idea for this setting is as follows. A commuter can visit a facility at the origin station before starting a trip or can access the facility at the destination station after completing the travel. In this case, flows having a facility at the origin or destination station have strong coverage, while those at transfer stations (but not origin or destination) have medium coverage, and those at other stations along the travel path have weak coverage. Another possibility is to consider stations that are not included in the original path but are accessible if some level of deviation from the path is allowed. A facility located within a

reasonable deviation distance for a given flow can provide weak coverage, while a facility on the path but not at an ODT station provides medium level coverage. Yet another possibility may be to differentiate between the desirability of having a facility at an origin station instead of a destination station. A model with four or more coverage levels could be applied by assuming that the coverage strength decreases according to the level of deviation from the original travel path. Depending on the situations and types of facilities considered, R-FILM with multiple coverage levels may be effectively used for the railway application.

4. Tokyo metropolitan railway network and flow data

The latter part of the paper is devoted to a case study of the rail network for the Tokyo metropolitan census area [34]. In addition to Tokyo, the census area consists of Kanagawa, Saitama, and parts of Chiba, Ibaraki, Gunma, Tochigi and Yamanashi Prefectures. The census area has a well-developed railway network, and many people use railway lines for commuting. In fact, the total flow volume for commuting trips that use railway lines in this area is roughly 7.9 million people. Also, a remarkable feature of commuting trips is that, unlike private trips, the route is always the same. For these reasons, commuting trips are important when determining facility locations over the railway network in the target area. Work-commuting trips and school-commuting trips by the 7.9 million commuters account for 79.7% and 20.3%, respectively, of the total flow volume. Fig. 2 shows the target network used in the analysis. The network was constructed from stations included in the census data for the Tokyo metropolitan area in 2010 (The Institute for Transport Policy Studies [34]). The target area of the census data comprises those municipalities that satisfy the following conditions: (i) the travel time to Tokyo station is no more than two hours, and (ii) the number of commuters whose destinations are within the 23 wards of Tokyo exceeds 500 and the proportion of such commuters exceeds 3% of all commuters in the municipality. In total, 37 different railway companies operate 136 railway lines. The total number of railway stations is 1,470, where sets of stations among which transfer by foot is possible are counted as a single station. The set of all 1470 stations is taken as the set of candidate locations of facilities for the location analysis in the next section.

Fig. 3 shows an enlarged view of the central part of the network. Six stations indicated by number one through six in Fig. 3 are Shinjuku, Ikebukuro, Shibuya, Yokohama, Akihabara, and Tokyo, which are the busiest stations and act as business and commercial centers. The circular railway line seen in Fig. 3 is the Yamanote Line, which is operated by the East Japan Railway Company (JR East). It is one of the busiest

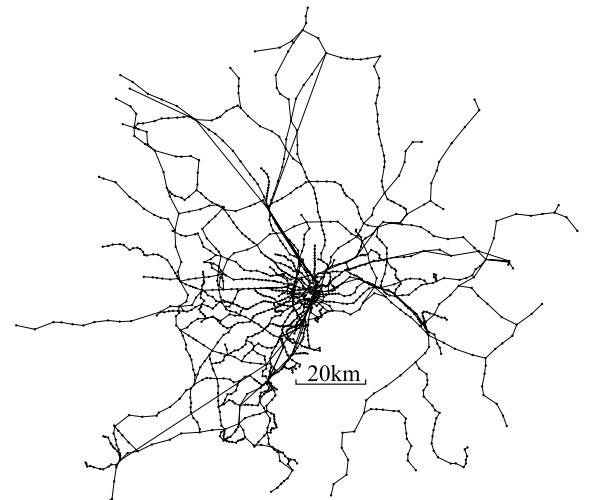


Fig. 2. Tokyo metropolitan railway network.

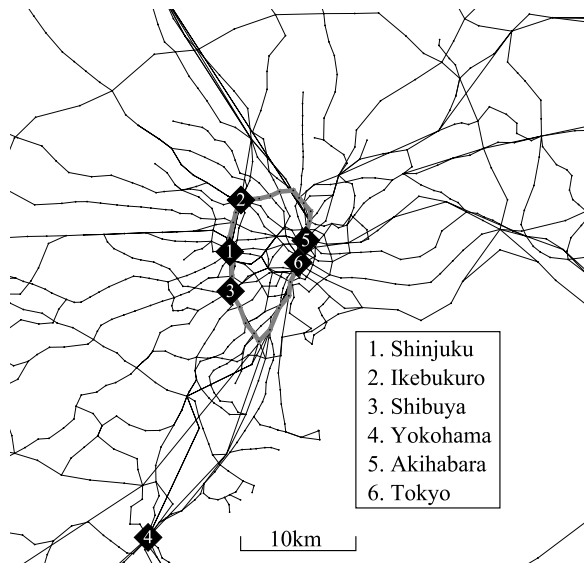


Fig. 3. Enlarged view of the central area.

and most important lines in this area, connecting most of Tokyo's major terminal stations, including five of the six stations shown in Fig. 3. All six stations are shown to be good locations by both single- and multi-facility analysis, as we will see in later sections.

Next, we explain how to construct railway flow data from census data. The census dataset is constructed from a large-scale questionnaire survey about commuting, which include both work-commuting trips and school-commuting trips. The data provides 100,207 different railway flow paths and the corresponding flow volume for each flow. Each flow data item is composed of the set of railway lines used from origin to destination stations, and the corresponding flow volume. The total flow volume for all 100,207 flows is about 7.9 million (7,895,066) one-way commuting trips. Note that there can be several different flow paths for the same origin-destination station pairs. We constructed coverage coefficients α_{qk}^w and α_{qk}^s and flow volume f_q from this data. Concretely, for each flow, the ODT stations are recorded as candidate locations for strong coverage, and other stations in the path are included as candidate locations for weak coverage in the input file. All values of f_q are recorded in a separate input file.

An example of a flow path is shown in Fig. 4. The origin station of this flow path is Hiyoshi station on the Tokyu Toyoko Line and the destination station is Ueno station on the JR Yamanote Line. This flow uses three different railway lines: Tokyu Toyoko Line from Hiyoshi to Musashikosugi station, JR Yokosuka Line from Musashikosugi to Tokyo

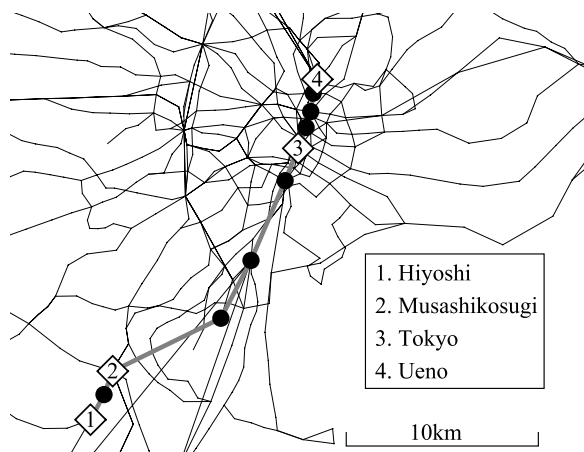


Fig. 4. Example of commuting travel path.

station, and JR Yamanote Line from Tokyo to Ueno. There are four candidate stations for strong coverage (origin [Hiyoshi], destination [Ueno], and the two transfer stations [Musashikosugi and Tokyo]) as shown by diamonds with a station number inside, as in Fig. 4. In Fig. 4, stations that could provide weak coverage are shown by black disks. It is required to locate a facility at one of the four ODT stations for this flow to be strongly covered. Similarly, at least one facility has to be located among the black-disk stations for this flow to be weakly covered.

