Cluster-Based Hierarchical Model for Urban Transit Hub Location Planning

Formulation, Solution, and Case Study

Jie Yu, Yue Liu, Gang-Len Chang, Wanjing Ma, and Xiaoguang Yang

A cluster-based hierarchical location model for the selection of the proper locations and scales of urban transit hubs was developed with the objective of minimizing the demand-weighted total travel time. As an improvement to previous work, the proposed model has the following unique features: (*a*) it incorporates a hierarchical hub network topology that uses the concept of hub hierarchy establishment, route categorization, and service zone clustering to capture the critical operational issues for the transit network in an efficient manner and (*b*) it extends the previous nonhierarchical model to account for the impacts of hubs with various hierarchies as well as their interactions with lane use restrictions. An enhanced set of formulations along with the linearization approach was used to reduce significantly the number of variables and the computing time required to achieve the global optimum. The results of a case study in Suzhou Industrial Park in China revealed that the proposed model and solution method are quite promising for use in the planning of hub locations for the transit network. Sensitivity analysis of the performance of the system was also done to assist planners with the selection of the hierarchical structure and the design of transit routes.

Contending with traffic congestion has emerged as one of the more pressing issues during the process of urbanization. An increasing number of researchers have recognized that the development of transit-oriented urban transportation systems is one of the potentially effective strategies that can be used to relieve traffic congestion. In recent years, many big cities have been dedicated to the development of public transportation systems that are efficient from both the planning and the operation perspectives. Transit hubs are fundamental facilities in the urban transit system and are designed to provide switching points for intermodal flows and to feature seamless pedestrian connections. Properly located transit hubs significantly improve the effectiveness of limited transportation resources and the quality of transit services. Therefore, the transit hub location problem usually serves as the basis and the first step of the urban transit planning process.

Planning of the transit hub location is a branch of the hub location problem. Since O'Kelly first formulated a quadratic single-assignment model of the interaction of hub facilities from an operations research point of view (*1*), this subject has attracted the attention of researchers from a variety of fields, such as telecommunications, airline passenger management, and logistics. Most studies of hub location are concentrated on two basic types of models: the single-assignment model and the multiple-assignment model, depending on how nonhub nodes are connected to the hubs (*2, 3*).

In the single-assignment model, each node is connected to a single hub (*4*) and there is no sorting at the origin because all flow must travel to the same hub. However, the multiple-assignment model allows each node to be connected to more than one hub and sorting must take place at each origin that interacts with more than one hub (*5*). With the objective of minimizing the total travel cost, these two basic models require all services between the nonhub nodes to be connected to a hub, which is known to be strict hubbing policy.

To deal with more realistic characteristics of hub networks, researchers have explored different extensions, including the addition of a fixed cost to the objective function so that the trade-offs between travel costs and fixed costs are captured (*6*), the incorporation of a capacity constraint into the model by limiting the flows entering a hub under its capacity (*7*), or the use of a nonrestrictive hubbing policy that allows every pair of nodes to interact directly with each other (*8*). Sung et al. proposed a cluster-based hub location model with a nonrestrictive policy (*9*). In their model, exactly one hub location is assigned to a cluster to be opened, and traffic flows between nodes can be routed either directly or via hubs.

Despite the promising progress in hub location studies, a reliable optimal transit hub location model that is capable of capturing the critical operational issues in response to the needs for the development of transit-oriented urban transport systems is still lacking. In a recent study, Yu et al. developed an optimization model that can be used to choose urban transit hub locations on the basis of the cluster-based concept from the traditional hub location problem and proposed that a reformulation and linearization approach be implemented to solve the model (*10*). The impacts of several critical factors, such as the number of hubs and the travel time discount coefficient on the system measure of effectiveness (MOE), were also investigated. Applications of their model to a case study yielded promising results compared with the results achieved by use of a nonhub policy. However, their model did not account for hub hierarchies and their impacts on economies of scale.

Because of the hierarchical nature of transit service facilities and routes, it is essential to design hierarchical hubs to serve the transit system better. For example, rail transit routes could connect region-level hubs, whereas arterial routes may be suitable for the linkage of area-level hubs and branch or local routes may be used to connect local-level hubs and demand origins or destinations. In a review of the literature, hierarchical facility location problems have been formulated and widely applied in health care systems (*11*), solid waste management systems (*12*), production–distribution systems (*13*), and

J. Yu, Y. Liu, and G.-L. Chang, Department of Civil Engineering, University of Maryland, College Park, MD 20742. W. Ma and X. Yang, Department of Traffic Engineering, Tongji University, Shanghai 200092, China. Corresponding author: J. Yu, yujie@umd.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2112, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 8–16.
DOI: 10.3141/2112-02

education systems (*14*), among others. Most of those problems are formulated as hierarchical *p*-median models, set-covering models, or fixed-charge models (*15*). However, only a few studies that have dealt with the hierarchical location problem for a transit network under the cluster-based framework have been conducted. Hence, this research focuses on completing the following critical tasks:

- Design a hierarchical hub network topology that integrates the concept of hub hierarchy establishment, route categorization, and service zone clustering to efficiently capture the critical operational issues for the transit network;
- Formulate a hierarchical hub location model that is capable of capturing the impacts of hubs with various hierarchies and producing detailed output information, including optimal locations, hierarchies, and scales of hubs to assist the responsible agencies with prioritizing limited budgets for hub construction;
- Apply the enhanced formulation and linearization approach to solve the proposed model and yield the tractable solution for large-scale real-world applications; and
- Test the proposed model with an example and perform sensitivity analysis of the critical factors that may affect the performance of the model, such as the proportions of different hierarchies and travel time discount coefficients between hubs.

