Online Supplement to Approximate and exact reformulations of a capacity expansion model for single-allocation hub-and-spoke networks with congestions

S.1. Literature review

(1) HSN congestion problems

Employing capacity constraint, using a power-law cost function, and applying queuing theory are three typical ways of modeling congestion effects in the literature. Regarding hub congestion problems with considering capacity constraints, Marianov and Serra (2003) identified optimal hub locations in airline networks and formulated probabilistic constraints to limit the probability of more airplanes in a queue with a predetermined probability level. By imposing a finite queue capacity, Rahimi et al. (2016) modeled each hub as an M/M/c/K queuing system to control the congestion at this hub. Concerning hub congestions with considering power-law cost functions, Elhedhli and Hu (2005), Kian and Kargar (2016), and Alkaabneh et al. (2019) have presented such consideration in uncapacitated single allocation versions and incorporated power-law cost into the optimization objective. In multiple allocation versions, de Camargo et al. (2009) also used a power-law function to model hub congestion effects and developed a MINLP model and a generalized Benders decomposition algorithm. Moreover, by appropriately modifying the power-law function or employing other convex functions, hub congestion effects are investigated in some studies. For example, when a predefined flow threshold is exceeded, the power-law function is activated to capture hub congestion effects in de Camargo et al. (2011). Najy and Diabat (2020) introduced a concave piecewise-linear function and a convex piecewise-linear function to capture economies of scale and hub congestion effects, respectively, and examined the impact of these two conflicting effects on HSN design. Addressing prevalent queuing systems in modeling hub congestions, Elhedhli and Wu (2010) presented M/M/1 queuing systems for the single allocation p-hub location with capacity selection and represented the congestion cost as a Kleinrock average delay function. Based on M/M/c queuing systems, Zhalechian et al. (2017) stated a fuzzy multi-objective hub location problem, in which responsiveness and social responsibility under uncertainty are integrated into decisions. In addition, by modeling hubs as M/G/1 queue systems and representing congestion levels as expected queue lengths at hubs, Azizi et al. (2018) and Bhatt et al. (2021) both studied single allocation p-hub location problems with stochastic demand, congestion, and capacity selection.

However, studies considering simultaneous congestions in both hubs and hub links are limited. Hu et al. (2018) described the joint design of the fleet size and the number, locations, and capacities of hubs in a multiple allocation HSN and explicitly modeled node and road congestions by imposing node capacity constraints and treating the route travel time as an increasing function of the number of trucks on the route, respectively. Mohammadi et al. (2019) focused on hub and link uncertainties, and regarded each hub as an M/M/1 queue system and modeled congestion on links through the BPR function to address stochastic waiting time at hubs and transportation time on links. Karimi-Mamaghan et al. (2020) presented a bi-objective MINLP model for a single allocation multi-commodity HSN problem under congestions, in which a general GI/G/c queuing system is used to account for the hub congestion of the flow for multi-product and a stochastic traffic model employing the BPR function is utilized to evaluate the congestion of flow traversing hub links.

With respect to solution methods for HSN design under congestions, most studies resorted to heuristic, relaxation, and decomposition methods because addressing congestion effects typically results in MINLP formulations that are very difficult to solve. For example, Elhedhli and Hu (2005), and Alkaabneh et al. (2019) developed piecewise tangent approximation along with Lagrangian heuristic, and Lagrangian heuristic as well as greedy randomized adaptive search procedure to solve the proposed model, respectively. De Camargo et al. (2011) presented a combined algorithm that considers outer approximation and Benders decomposition methods to tackle the established nonlinear model. Karimi-Mamaghan et al. (2020) focused on a novel metaheuristic that hybridizes non-dominated sorting genetic algorithm-II and a learning-based iterated local search algorithm to deal with large-sized problem instances. However, studies developing exact solution methods that do not rely on any relaxation or decomposition and linear approximation methods with provable and desirable approximation accuracy guarantees are rare. To pursue this, only Kian and Kargar (2016) and Bhatt et al. (2021) developed reformulations by applying SOCP technologies for the hub location problem with nonlinear congestion functions. In general, these reformulation methods are easier to be followed and favored by practitioners since these reformulations can be solved directly using solvers such as GUROBI and CPLEX. Therefore, in this regard, it deserves more attention to explore tractable reformulations resting on the solvers when addressing HSN design under congestions.