Fig. 5 (a) and (b) illustrate the total number of trips originating from and terminating at each station, respectively, for all 100,207 flow paths. Fig. 5 (a) shows that origins are widely dispersed over the entire network and tend to have large values at some distance away from the outer perimeter of the JR Yamanote Line. In the other direction, it can be seen from Fig. 5 (b) that the destinations are much more concentrated near the center of Tokyo compared with origin stations, and several stations along the JR Yamanote Line have high values, particularly the five stations shown in Fig. 3. Note that the travel routes of morning and evening trips are the same because the above data are for commuting trips. Therefore, we will not differentiate these two cases in the following analysis. Table 1 shows the travel time distribution for all flows, which are created from the aggregated data in the census data for commuter traffic in the Tokyo area (The Institute for Transport Policy Studies [34]). The table columns show, in order, the range for the travel time (t_{range}), the corresponding flow volume (flow_vol), the relative volume in percentage (rate), and the cumulative percentage (r_{cum}). The average commuting travel time is 70 min. More than 60% of all commuters spend at least 60 min on their trips, and more than 20% of travelers spend 90 min or longer.

Next, we focus on the number of railway lines each commuter uses to move between the origin and destination stations. These data are quite important since this information is related to the number of transfers for each flow. For example, use of three different railway lines requires commuters to change trains twice (thus involving two transfer stations). The data are summarized in Table 2, which shows (in order) the number of railway lines (nline), the corresponding flow volume (flow_vol), the volume in percentage (rate) and the cumulative percentage (r_{cum}). The mode of the number of railway lines is two (one transfer), accounting for more than 40% of all trips. The second-most frequent number of lines is three (two transfers), accounting for more than a quarter for all trips. The average number of railway lines used for all trips is 2.27.

Next, we focus on six stations: Shinjuku, Ikebukuro, Shibuya, Yokohama, Akihabara and Tokyo. These stations are major transport hubs and each connects to several different railway lines. Also, these stations are some of the busiest stations as business and commercial centers. It is important to analyze characteristics of flow patterns concerning these stations since they are frequently selected in optimal solutions, as we will show in the next section. Table 3 illustrates the total flow volume that can be covered when a single facility is located at each of the six stations. The numbers in the first, the second and the third row show the volume of strongly covered flows (SCF), the volume of weakly covered flows (WCF), and the combined volume of SCF and WCF, respectively. The latter value is equivalent to the objective value of the original FILM when one facility is located at each station. Shinjuku, which is the largest station by traffic volume in the Tokyo area, has a much larger value than those of the other stations except for the value of WCF for Akihabara. A group of three stations, Ikebukuro, Shibuya and Yokohama, have the next largest values after Shinjuku. Akihabara is slightly different from the other stations. While Akihabara has a large value under the original FILM objective, almost comparable to the three stations above, the SCF is much smaller than for those stations. This indicates that although it performs well under the original FILM objective, Akihabara is less attractive for opening a facility due to the lower accessibility to many commuters from the R-FILM viewpoint.

Next, we investigate the spatial flow patterns for the six stations. For

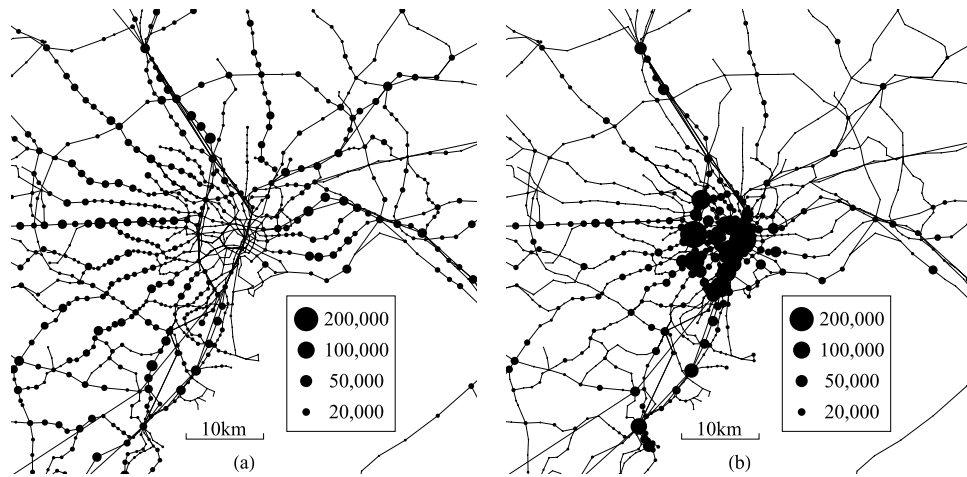


Fig. 5. Distribution of origins and destinations.

Table 1

Travel-time distribution for all flows.

t_range [min]	flow_vol	rate [%]	r_cum[%]
0–29	252,670	3.317	3.317
30–59	2,458,031	32.270	35.587
60–89	3,141,076	41.237	76.824
90–119	1,354,407	17.781	94.605
120–	410,952	5.395	100.000

Table 2

Distribution of the number of railway lines used for all flows.

nlines	flow_vol	rate [%]	r_cum[%]
1	1,727,304	21.878	21.878
2	3,267,517	41.387	63.265
3	2,125,089	26.917	90.182
4	611,575	7.746	97.928
5	127,869	1.620	99.548
6 or more	35,712	0.452	100.000

each station, we extract those flows (from among all 100,207 flows) whose travel path includes the target station (e.g., Tokyo), which is marked by the white disk in Figs. 6–11. Figs. 6–11 show the total traffic volume passing through each of the 1470 stations for the travel paths containing the target station. In each case, the map on the left is for SCF only and the map on the right is for SCF plus WCF. To explain the meaning of these figures in more detail, we focus on the case of Tokyo station shown in Fig. 11. The example flow path shown in Fig. 4 transfers from the JR Yokosuka Line to the JR Yamanote Line at Tokyo station, making Tokyo station an ODT station for that flow. The volume of this flow path is counted at each of the 11 stations shown in Fig. 4 from the origin station (Hiyoshi) to the destination station (Ueno). By counting traffic volume in a similar manner for all of the flow paths that use Tokyo station as an ODT station, we obtain Fig. 11 (a). Fig. 11 (b) is obtained similarly by targeting all the flows that use Tokyo station as either a non-ODT or an ODT station. The value for each station is represented by the area of the disk (the value of the target station is

represented by the area of the white disk).