This paper is organized as follows. The next section details the multihierarchy hub network topology on which the hub location optimization model is based. The model formulation, including the objective function and operational constraints, is then presented. A set of enhanced reformulations along with linearization approach is illustrated. The performance of the proposed model is evaluated, and its sensitivity is analyzed by use of a case study in Suzhou Industrial Park in China. Concluding comments along with future extensions of this study are reported in the last section.

MULTIHIERARCHY HUB NETWORK TOPOLOGY DESIGN

This section proposes a multihierarchy hub network topology that can be summarized as hub hierarchy structuring, route categorizing, and service zone clustering. As shown in Figure 1, a transit network is clustered into several service zones centered with hubs with various hierarchies to consolidate flows from demand origins to destinations. Transit routes with different capacities and service levels serve as the connections between those hubs to ensure the economies of scale obtained.

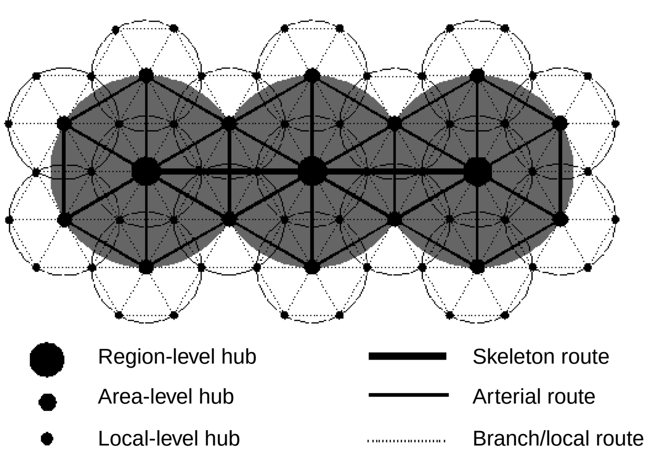

FIGURE 1 Multihierarchy hub network topology.

Hub Hierarchy Structuring

In this study, according to the differences in serving capacities, flow characteristics, and importance, transit hubs are classified into three hierarchies: region level, area level, and local level, as shown in Figure 1. The region-level transit hub serves as the key infrastructure in the entire network to facilitate large-scale flow exchanges. As an intermediate facility, the area-level hub functions to connect flows between the local-level hub and the region-level hub. The local-level hub collects flows from local demand origins and destinations and dispatches them to the upper-level hubs.

Route Categorization

Corresponding to the different hierarchies of hubs, the transit routes between them can also be categorized into three classes: skeleton routes, arterial routes, and branch and local routes, as shown in Figure 1 and as described below:

Hub Level	*Region-Level Hub*	*Area-Level Hub*	*Local-Level Hub*
Region	Skeleton routes	Arterial routes	Branch and local routes
Area		Arterial routes	Branch and local routes
Local			Branch and local routes

Service Zone Clustering

The traditional cluster-based hub network structure could result in an outcome in which the region-level or area-level hub locations selected are concentrated only in highly populated areas, which may adversely reduce the economies of scale because of a relatively short distance between those hubs and which may also conflict with land use restrictions or other political as well as administrative regulations. To contend with this issue, this study defines various level-of-service zones to correspond to various hierarchies of hubs. A service zone consists of either one cluster or a bunch of clusters. Each service zone has a well-defined coverage range within which one hub with a prespecified hierarchy must be located, as shown in Figure 2. Note that hubs with higher hierarchies usually function to serve zones larger than those served by hubs with lower hierarchies.

HIERARCHICAL LOCATION MODEL

Model Assumptions

To yield a tractable solution for the proposed formulations with realistic constraints, this study uses the following assumptions.

Assumption 1. Discrete Location Nature

The target study area for this study can be divided into different traffic analysis zones (TAZs). The origins and destinations of demand, as well as the hubs, are assumed to be at the centroids of those TAZs, denoted as nodes.

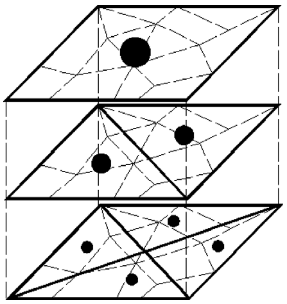

FIGURE 2 Service zones for hubs with different hierarchies.

Assumption 2. Cluster-Based Single Hub Allocation Policy

All nodes in the target network are partitioned into clusters in advance on the basis of the geographic relations between neighboring TAZs, land use restrictions, or other constraints. Only one hub with a designated hierarchy is located in any cluster. Moreover, all the hubs are assumed to be fully interconnected.