(2) Capacity expansion problems

The study on capacity expansion problems has received a growing concern in various areas such as air transportation, rail freight, telecommunication, manufacturing, and express delivery (Magnanti and Wong, 1984; Marín and Jaramillo, 2008; Taghavi and Huang, 2016; Zhao et al., 2018). Bärmann et al. (2017) built a mixed-integer programming formulation to study a multiperiod expansion of the German railway network, in which an optimal investment strategy is decided to upgrade the most profitable rail freight lines. They developed a novel multiple-knapsack decomposition approach to solve the formulated model. Şafak et al. (2022) proposed a two-stage stochastic MINLP for the airline flight network expansion by introducing new flights and considering the impact of departure time decisions on the probability distribution of random demands. Strong SOCP and McCormick inequalities, and an exact scenario group-wise decomposition approach along with a new bounding method are devised to solve their model. By upgrading existing spokes into temporary hubs, Zhang and Liu (2022) established a two-stage bi-objective robust model to expand the hub capacity in HSN structured express transportation during periods with surging demands. An enhanced column-and-constraint generation algorithm is designed, and a case study on a realistic express logistics company in China is conducted in their study. To effectively handle surged package volumes due to unprecedented supply chain disruptions and extraordinary growth in online shopping, Ashraf et al. (2022) investigated the occurrence of the Braess Paradox in third-party logistics when adding sort routes to expand hub capacities.

Since capacity expansion decisions significantly affect congestion effects, related studies have developed different formulations to model their interactions. For a multi-commodity tree network design arising in telecommunication and transportation networks, Miranda et al. (2011) took capacity expansion and congestion effects into account by using a modified function that convexified a series of Kleinrock delay functions, and presented a generalized Benders decomposition to tackle the nonlinear model. Liu and Wang (2015) addressed a continuous road network design problem in traffic management under stochastic user equilibrium, in which the network performance is optimized by expanding road capacity, and a BPR travel time function is employed to capture the relationship between congestion and capacity expansion. Due to the resulting nonlinear nonconvex programming, they developed outer approximation techniques and a global optimization solution algorithm based on a range reduction technique to solve the problem. By incorporating airport delay levels as functions of capacity utilization rates (or capacity expansion decisions), Sun and Schonfeld (2015) formulated a stochastic MINLP to optimize airport capacity expansions under demand uncertainty, and designed an outer approximation method to solve the model. Paraskevopoulos et al. (2016) described a fixed-charge multi-commodity network design problem with simultaneous consideration of node congestion and node capacity expansion. Specifically, a BPR congestion cost function is used to model the relationship between the total amount of flows via a node and the upgrade capacity of this node, and a SOCP reformulation and an evolutionary algorithm are derived to solve the resulting nonlinear integer programming.

Although capacity expansion problems have been extensively investigated in different network topologies, studies addressing such problems in the context of HSNs appear to be lacking. Meanwhile, addressing the interaction and integration between capacity expansion and congestion effects in an HSN topology needs further study. In particular, since the consideration of capacity expansion decisions in general leads to a multivariable nonlinear congestion function, it is necessary to further exploit more effective solution methods such as reformulation methods and exact solution algorithms.

Table S-1. Review of related studies.

Reference	Modeling viewpoint									Methodological viewpoint				
	Network topology		Decision problem		Congestion		Model			Linear approximation		Exact method		Heuristic method
	HSN	Multi-commodity/other networks	Network design	Network expansion	Hub/ node	Hub link/arc	MILP	Convex MINLP	Non-convex MINLP	Error non-guarantee	Error guarantee	SOCP	Decomposition or others	
Marianov and Serra (2003)	✓		✓		✓		✓							✓
Elhedhli and Hu (2005)	✓		✓		✓			✓		✓				✓
De Camargo et al. (2009)	✓		✓		✓			✓					✓	
Elhedhli and Wu (2010)	✓		✓		✓				✓	✓				✓
De Camargo et al. (2011)	✓		✓		✓			✓					✓	
Miranda et al. (2011)		✓		✓		✓		✓					✓	
Liu and Wang (2015)		✓		✓		✓			✓	✓			✓	
Sun and Schonfeld (2015)		✓		✓	✓			✓					✓	
Kian and Kargar (2016)	✓		✓		✓			✓				✓		
Paraskevopoulos et al. (2016)		✓		✓	✓				✓			✓		✓
Bärmann et al. (2017)		✓		✓			✓						✓	
Azizi et al. (2018)	✓		✓		✓				✓	✓			✓	✓
Hu et al. (2018)	✓		✓		✓	✓			✓					✓
Alkaabneh et al. (2019)	✓		✓		✓			✓		✓				✓
Mohammadi et al. (2019)	✓		✓		✓	✓			✓	✓				✓
Najy and Diabat (2020)	✓		✓		✓		✓						✓	
Karimi-Mamaghan et al. (2020)	✓	✓	✓		✓	✓			✓	✓				✓
Bhatt et al. (2021)	✓		✓		✓				✓			✓		
Şafak et al. (2022)		✓		✓					✓			✓	✓	
Zhang and Liu (2022)	✓			✓			✓						✓	
This work	✓			✓	✓	✓			✓		✓	✓		