Figs. 6–11 show that the values differ greatly among the 1470 stations: some of the stations near the target (white) station have values that are quite large whereas others have values that are either negligibly small or zero. It can be seen in Fig. 6 that the flow volumes involving Shinjuku are especially large in the western area of Tokyo. Figs. 7 and 8 illustrate that Ikebukuro and Shibuya attract flows from the northwest area and southwest area, respectively. These results coincide with generally accepted impressions in the Tokyo area that Ikebukuro and Shibuya are gateway stations for Saitama Prefecture and Kanagawa Prefecture, located north and south of Tokyo, respectively. Next, the flow patterns involving Yokohama are much different from the five other stations due to its location. Yokohama is a large terminal station in the south of Tokyo, and it attracts a large volume of flows from the surrounding area. It can also be seen that many people whose origin is in the southern part of the Kanagawa Prefecture and destination is the center of Tokyo pass through Yokohama. Fig. 10 shows that the spatial flow pattern for Akihabara is different from those for Ikebukuro and Shibuya because it attracts flows from north and east. Fig. 10 shows that the volume of SCF for Akihabara is much smaller than the volume of SCF and WCF combined. Figs. 10 and 11 show that although Tokyo and Akihabara stations are relatively close together, their flow patterns differ somewhat because different railway lines pass through the two stations. Lastly, the flow pattern for Yokohama is distinct among the six stations because flows that contain Yokohama comprise two main types of flow, namely those with destinations near Yokohama and those with destinations in the center of Tokyo. Because several railway lines with different stopping stations connect the southern part of the network with the center of Tokyo, some stations between Yokohama and the center of Tokyo have large values, such as Kawasaki and Shinagawa.

5. Analysis of covered flows for each station

Using the network data and flow data described in the previous section, we analyze the features when only one facility is placed, considering each station separately as the location. In the following, to illustrate an example for R-FILM, we assume that strong coverage is five

Table 3

Flow volume for six large stations.

	Shinjuku	Ikebukuro	Shibuya	Yokohama	Akihabara	Tokyo
SCF	853,178	645,941	624,455	557,247	285,467	436,423
WCF	569,939	390,477	308,588	343,211	543,537	370,383
SCF and WCF	1,423,117	1,036,418	933,043	900,458	829,004	806,806

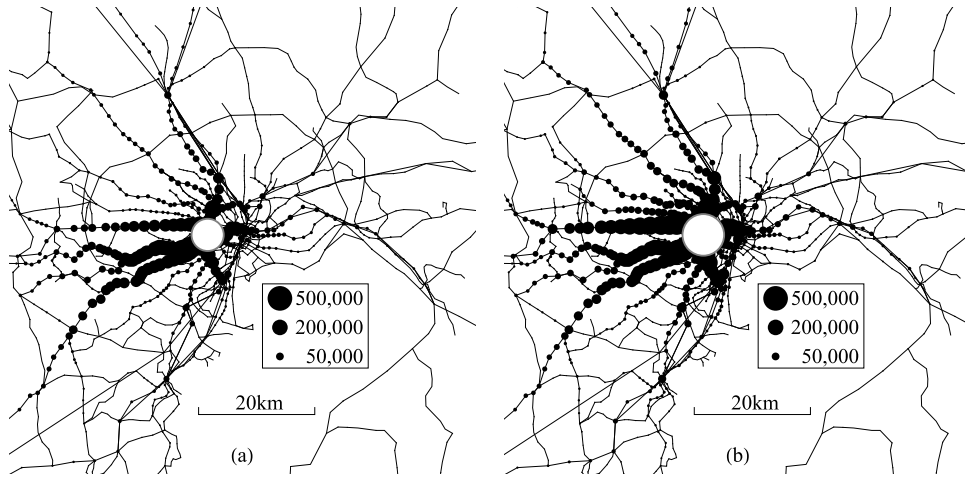


Fig. 6. Flows involving Shinjuku (left: only SCF; right: both SCF and WCF).

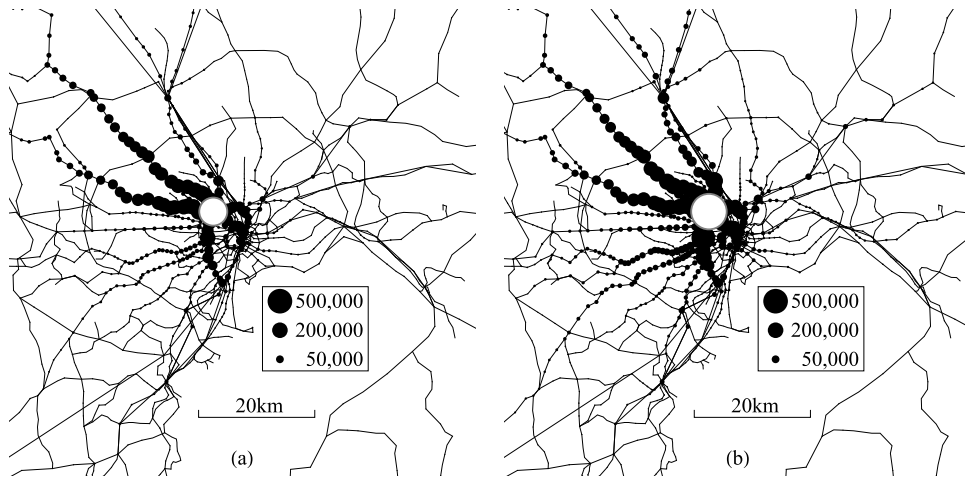


Fig. 7. Flows involving Ikebukuro (left: only SCF; right: both SCF and WCF).

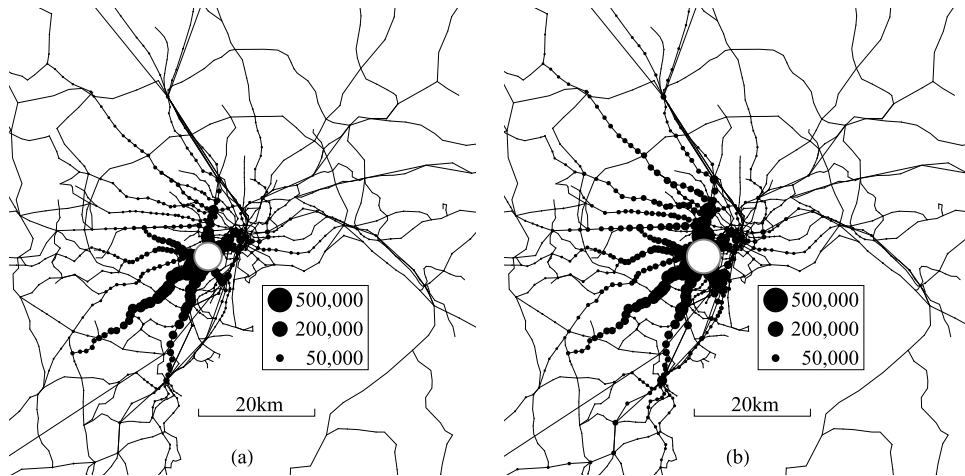


Fig. 8. Flows involving Shibuya (left: only SCF; right: both SCF and WCF).

as desirable as weak coverage (i.e., $\theta_s = 1.0$ and $\theta_w = 0.2$). Note that when applying R-FILM to a specific real-world situation, appropriate parameter values for θ_s and θ_w should be used depending on the situation and types of facility.