Assumption 3. Service Zone-Based Hub Allocation Policy

Clusters are further grouped into different levels of transit service zones with preset rules. A hub with a designated hierarchy must be located in each service zone.

Assumption 4. Nonrestrictive Policy

A nonrestrictive policy means that flows between origins and destinations may be sent either directly or through a hub(s), and the number of hub stops is no more than two. Under the nonrestrictive policy, if a node is assigned to a hub, any flow to or from that node must go via the hub or does not involve hubs at all (nonstop service). Therefore, the paths from origin node i to destination node j could have three possible choices, as shown in Figure 3: (*a*) nonstop, in which transit flows are transported directly from i to j; (*b*) one hub stop, in which transit flows are transported from i to j, with hub k being the transfer point; or (*c*) two hub stops, in which transit flows are transported from node i to node j via both hub k and hub m along the route $i \rightarrow k \rightarrow m \rightarrow j$.

Assumption 5

Hubs in the network are classified into three hierarchies, and travel time discount coefficients for skeleton and arterial transit routes exist.

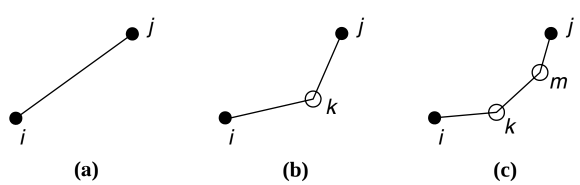

FIGURE 3 Paths for transit flows between origins and destinations: (*a*) nonstop, (*b*) one hub stop, and (*c*) two hub stops.

Assumption 6

The number of various hierarchies of the hubs is predetermined.

Assumption 7

The transfer time (including the walking time and the waiting time) at the hubs is assumed to be constant.

Notation

To facilitate the presentation, all definitions and notations used hereafter are summarized below.

Parameters and Sets

- G = target transit network;
- N = set of nodes in the target network (origin or destination);
- H = set of possible hierarchies for the target hub network;
- $i, j \in N$ = index of each node;
- k, m = index of a node that is potentially located with a hub;
- $u, v \in H$ = index of hierarchy for a transit hub;
- (i, j) = link (route) between nodes i and j;
- p = the number of clusters (hubs) in the target network;
- p^u = number of hubs with the hierarchy of u in the target network;
- $S^u \subseteq G$ = service zone within which one hub with hierarchy u must be located to provide satisfactory service, $i, j \in S^u \supseteq C_r$;
- C_r = cluster r in the target network ($r = 1, \ldots, p$);
- C_i = cluster to which node i is assigned ($i = 1, \ldots, N$);
- w_{ij} = flows from node i to node j (in number of trips);
- α_{km}^{uv} = travel time discount coefficient between nodes k and m if they are located with hubs that belong to hierarchy u and v, respectively;
- t_{ij} = nonstop average travel time from node i to j (in minutes);
- t_k = transfer time at hub k (in minutes); and
- t_{ijkm} = hub stop average travel time from node i to node j via hub nodes k and m (in minutes), where

$$t_{ijkm} = t_{ik} + t_k + \sum_{u \in H} \sum_{v \in H} \delta_k^u \delta_m^v \alpha_{km}^{uv} t_{km} + t_m + t_{mj} \qquad (t_{km} = 0 \text{ if } m = k)$$

Model Variables

Four sets of binary decision variables, x_{ij}, x_{ijkm}, y_k, and δ_k^u are defined:

$$x_{ij} = \begin{cases} 1 & \text{if flows between nodes } i \text{ and } j \text{ are transported with no stop} \\ 0 & \text{otherwise} \end{cases}$$

$$x_{ijkm} = \begin{cases} 1 & \text{if flows between nodes } i \text{ and } j \text{ are transported with one or two hub stops } (k = m, \text{ one hub stop}; k \neq m, \text{ two hub stops}) \\ 0 & \text{otherwise} \end{cases}$$

$$y_k = \begin{cases} 1 & \text{if a hub is located at node } k \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_k^u = \begin{cases} 1 & \text{if the hierarchy of a hub located at node } k \text{ is } u \\ 0 & \text{otherwise} \end{cases}$$

Formulation

The hierarchical location model for transit hubs is formulated as follows:

$$\min \sum_{i \in N} \sum_{j \in N} \left(t_{ij} x_{ij} + \sum_{k \in C_i} \sum_{m \in C_j} t_{ijkm} x_{ijkm} \right) w_{ij} \tag{1}$$

subject to

$$\sum_{k \in C_r} y_k = 1, r = 1, \ldots, p \tag{2}$$

$$x_{ij} + \sum_{k \in C_i} \sum_{m \in C_j} x_{ijkm} = 1 \qquad \forall i, j \tag{3}$$

$$\sum_{m \in C_j} x_{ijkm} \le y_k \qquad \forall j, k; i \in C_k \tag{4}$$

$$\sum_{m \in C_i} x_{ijkm} \le y_m \qquad \forall i, m; j \in C_m \tag{5}$$

$$t_{ijkm} = t_{ik} + t_k + \sum_{u \in H} \sum_{v \in H} \delta_k^u \delta_m^v \alpha_{km}^{uv} t_{km} + t_m + t_{mj} \quad \forall i, j, k \in C_i, m \in C_j \tag{6}$$