S.2. Proof of Proposition 1

Proof. Given any $S_k \in \left[2^{b-1}, 2^b\right]$, $b = 1,2,\ldots,|B_k|-2$, $\forall k \in H$, the linear equation via points $\left(2^{b-1}, \frac{2^{b-1}}{1+2^{b-1}}\right)$ and $\left(2^b, \frac{2^b}{1+2^b}\right)$ is expressed as

$$\tilde{F}(S_k) = \frac{\frac{2^b}{1+2^b} - \frac{2^{b-1}}{1+2^{b-1}}}{2^b - 2^{b-1}} \cdot \left(S_k - 2^{b-1}\right) + \frac{2^{b-1}}{1+2^{b-1}} = \frac{S_k + 2^{2b-1}}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}, \forall k \in H .$$

Then, we can write the following equation (S-1).

$$\frac{F(S_k) - \tilde{F}(S_k)}{F(S_k)} = \frac{\frac{S_k}{1+S_k} - \frac{S_k + 2^{2b-1}}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}}{\frac{S_k}{1+S_k}}, \forall k \in H \qquad \text{(S-1)}$$

Taking the first-order derivative of (S-1) on S_k, we obtain

$$\frac{d\left(\frac{F(S_k) - \tilde{F}(S_k)}{F(S_k)}\right)}{d(S_k)} = \frac{\frac{2^{2b-1}}{(S_k)^2} - 1}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}, \forall k \in H .$$

Let the first-order derivative of (S-1) be 0, we can derive that $S_k = \sqrt{2^{2b-1}}$, $\forall k \in H$. Taking the second-order derivative of (S-1) on S_k, we obtain

$$\frac{d\left(\frac{\frac{2^{2b-1}}{(S_k)^2} - 1}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}\right)}{d(S_k)} = \frac{-2 \cdot 2^{2b-1}}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)\cdot(S_k)^3} < 0, \forall k \in H .$$

Therefore, $\frac{F(S_k) - \tilde{F}(S_k)}{F(S_k)}$ achieves the maximum value at $S_k = \sqrt{2^{2b-1}}$ ($\forall k \in H$), namely

$$\sup_{S_k \in \left[2^{b-1}, 2^b\right]} \left[\frac{F(S_k) - \tilde{F}(S_k)}{F(S_k)}\right] = 1 - \frac{\left(2^{b-0.5}+1\right)^2}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}, \forall k \in H .$$

Moreover, it is noted that the maximum value of $\frac{F(S_k) - \tilde{F}(S_k)}{F(S_k)}$ is a function of b. Taking the first-order derivative of this maximum value on b, we obtain that

$$\frac{d\left(1-\frac{\left(2^{b-0.5}+1\right)^2}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}\right)}{d(b)}=\frac{\left(2^{b-0.5}+1\right)\cdot\ln(2)\cdot 2^{b-1}\cdot\left[\left(\sqrt{2}+1-2^{2.5}\right)\cdot 2^{b-1}+3-2^{1.5}\right]}{\left(1+2^b\right)^2\cdot\left(1+2^{b-1}\right)^2}<0\text{, }(b\geq 1).$$

Thus, the maximum value of $\frac{F(S_k)-\tilde{F}(S_k)}{F(S_k)}$ monotonically decreases on b ($b\geq 1$). That is, $1-\frac{\left(2^{b-0.5}+1\right)^2}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}\leq 1-\frac{\left(2^{1-0.5}+1\right)^2}{\left(1+2^1\right)\cdot\left(1+2^{1-1}\right)}=0.0286$. In addition, since $F(S_k)\geq\tilde{F}(S_k)$ (see Figure S-1), $F(S_k)>0$ and $\tilde{F}(S_k)>0$, it is derived that $\varepsilon_k^0=\left|\frac{F(S_k)-\tilde{F}(S_k)}{F(S_k)}\right|=\frac{F(S_k)-\tilde{F}(S_k)}{F(S_k)},\forall k\in H$. Then, $\varepsilon_k^0\leq 1-\frac{\left(2^{b-0.5}+1\right)^2}{\left(1+2^b\right)\cdot\left(1+2^{b-1}\right)}\leq 0.0286<3\%$ $(b\geq 1)$.