The areas of the disks in Fig. 12 reflect the objective value under FILM and R-FILM when a single facility is located separately at each station. In both cases, the largest objective value is attained when a

facility is placed at Shinjuku, where the value is 1,423,117 for FILM and 957,133 for R-FILM. In both Fig. 12 (a) and (b), the area of disks is determined relative to the Shinjuku value: 1,423,117 for (a) and 957,133 for (b), as each legend shows. For both FILM and R-FILM, stations in the central area tend to have higher objective values. Closer comparison of Fig. 12 (a) and (b) reveals that the R-FILM value is relatively large at major transport hubs such as Shinjuku, Ikebukuro, and

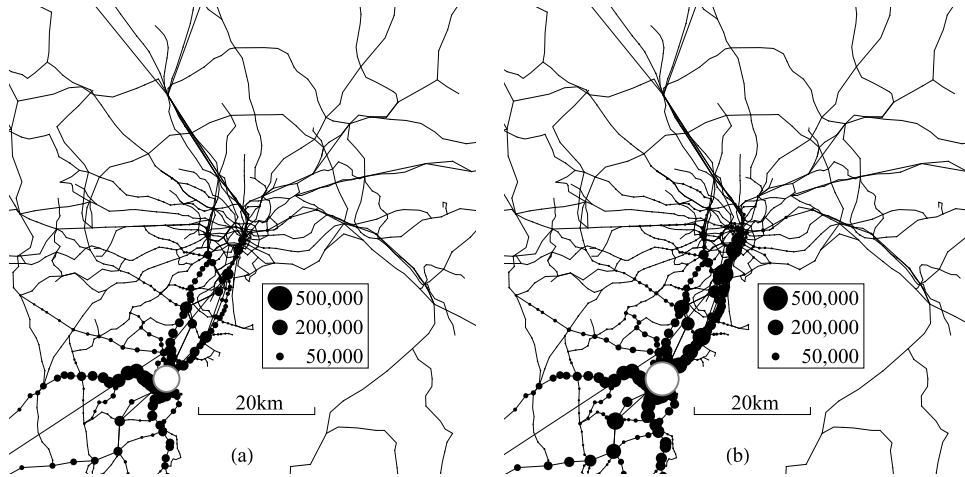


Fig. 9. Flows involving Yokohama (left: only SCF; right: both SCF and WCF).

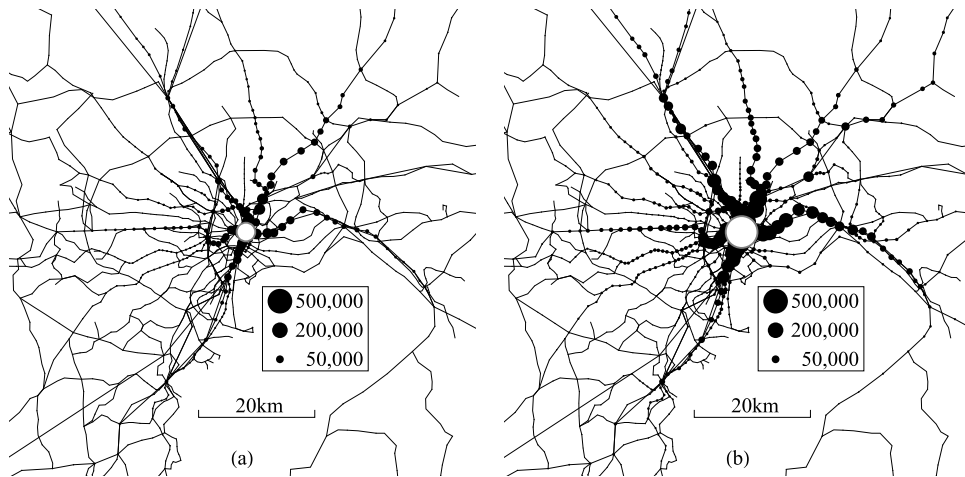


Fig. 10. Flows involving Akihabara (left: only SCF; right: both SCF and WCF).

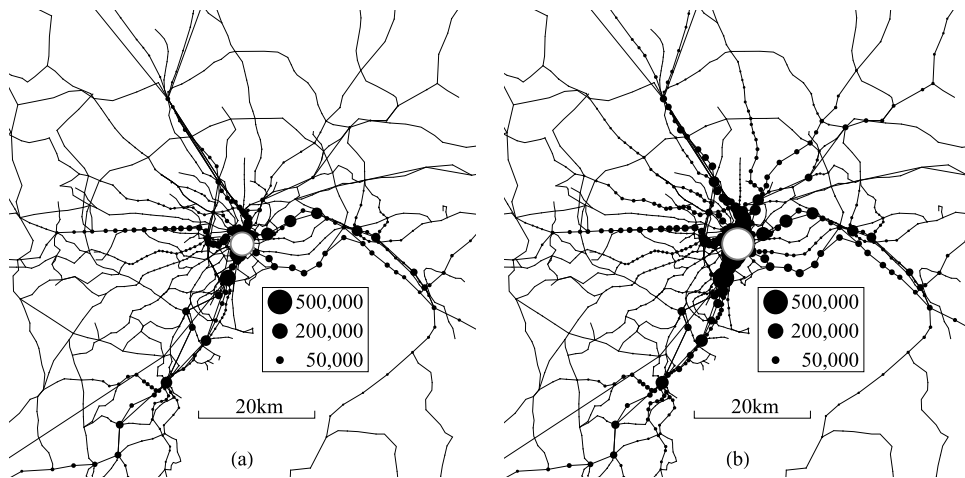


Fig. 11. Flows involving Tokyo (left: only SCF; right: both SCF and WCF).

Shibuya, where many people transfer to other lines. The numbers of people that use these stations as transfer stations are 593,726 for Shinjuku, 502,793 for Ikebukuro, and 474,071 for Shibuya, making them the top three stations among 1470 stations. Fig. 13 illustrates the 20 stations in terms of the objective values in Fig. 12 (a) and (b). Some important stations are highlighted by including the rank of the station in the disk for later reference. Table 4 compares the objective values of

the two models for the top 20 stations, showing table, rank, station name (StName), objective value (ObjVal), and objective value relative to Shinjuku (rVal), in that order. Note that the FILM objective values are larger than the R-FILM ones because FILM is a special case of R-FILM obtained by setting $\theta_s = \theta_w = 1.0$. On one hand, it can be seen that both models rank Shinjuku, Ikebukuro, Shibuya, and Yokohama as the top four stations, in that order. On the other hand, the ranking from