$$\sum_{k \in N} \delta_k^u = p^u \qquad \forall u \tag{7}$$

$$\sum_{u \in H} p^u = p \tag{8}$$

$$\sum_{u \in H} \delta_k^u = y_k \qquad \forall k \tag{9}$$

$$\sum_{k \in S^u} \delta_k^u = 1 \qquad \forall u \tag{10}$$

$$x_{ij}, x_{ijkm}, y_k \in \{0, 1\} \qquad \forall i, j, k, m \tag{11}$$

The objective function (Equation 1) aims at minimizing the demand-weighted total travel time in the system. Equation 2 establishes the limit that only one of the nodes in each cluster should be selected as a hub. Equation 3 means that flows between nodes i and j are transported via either the nonstop service or the hub stop service. Equations 4 and 5 prevent any hub services unless the node is selected as a hub. Equation 6 models the impact of hub hierarchies on the travel times via hubs. Equation 7 sets the total number of hubs allowed to be built with the hierarchy of u in the target network, and Equation 8 establishes the limit that the sum of hubs with various hierarchies must be equal to the total number of hubs in the target network. Equation 9 establishes the limit that there is only one hierarchy for a hub if it is located at node k. Equation 10 defines the serving zone for a hub with the hierarchy of u, which means that one hub with hierarchy u must be located to serve an area denoted by S^u. Equation 11 is a standard integrality constraint.

SOLUTION APPROACH

Note that the proposed model formulation is an integer nonlinear problem because of the nonlinearity of Equation 6 and may result in a huge number of variables because of the large dimensionality of x_{ijkm}. Therefore, it will be quite difficult to obtain the global optimal solution for the proposed model, especially when it is applied to a large hub network. To address this issue, this study has taken the following principal steps: (*a*) model reformulation to reduce the dimensionality of the original model with a set of new variables and bilinear constraints and (*b*) model linearization to ensure the existence of global optimal solutions.

Step 1. Model Reformulation

In the target hub network designed, each node belongs to a cluster and all candidate locations for a hub are also related to a cluster. Therefore, if one can calculate the total system travel time in a cluster-based way rather than a node-based way, the dimensionality of the proposed model will be significantly reduced. It is noticeable that the total travel time between two clusters (Case 1) or within a cluster (Case 2) would easily be determined if the hub and its hierarchy were determined for those clusters, as illustrated in Figure 4.

Case 1

If nodes k and m ($k < m$ and $C_k \neq C_m$) are selected as hubs with hierarchies of u and v (two hub stops), the demand-weighted total travel time to transport flows from i to j (T_{ij}^{uv}) is

$$\begin{aligned} T_{ij}^{uv} &= \min\left\{ w_{ij} t_{ij}, w_{ij} \left(t_{ik} + t_k + \alpha_{km}^{uv} t_{km} + t_m + t_{mj} \right) \right\} \\ &\quad + \min\left\{ w_{ji} t_{ji}, w_{ji} \left(t_{jm} + t_m + \alpha_{mk}^{vu} t_{mk} + t_k + t_{ki} \right) \right\} \end{aligned} \tag{12}$$

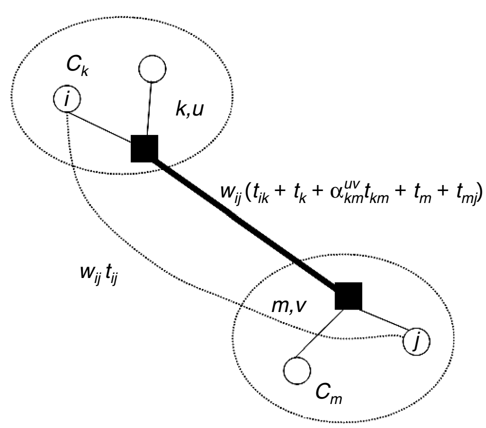

FIGURE 4 Minimum demand-weighted total travel time between clusters.

The total demand-weighted travel time between clusters C_k and C_m (T_{km}^{uv}) will then be

$$T_{km}^{uv} = \sum_{i \in C_k} \sum_{\substack{j \in C_m \cap k<m, \\ C_k \neq C_m}} T_{ij}^{uv} \tag{13}$$

Case 2

In Case 2, k is equal to m, and so the total demand-weighted travel time within clusters C_k (T^k) will be

$$T^k = \sum_{i \in C_k} \sum_{j \in C_k} T_{ij} \tag{14}$$

$$T_{ij} = \min\left\{w_{ij}t_{ij}, w_{ij}\left(t_{ik}+t_k+t_{kj}\right)\right\} + \min\left\{w_{ji}t_{ji}, w_{ji}\left(t_{jk}+t_k+t_{ki}\right)\right\} \tag{15}$$