Finally, considering that $S_k\in\left[2^{b-1},2^b\right]$, the increase in S_k naturally results in the increase in the value of b. Thus, as the value of S_k increases, the value of ε_k^0 decreases because the maximum value of $\frac{F(S_k)-\tilde{F}(S_k)}{F(S_k)}$ monotonically decreases on b ($b\geq 1$). □

S.3. Proof of Proposition 2

Proof. Given any $S_k\in\left[2^{b-1},2^b\right]$, $b=1,2,...,|B_k|-2$, $\forall k\in H$, the following equation system (S-2) can be constructed based on the H-LA reformulation.

$$\begin{cases}\sum_{l\in L}r_{k,l}=\frac{2^{b-1}}{1+2^{b-1}}\cdot g_{k,b-1}^0+\frac{2^b}{1+2^b}\cdot g_{k,b}^0,\forall k\in H\\ g_{k,b-1}^0+g_{k,b}^0=1,\forall k\in H\end{cases}\tag{S-2}$$

By solving equation system (S-2), we obtain that $g_{k,b-1}^0=\frac{\left[2^b-\left(1+2^b\right)\cdot\sum_{l\in L}r_{k,l}\right]\cdot\left(1+2^{b-1}\right)}{2^b-2^{b-1}}$ and $g_{k,b}^0=\frac{\left[\left(1+2^{b-1}\right)\cdot\sum_{l\in L}r_{k,l}-2^{b-1}\right]\cdot\left(1+2^b\right)}{2^b-2^{b-1}}$. Then, $\hat{S}_k$ can be correspondingly expressed as

$$\hat{S}_k=\left(1+2^{b-1}+2^b+2^{b-1}\cdot 2^b\right)\cdot\sum_{l\in L}r_{k,l}-2^{b-1}\cdot 2^b,\forall k\in H.$$

Subsequently, it is proved that $\varepsilon_k^1=\left|\frac{S_k-\hat{S}_k}{S_k}\right|=\frac{\hat{S}_k-S_k}{S_k},\forall k\in H$. Denoting $\hat{S}_k-S_k=\frac{-\left[\left(1+2^{b-1}\right)\cdot\sum_{l\in L}r_{k,l}-2^{b-1}\right]\cdot\left[\left(1+2^b\right)\cdot\sum_{l\in L}r_{k,l}-2^b\right]}{1-\sum_{l\in L}r_{k,l}}$ ($\forall k\in H$) and taking derivative of $\hat{S}_k-S_k$ on

$\sum_{l\in L} r_{k,l}$, we can derive that $\hat{S}_k - S_k$ is concave when $\sum_{l\in L} r_{k,l} \in \left[\frac{2^{b-1}}{1+2^{b-1}}, \frac{2^b}{1+2^b}\right]$. This means that it has maximum value and minimum value. It is noted that its minimum value is equal to 0 when $\sum_{l\in L} r_{k,l} = \frac{2^{b-1}}{1+2^{b-1}}$ or $\sum_{l\in L} r_{k,l} = \frac{2^b}{1+2^b}$. Thus, $\hat{S}_k - S_k \geq 0$. Again, since $S_k > 0$, $\varepsilon_k^1 = \left|\frac{S_k - \hat{S}_k}{S_k}\right| = \frac{\hat{S}_k - S_k}{S_k}, \forall k \in H$.

Then, we can further construct the following equation (S-3).

$$\frac{\hat{S}_k - S_k}{S_k} = -\left(1+2^{b-1}\right)\cdot\left(1+2^b\right)\cdot\sum_{l\in L} r_{k,l} + \left(2^{b-1}+2^b+2\cdot 2^{b-1}\cdot 2^b\right) - \frac{2^{b-1}\cdot 2^b}{\sum_{l\in L} r_{k,l}}, \forall k \in H \qquad \text{(S-3)}$$

Taking the first-order derivative of (S-3) on $\sum_{l\in L} r_{k,l}$, we obtain that

$$\frac{d\left(\frac{\hat{S}_k - S_k}{S_k}\right)}{d\left(\sum_{l\in L} r_{k,l}\right)} = -\left(1+2^{b-1}\right)\cdot\left(1+2^b\right) + \frac{2^{b-1}\cdot 2^b}{\left(\sum_{l\in L} r_{k,l}\right)^2}, \forall k \in H.$$

Let the first-order derivative of (S-3) be 0, we can derive that $\sum_{l\in L} r_{k,l} = \sqrt{\frac{2^{b-1}\cdot 2^b}{\left(1+2^{b-1}\right)\cdot\left(1+2^b\right)}}$, $\forall k \in H$.