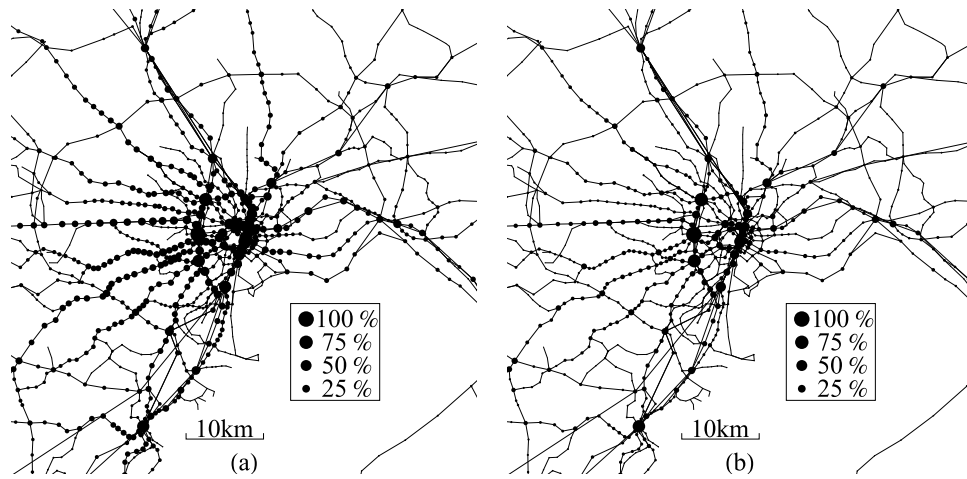


Fig. 12. Objective value for each station for one-facility case: (a) FILM, (b) R-FILM.

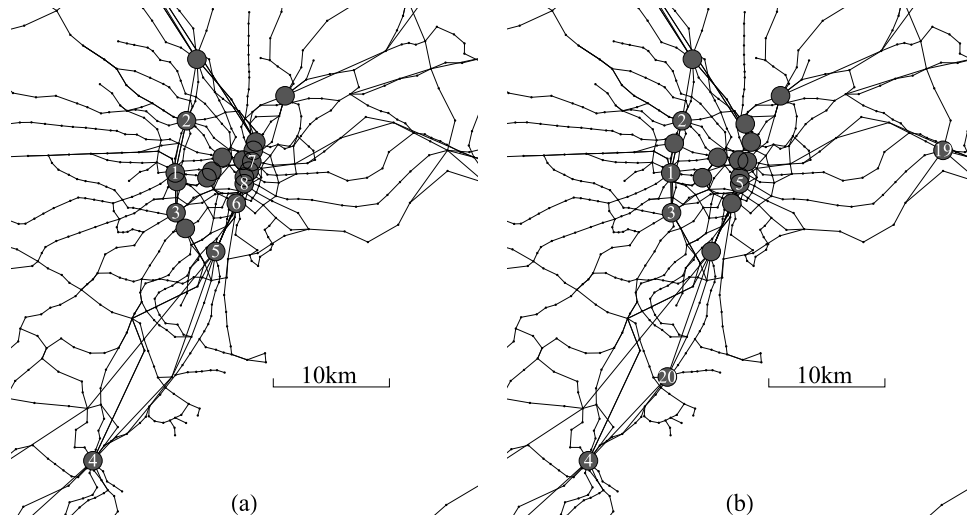


Fig. 13. Top 20 one-facility locations: (a) FILM, (b) R-FILM.

the fifth place is different between the two models. In FILM, all 20 stations are located in or near the center, whereas R-FILM includes some terminal stations located distantly from the city center, such as Omiya, Nishi-Funabashi, and Kawasaki as can be understood from Fig. 13 (b) and Table 4.

Also, from Table 4, it can be seen that the objective value relative to Shinjuku for R-FILM is smaller than for FILM when stations of the same rank are compared. This indicates the real-world superiority of some large terminal stations when the access pattern of train passengers is taken into consideration. The level of commercial accumulation near Shinjuku, Ikebukuro, and Shibuya is quite high, and there are many facilities near these stations, such as department stores, electronic and home appliance stores, large book stores, and various kinds of vocational schools. This reflects the importance of incorporating different levels of accessibility to stations based on the types of stations for each flow.

It is interesting to compare the ranking of R-FILM objective values shown in Table 4 with that of passengers for each station, which can be obtained by adding the values of Fig. 5 (a) and (b). Stations with relatively few passengers but high rank in Table 4 are considered to be important stations. Two such examples are Ueno and Kita-Senju. Ueno ranks eighth in the R-FILM objective value but 36th in terms of passengers. Kita-Senju ranks ninth in the R-FILM objective value but 133th in terms of passengers. These two stations have several railway lines and are used as transfer stations by many passengers, which is an

important advantage for covering flows by a facility located at such stations.

6. Analysis of optimal solutions for multi-facility case

This section explores optimal solutions for multi-facility cases. Using the flow data and railway network data, we used the mathematical programming solver Gurobi Optimizer 7.5.2 to obtain optimal solutions for FILM and R-FILM with $m = 1$ to $m = 20$. Our hardware was a PC with an Intel Core i7-7700 (4.20 GHz) processor and 32 GB of RAM. In the numerical experiments, we treated flow variables y_q (FILM), y_q^s , and y_q^w (R-FILM) as between 0 and 1, instead of as pure binary variables. In the optimal solutions obtained under this assumption, all flow variables were binary, and so they are also optimal solutions in the original formulation. The average computational time for FILM for 20 instances was 61 sec, while that for R-FILM was 338 sec. These values are quite reasonable considering that the problem instances are based on a large-scale network and flow data arising from real-world application.

Tables 5 and 6 list station names (S1 through S8) for the first eight optimal solutions for FILM and R-FILM, respectively, where $\theta_s = 1.0$ and $\theta_w = 0.2$ for R-FILM. In Figs. 14 and 15, selected stations for FILM and R-FILM with $m = 3$ and $m = 8$ are compared; the number in each diamond is the station number in Tables 5 and 6.

As we saw in the previous section, the best location for a single

Table 4
Top 20 stations for FILM and R-FILM objective value.

FILM				R-FILM			
rank	StName	ObjVal	rVal[%]	rank	StName	ObjVal	rVal[%]
1	Shinjuku	1,423,117	100.0	1	Shinjuku	967,166	100.0
2	Ikebukuro	1,036,418	72.8	2	Ikebukuro	724,036	74.9
3	Shibuya	933,043	65.6	3	Shibuya	686,173	70.9
4	Yokohama	900,458	63.3	4	Yokohama	625,889	64.7
5	Shinagawa	891,171	62.6	5	Tokyo	510,500	52.8
6	Shimbashi	841,196	59.1	6	Shinagawa	466,211	48.2
7	Akihabara	829,004	58.3	7	Akihabara	394,174	40.8
8	Tokyo	806,806	56.7	8	Ueno	350,914	36.3
9	Yoyogi	738,357	51.9	9	Kita-Senju	344,416	35.6
10	Iidabashi	688,099	48.4	10	Shimbashi	335,085	34.6
11	Ueno	673,241	47.3	11	Omiya	323,272	33.4
12	Ochanomizu	653,693	45.9	12	Ochanomizu	290,727	30.1
13	Okachimachi	626,579	44.0	13	Iidabashi	269,638	27.9
14	Kita-Senju	620,614	43.6	14	Takadanobaba	264,950	27.4
15	Kanda	598,097	42.0	15	Nippori	246,051	25.4
16	Ichigaya	580,684	40.8	16	Akabane	232,863	24.1
17	Akabane	553,545	38.9	17	Otemachi	228,010	23.6
18	Otemachi	537,279	37.8	18	Yotsuya	218,208	22.6
19	Yotsuya	516,470	36.3	19	Nishi-Funabashi	217,227	22.5
20	Ebisu	501,167	35.2	20	Kawasaki	214,469	22.2