To accommodate the cluster-based travel cost calculation approach described above, auxiliary binary variables δ_{km} and δ_{km}^{uv} ($k < m$ and $C_k \neq C_m$) are introduced to replace the variables x_{ij} and x_{ijkm}, where

$$\delta_{km} = \begin{cases} 1 & \text{if and only if hubs are set in both nodes } k \text{ and } m \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{km}^{uv} = \begin{cases} 1 & \text{if hubs set at nodes } k \text{ and } m \text{ have hierarchies} \\ & \text{of } u \text{ and } v\text{, respectively} \\ 0 & \text{otherwise} \end{cases}$$

The original model can be transformed into the following formulation:

$$\min \sum_{k<m \wedge C_k \neq C_m} \sum_{u \in H} \sum_{v \in H} T_{km}^{uv} \delta_{km}^{uv} + \sum_{k \in N} T^k y_k \tag{16}$$

subject to

$$\sum_{k \in C_r} y_k = 1, r = 1, \ldots, p \tag{17}$$

$$\delta_{km} = y_k \cdot y_m \qquad k < m, C_k \neq C_m \tag{18}$$

$$\sum_{u \in H} \sum_{v \in H} \delta_{km}^{uv} = \delta_{km} \qquad k < m, C_k \neq C_m \tag{19}$$

$$\sum_{k \in N} \delta_k^u = p^u \qquad \forall u \tag{20}$$

$$\sum_{u \in H} p^u = p \tag{21}$$

$$\sum_{u \in H} \delta_k^u = y_k \qquad \forall k \tag{22}$$

$$\sum_{k \in S^u} \delta_k^u \geq 1 \qquad \forall u \tag{23}$$

$$\delta_{km}^{uv}, \delta_{km}, y_k, \delta_k^u \in \{0, 1\} \tag{24}$$

The objective function (Equation 16) is the demand-weighted total travel time between different clusters and within any cluster and is equivalent to the original one (Equation 1). Equation 17 is the same as the constraint established by Equation 2. The constraints established by Equation 18 ensure that δ_{km} is 1 if and only if y_k and y_m are both equal to 1. Equation 19 establishes the limit that hubs set at nodes k and m can have only one set of hierarchies. Equations 20 to 23 retain the meanings of the original formulation.

Step 2. Model Linearization

After the reformulation, the high-dimension variables x_{ijkm} have been replaced with lower-dimensionality auxiliary variables δ_{km} and δ_{km}^{uv} and the nonlinear constraints of Equation 6 have been eliminated. These changes make the proposed model tractable for large-scale applications. However, a new set of bilinear constraints (Equation 18) was added to the reformulation. To have the global optimal solutions for the reformulated model, the newly added bilinear constraints (Equation 18) were linearized with the following functions:

$$\delta_{km} \geq 0 \tag{25}$$

$$\delta_{km} \leq y_m \tag{26}$$

$$\delta_{km} \geq y_k + y_m - 1 \tag{27}$$

$$\delta_{km} \leq y_k \tag{28}$$

Equations 25 to 28 can always ensure that Equation 18 holds. Thus, by replacing Equation 18 with Equations 25 to 28, the model becomes a mixed-integer program (MIP) that can be solved with the existing MIP solvers within a reasonable time frame.

CASE STUDY

To illustrate the applicability of the proposed hierarchical location model and solution approach, this study uses the transit network in Suzhou Industrial Park in China as a case study.

Network Layout

As shown in Figure 5, the study network is divided into 58 TAZs, which can be further classified into 25 clusters on the basis of geographic and administrative restrictions. Each of the clusters contains exactly one hub with a specified hierarchy. According to the network topology designed as described earlier in the paper, this case study is given five and two service zones for area-level hubs and region-level hubs, respectively. The node identifiers (IDs) within those service zones are summarized in Table 1. Therefore, the numbers of region-level, area-level, and local-level hubs are two, five, and 18, respectively.

Optimization Model Settings

To implement the proposed model, the following information should be available as inputs:

- The origin–destination matrix and the average travel time matrix of the study network;
- The clustering rule for the study network;
- The travel time discount coefficients for skeleton routes, arterial routes, and branch and local routes, denoted by α_1, α_2, and α_3, respectively (here, α_1:α_2:α_3 = 0.3:0.5:0.7 is used);

FIGURE 5 Study network layout: (*a*) 58 TAZs in the study network, (*b*) 25 clusters in study network, (*c*) five service zones for five area-level hubs, and (*d*) two service zones for two region-level hubs.

- The total number of hubs and the hierarchical structure (here, there are two region-level hubs, five area-level hubs, and 18 local-level hubs);
- The service zone corresponding to each hierarchy of hub, as shown in Figure 5; and
- A constant transfer time at hubs (here, 3 min is used).

On the basis of the input information, responsible agencies can then use the model proposed here to obtain the optimal hubbing policy, which includes the following four types of information:

- System MOE, that is, the total demand-weighted travel time for the study network;

TABLE 1 Service Zone Clustering Rules in Study Network

Corresponding Hub Hierarchy	Service Zone Index	TAZ IDs
Region level	1	{1–9, 12, 37–39, 47–51}
	2	{10–11, 13–36, 40–46, 52–58}
Area level	1	{1–9,12, 37–39}
	2	{10–11, 13–19, 20–28}
	3	{40–46}
	4	{32–36}
	5	{29–31, 47–58}

- The list of open hubs, their locations, and their hierarchies;
- The scale of the hubs (flows in and out); and
- The route assignment of transit flows between TAZs.

Optimization Results

The proposed model was implemented in the LINGO MIP Solver program, and the optimal locations (TAZ IDs) of the open hubs are shown in Figure 6 and Table 2. The total demand-weighted travel time under the hierarchical hubbing policy is 68,789 h. It has been reduced by about 18.8% compared with the system performance of 84,731 h without the hubbing policy (*10*).

Furthermore, on the basis of the optimal locations of hubs, the route assignment of each transit origin–destination pair can easily be determined with the following equations:

$$\underset{i\to j,i,j\in C_k}{\text{path}} = \begin{cases} i-j & \text{if } t_{ij} < t_{ik} + t_k + t_{kj} \\ i-k-j & \text{otherwise} \end{cases} \tag{29}$$

$$\underset{i\to j,i\in C_k,j\in C_m \cap k<m, C_k\neq C_m}{\text{path}} = \begin{cases} i-j & \text{if } t_{ij} < t_{ik} + t_k + \alpha_{km}^{uv} t_{km} \\ & \quad + t_m + t_{mj} \\ i-k-m-j & \text{otherwise} \end{cases} \tag{30}$$

FIGURE 6 Optimal locations of the transit hubs from the model.

Equation 29 determines the route assignment within cluster *k,* and Equation 30 determines the flows between clusters *k* and *m.* The route assignment information then serves as the basis for the generation of the scales of the open hubs. If a hub is located at node *k,* then the hub scale (Q_k) can be computed with the following equations:

$$Q_k = Q_k' + Q_k'' \tag{31}$$

$$Q_k' = \sum_{i,j \in C_k} \left(w_{ij} + w_{ji} \right) \qquad \text{for } t_{ij} > t_{ik} + t_k + t_{mj};\ t_{ji} > t_{jk} + t_k + t_{ki} \tag{32}$$

TABLE 2 Optimal Results Generated from the Model

Hub Hierarchy	Hub Location (TAZ IDs)	Hub Scale (in trips/h)
Region level	6	15,574
	19	24,850
Area level	3	10,862
	24	7,979
	35	3,303
	42	3,916
	52	1,315
Local level	12	4,087
	15	4,667
	17	4,077
	26	2,580
	27	5,311
	28	1,774
	31	872
	33	1,957
	36	1,317
	38	2,904
	39	855
	40	2,635
	45	848
	46	921
	47	992
	49	872
	56	527
	58	1,216

$$Q_k'' = \sum_{\substack{i \in C_k, j \in C_m \cap k < m, \\ C_k \neq C_m}} \left(w_{ij} + w_{ji} \right) \qquad \text{for } t_{ij} > t_{ik} + t_k + \alpha_{km}^{uv} t_{km} + t_m + t_{mj}; \quad t_{ji} > t_{jm} + t_m + \alpha_{mk}^{vu} t_{mk} + t_k + t_{ki} \tag{33}$$

where Q_k' and Q_k'' represent the flow exchanges incurred at hub *k* within the cluster and between clusters, respectively. Q_k is then equal to the sum of Q_k' and Q_k''. The scales of the hubs in this case study were calculated with Equations 31 to 33 and are summarized in Table 2. This information will help planners determine the resources allocated to each hub to ensure that each hub operates under its flow capacity.

Comparison with Nonhierarchical Hub Location Model

Compared with the nonhierarchical model, the strength of the proposed hierarchical model lies in its ability to further achieve the economies of scale through different hierarchies of hubs and transit routes and to reflect lane use characteristics by defining the transit zones. However, it does not always outperform the nonhierarchical model under various hierarchical structures and with various travel time discount coefficients. With the aim of providing guidelines on choosing the proper hub network structure to transit planning agencies, this study has compared the performance of the proposed hierarchical model with that of the nonhierarchical model under various scenarios.

The sensitivity analysis was performed with the same study network of 25 hubs that were to be located. For the hierarchical model, the travel time discount coefficients for skeleton routes, arterial routes, and branch and local routes were fixed at 0.3, 0.5, and 0.7, respectively; and a total combinations of 35 hierarchical structures (the number of region-level hubs varied from one to five and the number of area-level hubs ranged from four to 10; thus, the number of local-level hubs changes from 10 to 20) were evaluated. For the nonhierarchical model, 25 hubs without hierarchies were selected, with the travel time discount coefficient for the routes between hubs varying from 0.5 to 0.9. The objective function values for the hierarchical model under various hierarchical structures and the objective function values for the nonhierarchical model under various travel time discount coefficients are summarized in Table 3 and Table 4. Comparison of the results obtained with the two models are illustrated in Figure 7.

Figure 7 uses *R* to represent the ratio of the hierarchical model to the nonhierarchical model in terms of the objective function values. Cells in light, medium, and dark gray indicate that the hierarchical model produces worse, slightly better, and better performances than the nonhierarchical model, respectively, under the given scenario.