Table 5
Optimal Solutions for FILM with $m = 1$ through $m = 8$.

m	S1	S2	S3	S4	S5	S6	S7	S8
1	Shinjuku							
2	Shinjuku	Yokohama						
3	Shinjuku	Yokohama	Akihabara					
4	Shinjuku	Yokohama	Akihabara	Ikebukuro				
5	Shinjuku	Yokohama	Akihabara	Ikebukuro	Shibuya			
6	Shinjuku	Yokohama	Akihabara	Ikebukuro	Shibuya	Otemachi		
7	Shinjuku	Yokohama	Akihabara	Ikebukuro	Shibuya	Otemachi	Shimbashi	
8	Shinjuku	Yokohama	Ueno	Ikebukuro	Shibuya	Otemachi	Shinagawa	Funabashi

Table 6
Optimal Solutions for R-FILM with $m = 1$ through $m = 8$.

m	S1	S2	S3	S4	S5	S6	S7	S8
1	Shinjuku							
2	Shinjuku	Yokohama						
3	Shinjuku	Yokohama	Ikebukuro					
4	Shinjuku	Yokohama	Ikebukuro	Shibuya				
5	Shinjuku	Yokohama	Ikebukuro	Shibuya	Tokyo			
6	Shinjuku	Yokohama	Ikebukuro	Shibuya	Tokyo	Kita-Senju		
7	Shinjuku	Yokohama	Ikebukuro	Shibuya	Tokyo	Kita-Senju	Akihabara	
8	Shinjuku	Yokohama	Ikebukuro	Shibuya	Tokyo	Kita-Senju	Akihabara	Omiya

facility is Shinjuku. In the two-facility case, the optimal solution for both models coincides: place facilities at Shinjuku and Yokohama. Yokohama ranks fourth in the objective value for both models after Ikebukuro (second) and Shibuya (third), as shown in Table 4. The reason that Yokohama is selected in the two-facility optimal solution can be understood as follows. The spatial flow pattern involving Yokohama is quite different from that of Shinjuku due to the large spatial separation of the two stations. In contrast, the flow patterns involving Ikebukuro and Shinjuku, or involving Shibuya and Shinjuku, are not much different in comparison with the difference between Yokohama and Shinjuku, as shown in Figs. 6–9. Thus, the percentage of flows covered by both facilities in (Shinjuku, Ikebukuro), or in (Shinjuku, Shibuya), is much larger than the percentage of flows covered by both of (Shinjuku, Yokohama), making the latter solution more efficient. In the three-facility case, optimal solutions were (Shinjuku, Yokohama, Akihabara) for FILM and (Shinjuku, Yokohama, Ikebukuro) for R-FILM,

where the location of the third facility differs. In the case of four or more facilities, the two models produce different optimal solutions. The tables also suggest that a greedy approach may be effective when constructing a solution method for the proposed model; this could be an important topic for future exploration.

In what follows, we pay attention to the optimal solutions for $m = 3$ and explore the reason that the two models produced different solutions. Since Yokohama, which is distant from the network center, has different flow pattern from other stations, we focus on (Shinjuku, Ikebukuro) for FILM and (Shinjuku, Akihabara) for R-FILM and analyze the covered flows for these two location choices. First, we examine flows whose travel path contains both (Shinjuku, Ikebukuro) or (Shinjuku, Akihabara).

Fig. 16 illustrates the aggregated flow volume at each station, with those containing (Shinjuku, Ikebukuro) shown in panel (a) and those containing (Shinjuku, Akihabara) shown in panel (b). The flow volume

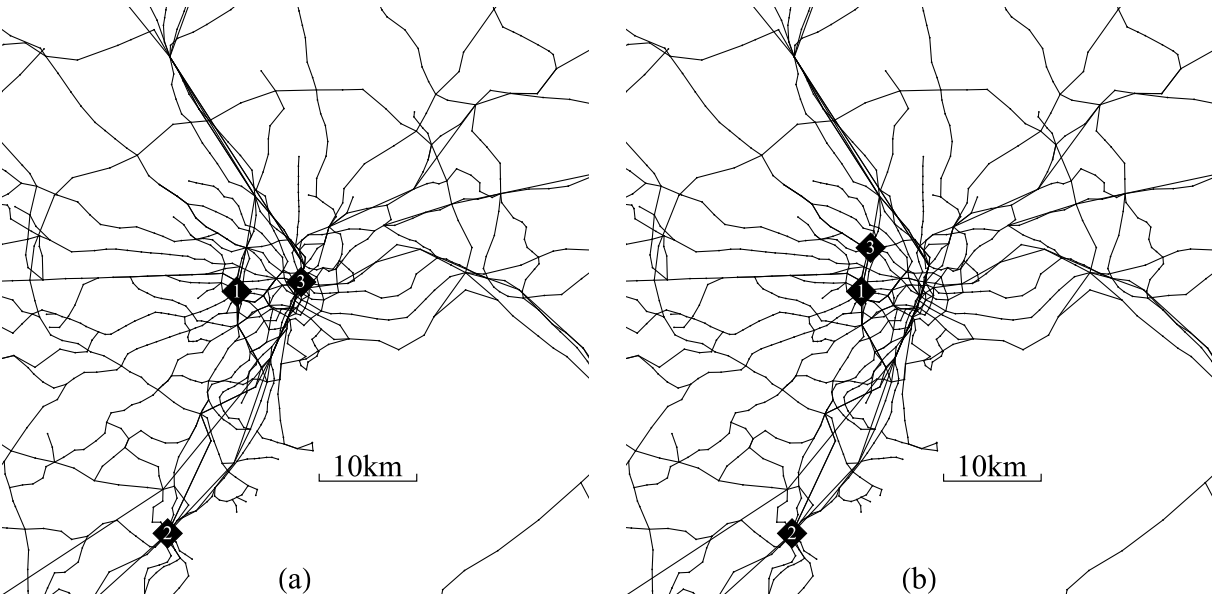


Fig. 14. Comparison of optimal solutions for FILM (a) and R-FILM (b) for $m = 3$.

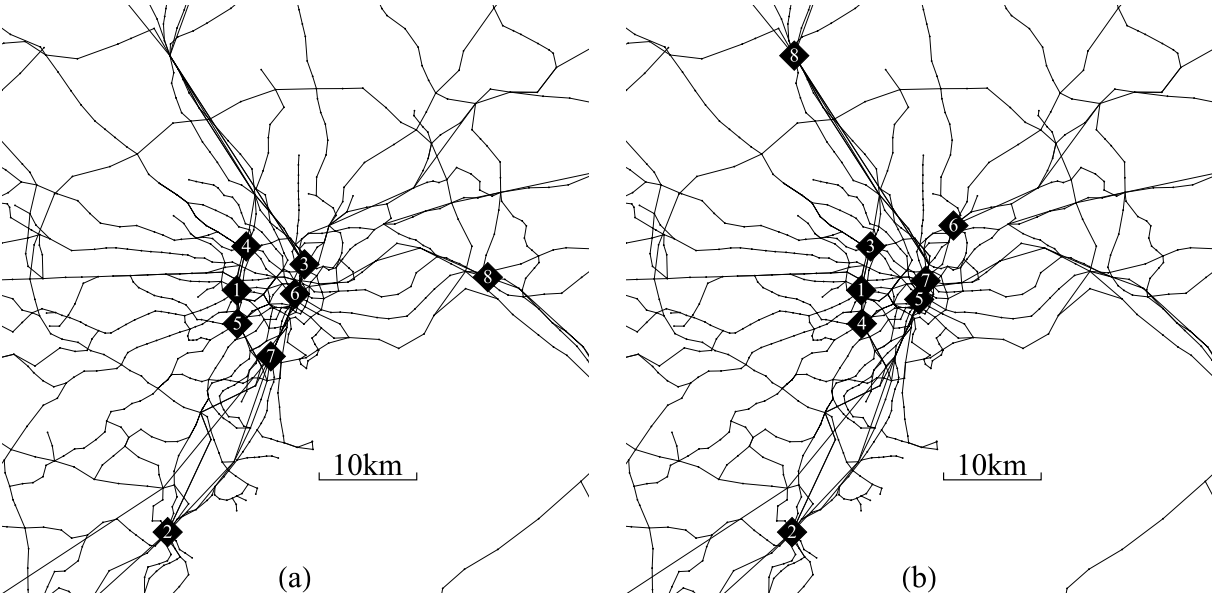


Fig. 15. Comparison of optimal solutions for FILM (a) and R-FILM (b) for $m = 8$.

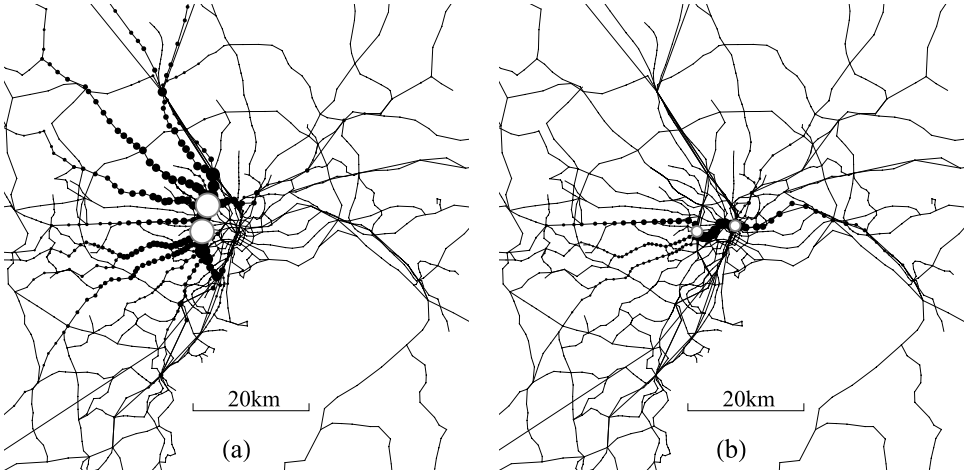


Fig. 16. Visualization of flows containing (Shinjuku, Ikebukuro) and (Shinjuku, Akihabara).

is represented by the areas of the disks; the scale is the same for both figures. Flows containing both Shinjuku and Ikebukuro are observed in the wider area (especially in the western area), but those containing both Shinjuku and Akihabara have lower volume and are concentrated mostly in the west–east direction. This result indicates that two facilities at (Shinjuku, Ikebukuro) double-cover many flows, whereas two facilities at (Shinjuku, Akihabara) are more efficient in terms of covering different types of flows. Because of this fact, the original FILM places facilities at (Shinjuku, Akihabara) for $m = 3$. The reason that (Shinjuku, Ikebukuro) is preferred to (Shinjuku, Akihabara) in R-FILM can be understood by referring to Table 3, Fig. 7, and Fig. 10. As mentioned in the previous section, while many travel paths contain Akihabara station, the number of passengers who use Akihabara as an ODT station is much smaller than the corresponding number for Ikebukuro.

The superiority of having facilities at (Shinjuku, Ikebukuro) over (Shinjuku, Akihabara), as suggested by R-FILM, matches the real-world situation. Shinjuku and Ikebukuro are major commercial centers in Tokyo, each having several huge department stores, large electronics stores, movie theaters, and many restaurants. The high degree of commercial accumulation around Shinjuku and Ikebukuro is confirmed by calculating the total floor area of retail within a circle of radius 500 m centered at each station using the Census of Commerce surveyed in 2014 by the Ministry of Economy, Trade and Industry of Japan [26]. The results show that while Shinjuku and Ikebukuro occupy first and second place, respectively, Akihabara is down in 25th. Also, Shinjuku and Ikebukuro are major transport hubs where many railway lines meet. Thus, there are many passengers who transfer to other lines at these stations, providing them easier access to facilities near the two stations. In summary, R-FILM reflects important railway-specific factors that are more accessible at ODT stations than at other stations. By differentiating coverage levels based on this feature, the proposed R-FILM can be used as a basic planning framework for deciding facility locations over railway networks.

7. Conclusions and future research directions

When providing facility services in cities with a developed railway network, the choice of station for placement of the facility is often important. In this paper, we focused on the situation where commuters visit a facility at railway stations along their travel path to obtain a service. We proposed a generalized flow interception location model that considers railway-specific features. Concretely, for railway users it is easier to access a facility located at an ODT station than to visit a facility at other stations on the travel path. Accessing a facility at a non-ODT station involves getting off the train to obtain the service and getting on another train to restart the trip after consuming the service, which is an obvious cost for train users. Focusing on this aspect, we developed R-FILM to incorporate two different coverages, based on the type of station providing coverage. R-FILM assumes that strong coverage is achieved when at least one facility is located at an ODT station, and weak coverage when only non-ODT stations have a facility on the travel path. We presented an integer programming formulation of R-FILM with two levels of coverage. To extend the model further, we also proposed R-FILM with multiple levels of coverage and considered several possible real-world applications of the three-coverage model.

The original FILM and the proposed R-FILM were both applied to the Tokyo metropolitan railway network (1,470 stations as candidate locations, 100,207 flow paths). The present case study is one of the largest examples ever considered within a FILM framework. By applying the proposed formulation and using a mathematical optimization solver, we obtained optimal solutions for FILM and R-FILM within a reasonable amount of time, locating up to 20 facilities. The results indicated that large terminal stations tend to be more heavily selected with R-FILM than with FILM.

There are a number of important future research directions using R-

FILM. First, application of the three-coverage model discussed in the paper is an interesting research topic. Models include deviation from the original travel path and differing treatments of origin, destination, and transfer stations. Depending on the situations and types of facilities considered, R-FILM with multiple coverage levels may be effectively used for railway application in various scenarios.

Considering locational analysis from a railway company's view point is another attractive research topic to explore. In Japan and elsewhere, railway companies often own and operate group companies in fields such as hotels, restaurants, and supermarkets. In such situations, they may be interested in locating facilities at stations belonging to their own railway lines. The locational competition between different railway companies is an important factor that should be reflected in the model.