As indicated in Table 3, Table 4, and Figure 7, the following findings can be made:

- Under the given clustering rules and with fixed travel time discount coefficients for skeleton routes, arterial routes, and branch and local routes, the objective function decreases when the numbers of region-level or area-level hubs increase in the network, which is quite understandable, because the economies of scale will increase correspondingly. This information will help planners find the proper hierarchical structure that both satisfies the budget constraints and maximizes the economies of scale.
- With any given hierarchical structure for the hierarchical model, there exists a threshold of the travel time discount coefficient (α) for the nonhierarchical model below which the hierarchical structure has no advantage over the nonhierarchical structure. For

TABLE 3 Results of Sensitivity Analysis as Objective Function Values of the Hierarchical Model Under Various Hierarchy Structures (in hours)

Number of Area-Level Hubs	Number of Region-Level Hubs				
	1	2	3	4	5
4	70,551.6	69,149.93	68,458.7	67,613.86	66,681.96
5	69,874.77	68,789.92	67,973.88	67,166.58	66,322.59
6	69,543.55	68,497.1	67,604.26	66,752.75	65,881.54
7	69,188.34	68,026.68	67,204.69	66,426.04	65,445.93
8	68,799.52	67,647.46	66,785.42	65,919.65	65,037.55
9	68,410.7	67,248.25	66,469.61	65,484.05	64,607.39
10	68,036.28	66,812.65	65,957.77	65,173.68	64,209.9

TABLE 4 Results of Sensitivity Analysis as Objective Function Values of the Nonhierarchical Model Under Various Travel Time Discount Coefficients (in hours)

Travel Time Discount Coefficient α	Objective Function Values
0.5	61,971.97
0.6	67,417.05
0.7	72,217.28
0.8	76,197.93
0.9	78,848.17

example, the hierarchical structure with three region-level hubs, five area-level hubs, and 17 local-level hubs (highlighted as bold rectangles in Figure 7) will not produce better system performance unless the travel time discount coefficient in the nonhierarchical structure is greater than 0.6 (the color of the highlighted cell turns from light gray to medium gray in Figure 7). This information will provide guidelines to help planners design speeds for transit routes.

CONCLUSIONS

This study has presented a cluster-based hierarchical transit hub location optimization model that is based on a hierarchical hub network topology and that employs the concepts of hub hierarchy structuring,

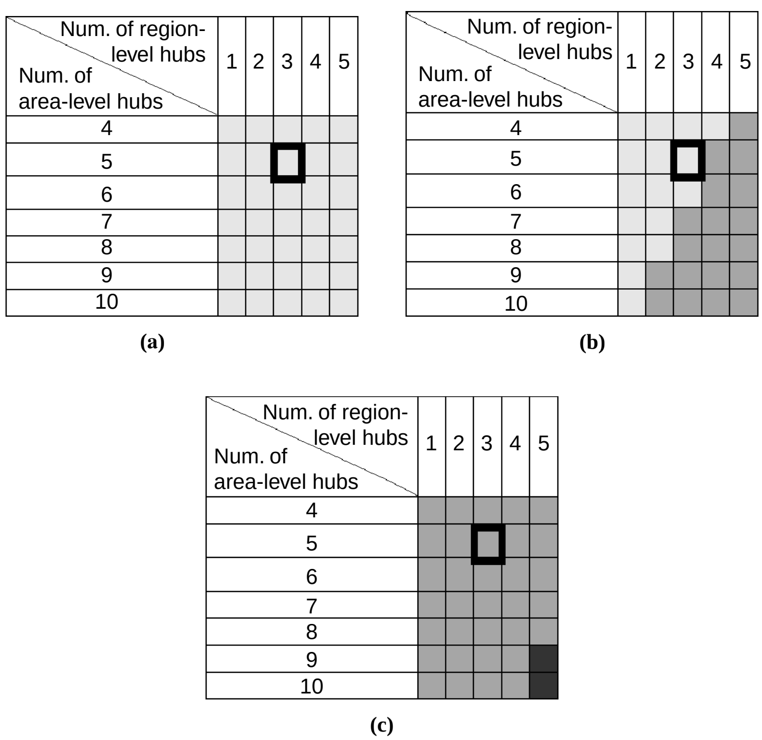

FIGURE 7 Comparison of the results of the hierarchical and the nonhierarchical models: (*a*) α = 0.5, (*b*) α = 0.6, (*c*) α = 0.7.

(continued on next page)

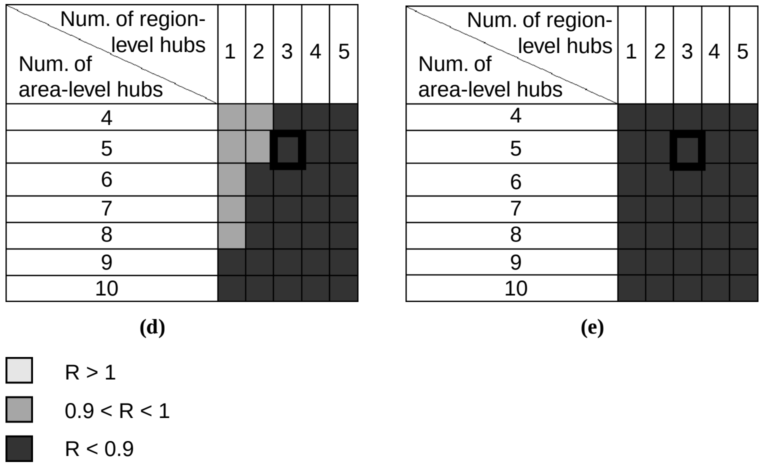

FIGURE 7 *(continued)* **Comparison of the results of the hierarchical and the nonhierarchical models: (*d*) α = 0.8, and (*e*) α = 0.9.**

route categorization, and service zone clustering to capture the critical operational issues for the transit network in an efficient manner. With the objective of minimizing the total demand-weighted travel time on the network, the proposed model not only can generate the optimal locations and hierarchies of open hubs in the target network but also yields the optimal scales of those hubs to assist planners with proper design and resource allocation. To ensure the efficiency of the model, this study has also presented a set of revised formulations and a linearization approach to reduce significantly the number of variables and solve the model with global optimality.

The model was successfully applied to the design of the hub network for the Suzhou Industrial Park in China and was shown to achieve a significant improvement in performance. Furthermore, sensitivity analyses of various hierarchical structures and travel time discount coefficients on system performance were performed. These will assist planners with choosing a proper hierarchical structure and will help them properly design transit routes.

Note that this paper has presented the findings of a preliminary evaluation and the results of a sensitivity analysis for the proposed model through a case study. More extensive tests or evaluations will be essential to assess the effectiveness of the proposed model under various transit flow patterns and hub hierarchical structures.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their constructive comments on an earlier version of this paper. The authors also acknowledge the financial support provided by National Natural Science Foundation of China.

REFERENCES

1. O'Kelly, M. E. A Quadratic Integer Program for the Location of Interacting Hub Facilities. *European Journal of Operational Research,* Vol. 32, 1987, pp. 393–404.
2. Deborah, B., and M. E. O'Kelly. Hub-and-Spoke Networks in Air Transportation: An Analytical Review. *Journal of Regional Science,* Vol. 39, 1999, pp. 275–295.
3. O'Kelly, M. E., B. Deborah, D. Skorin-Kapov, and J. Skorin-Kapov. Hub Network Design with Single and Multiple Allocation: A Computational Study. *Location Science,* Vol. 4, 1996, pp. 125–138.
4. Kim, D. K., C. H. Park, and T. J. Kim. Single Allocation Hub Network Design Model with Consolidated Traffic Flows. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2008,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 51–59.
5. Campbell, J. F. Integer Programming Formulations of Discrete Hub Location Problems. *European Journal of Operational Research,* Vol. 72, 1994, pp. 387–405.
6. O'Kelly, M. E. Hub Facility Location with Fixed Costs. *Papers in Regional Science: The Journal of the Regional Science Association International,* Vol. 71, 1992, pp. 293–306.
7. Rodriguez, V., M. J. Alvarez, and L. Barcos. Hub Location Under Capacity Constraints. *Transportation Research, Part E,* Vol. 43, 2007, pp. 495–505.
8. Aykin, T. Networking Policies for Hub-and-Spoke Systems with Applications to the Air Transportation System. *Transportation Science,* Vol. 29, 1995, pp. 201–221.
9. Sung, C. S., and H. W. Jin. Dual-Based Approach for a Hub Network Design Problem Under Non-Restrictive Policy. *European Journal of Operational Research,* Vol. 132, 2001, pp. 88–105.
10. Yu, J., Y. Liu, G.-L. Chang, and X.-G. Yang. Cluster-Based Optimization of Urban Transit Hub Locations: Methodology and Case Study in China. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2042,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 109–116.
11. Boffey, B., D. Yates, and R. D. Galvao. An Algorithm to Locate Prenatal Facilities in the Municipality of Rio de Janeiro. *Journal of the Operational Research Society,* Vol. 54, 2003, pp. 21–31.
12. Bottcher, H., and G. H. Rembold. Optimization Model for a Regional System of Solid-Waste and Wastewater Disposal Which Considers Legal Pollution Standards. *Environment and Planning, Part A,* Vol. 9, 1977, pp. 771–786.
13. Hinojosa, Y., J. Puerto, and F. R. Fernández. A Multi-Period Two-Echelon Multi-Commodity Capacitated Plant Location Problem. *European Journal of Operational Research,* Vol. 123, 2000, pp. 271–291.
14. Moore, G. C., and C. ReVelle. The Hierarchical Service Location Problem. *Management Science,* Vol. 28, No. 7, 1982, pp. 775–800.
15. Sahina, G., and H. Süralb. A Review of Hierarchical Facility Location Models. *Computers and Operations Research,* Vol. 34, 2007, pp. 2310–2331.

The Intermodal Transfer Facilities Committee sponsored publication of this paper.