Finally, when using the R-FILM framework for specific real-world decision-making situations, appropriate parameter values for θ_S and θ_W should be used. Suitable values for θ_S and θ_W depend on the situations and types of facilities of interest, and so large-scale questionnaire surveys may be required to estimate θ_S and θ_W .

Acknowledgements

The authors are grateful for the three anonymous reviewers for their careful reading of the original version of our manuscript and for providing valuable and insightful suggestions. This work was supported by JSPS KAKENHI Grant Number 18H01661 and 18H01662.

References

- [1] Averbakh I, Berman O. Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. *Eur J Oper Res* 1996;91(3):495–506.
- [2] Berman O, Bertsimas D, Larson RC. Locating discretionary service facilities, II: maximizing market size, minimizing inconvenience. *Oper Res* 1995;43(4):623–32.
- [3] Berman O, Hodgson MJ, Krass D. Flow-interception problems. In: Drezner Z, editor. *Facility location: a survey of applications and methods*. New York: Springer-Verlag; 1995. p. 389–426.
- [4] Berman O, Larson RC, Fouska N. Optimal location of discretionary service facilities. *Trans Sci* 1992;26(3):201–11.
- [5] Berman O, Drezner Z, Krass D. Generalized coverage: new developments in covering location models. *Comput Oper Res* 2010;37(10):1675–87.
- [6] Berman O, Krass D. The generalized maximal covering location problem. *Comput Oper Res* 2002;29(6):563–81.
- [7] Chung SH, Kwon C. Multi-period planning for electric car charging station locations: a case of Korean expressways. *Eur J Oper Res* 2015;242(2):677–87.
- [8] Church R, ReVelle C. The maximal covering location problem. *Papers Reg Sci Assoc* 1974;32(1):101–18.
- [9] Church R, Roberts KL. Generalized coverage models and public facility location. *Papers Reg Sci Assoc* 1983;53(1):117–35.
- [10] Drezner Z, Wesolowsky GO, Drezner T. The gradual covering problem. *Nav Res Logist* 2004;51(6):841–55.
- [11] García-Archilla B, Lozano AJ, Mesa JA, Perea F. GRASP algorithms for the robust railway network design problem. *J Heuristics* 2013;19(2):399–422.
- [12] Gutiérrez-Jarpa G, Laporte G, Marianov V. Corridor-based metro network design with travel flow capture. *Comput Oper Res* 2018;89:58–67.
- [13] Gutiérrez-Jarpa G, Obreque C, Laporte G, Marianov V. Rapid transit network design for optimal cost and origin-destination demand capture. *Comput Oper Res* 2013;40(12):3000–9.
- [14] Hakimi SL. Optimum locations of switching centers and the absolute centers and medians of a graph. *Oper Res* 1964;12(3):450–9.
- [15] Hodgson MJ. A flow-capturing location-allocation model. *Geogr Anal* 1990;22(3):270–9.
- [16] Hodgson MJ, Berman O. A billboard location model. *Geogr Environ Model* 1997;1(1):25–45.
- [17] Hodgson MJ, Rosing KE, Zhang J. Locating vehicle inspection stations to protect a transportation network. *Geogr Anal* 1996;28(4):299–314.
- [18] Hodgson MJ, Rosing KE, Storrier ALG. Applying the flow-capturing location-allocation model to an authentic network: Edmonton, Canada. *Eur J Oper Res* 1996;90(3):427–43.
- [19] Horner MW, Groves S. Network flow-based strategies for identifying rail park-and-ride facility locations. *Soc-Econ Plann Sci* 2007;41(3):255–68.
- [20] Kuby M, Lim S. The flow-refueling location problem for alternative-fuel vehicles. *Soc-Econ Plann Sci* 2005;39(2):125–45.
- [21] Kuby M, Lines L, Schultz R, Xie Z, Kim J-G, Lim S. Optimization of hydrogen stations in Florida using the flow-refueling location model. *Int J Hydrogen Energy* 2009;34(15):6045–64.
- [22] Körner M-C, Mesa JA, Perea F, Schöbel A, Scholz D. A maximum trip covering location problem with an alternative mode of transportation on tree networks and

- segments. *TOP* 2014;22(1):227–53.
- [23] Laporte G, Mesa JA. The design of rapid transit networks. In: Laporte G, Nickel S, Saldanha da Gama F, editors. *Location Science*, Springer, Cham; 2015. p. 581–94.
 - [24] Laporte G, Mesa JA, Ortega FA, Sevillano I. Maximizing trip coverage in the location of a single rapid transit alignment. *Ann Oper Res* 2005;136(1):49–63.
 - [25] Marković N, Ryzhov IO, Schonfeld P. Evasive flow capture: optimal location of weigh-in-motion system, tollbooths, and security checkpoints. *Networks* 2015;65(1):22–42.
 - [26] Ministry of Economy, Trade and Industry, The Census of Commerce in 2014. URL: <http://www.meti.go.jp/statistics/tyo/syogyo/mesh/download.html> (Last accessed: October 31, 2018).
 - [27] Tanaka K. Maximum flow-covering location and service start time problem and its application to Tokyo metropolitan railway network. *J Oper Res Soc Japan* 2011;54(4):237–58.
 - [28] Tanaka K, Furuta T. Flow capturing location problem for railway passengers: a case study of JR Yamanote line. *J City Plan Inst Japan* 2010;45(3):145–50. (In Japanese)
 - [29] Tanaka K, Furuta T. Locating flow capturing facilities on a railway network with two levels of coverage. *Proc Ninth Int Symp Oper Res Appl (ISORA' 10)* 2010:332–9.
 - [30] Tanaka K, Furuta T. Flow-capturing location problem with fixed costs—optimal location analysis on Keio railway network with flow coverage based on stopping patterns of trains. *Trans Oper Res Soc Japan* 2012;55:161–76. (In Japanese)
 - [31] Tanaka K, Furuta T. Locations and service start time of flow-covering facilities with multiple coverage levels. *J Oper Res Soc Japan* 2013;56(3):177–97.
 - [32] Tanaka K, Furuta T. A hierarchical flow capturing location problem with demand attraction based on facility size, and its Lagrangian relaxation solution method. *Geogr Anal* 2012;44(1):15–28.
 - [33] Tanaka K, Toriumi S. Extension of the maximum flow-covering location and service start-time problem to allow flexible consumption. *J Adv Mech Des, Syst, Manuf* 2016;10(3):13 pages. JAMDSM0044
 - [34] The Institute for Transport Policy Studies, Census data for commuter traffic in Tokyo area in 2010. URL: http://www.mlit.go.jp/sogoseisaku/transport/sosei_transport_tk_000034.html (Last accessed: October 31, 2018).
 - [35] Zeng W, Hodgson MJ, Castillo I. The pickup problem: Consumers' locational preferences in flow interception. *Geogr Anal* 2009;41(1):107–26.
 - [36] Zeng W, Castillo I, Hodgson MJ. A generalized model for locating facilities on a network with flow-based demand. *Netw Spat Econ* 2010;10(4):579–611.