Taking the second-order derivative of (S-3) on $\sum_{l\in L} r_{k,l}$, we can demonstrate that

$$\frac{d^2\left(\frac{\hat{S}_k - S_k}{S_k}\right)}{d\left(\sum_{l\in L} r_{k,l}\right)} = -\frac{2\cdot 2^{b-1}\cdot 2^b}{\left(\sum_{l\in L} r_{k,l}\right)^3} < 0, \forall k \in H.$$

Therefore, $\frac{\hat{S}_k - S_k}{S_k}$ achieves the maximum value at $\sum_{l\in L} r_{k,l} = \sqrt{\frac{2^{b-1}\cdot 2^b}{\left(1+2^{b-1}\right)\cdot\left(1+2^b\right)}}$ ($\forall k \in H$), namely

$$\sup_{\sum_{l\in L} r_{k,l} \in \left[\frac{2^{b-1}}{1+2^{b-1}}, \frac{2^b}{1+2^b}\right]} \left[\frac{\hat{S}_k - S_k}{S_k}\right] = \left(2^{b-1}+2^b+2\cdot 2^{b-1}\cdot 2^b\right) - 2\cdot\sqrt{2^{b-1}\cdot 2^b\cdot\left(1+2^{b-1}\right)\cdot\left(1+2^b\right)}, \forall k \in H.$$

Then, it is obtained that $\varepsilon_k^1 \leq \left(2^{b-1}+2^b+2\cdot 2^{b-1}\cdot 2^b\right) - 2\cdot\sqrt{2^{b-1}\cdot 2^b\cdot\left(1+2^{b-1}\right)\cdot\left(1+2^b\right)}$ ($b \geq 1$). □

S.4. Proof of Proposition 3

Proof. Note that the congestion cost for hubs can be essentially represented as the term $\sum_{k\in H} c_k^0 \cdot \hat{S}_k$ subjected to additional constraints in the H-LA reformulation. In the original model [M1], the congestion cost for hubs can be expressed as the term $\sum_{k\in H} c_k^0 \cdot S_k$. Since $\hat{S}_k - S_k \geq 0$, which has been proved in Proposition 2, $\sum_{k\in H} c_k^0 \cdot \hat{S}_k \geq \sum_{k\in H} c_k^0 \cdot S_k$. Therefore, the solution to the H-LA reformulation provides the upper bound of the congestion cost for hubs, compared to the original model [M1]. □

S.5. Proof of Proposition 4

Proof. Since $\sum_{l\in L} z_k^l = 1$ and $z_k^l \in \{0,1\}$, it must guarantee that for each k, we obtain $z_k^l = 1$ for some $l \in L$ and $z_k^{l*} = 0$ for all other $l* \in L \setminus \{l\}$. Then, it is derived that $\sum_{l\in L} \Gamma_k^l \cdot z_k^l \in \left\{\Gamma_k^1, \Gamma_k^2, ..., \Gamma_k^{|L|}\right\}$, $\forall k \in H$. Furthermore, it is easy to demonstrate that $1/\left(\sum_{l\in L} \Gamma_k^l \cdot z_k^l\right) \in \left\{1/\Gamma_k^1, 1/\Gamma_k^2, ..., 1/\Gamma_k^{|L|}\right\}$. Meanwhile, it is observed that $\sum_{l\in L} \left(1/\Gamma_k^l \right) \cdot z_k^l \in \left\{1/\Gamma_k^1, 1/\Gamma_k^2, ..., 1/\Gamma_k^{|L|}\right\}$ based on the fact that $\sum_{l\in L} z_k^l = 1$ and $z_k^l \in \{0,1\}$. These suggest that $1/\left(\sum_{l\in L} \Gamma_k^l \cdot z_k^l\right)$ and $\sum_{l\in L} \left(1/\Gamma_k^l \right) \cdot z_k^l$ take the value from the same set. Finally, it is shown that $1/\left(\sum_{l\in L} \Gamma_k^l \cdot z_k^l\right) = 1/\Gamma_k^{l*}$ and $\sum_{l\in L} \left(1/\Gamma_k^l \right) \cdot z_k^l = 1/\Gamma_k^{l*}$ when $z_k^{l*} = 1$ ($l* \in L$). Therefore, $1/\left(\sum_{l\in L} \Gamma_k^l \cdot z_k^l\right) = \sum_{l\in L} \left(1/\Gamma_k^l \right) \cdot z_k^l$, $\forall k \in H$. Then, we can derive that

$$\begin{aligned}
\sum_{i\in N} O_i \cdot x_{i,k} \Big/ \sum_{l\in L} \Gamma_k^l \cdot z_k^l &= \left(\sum_{i\in N} O_i \cdot x_{i,k}\right) \cdot \left[1/\left(\sum_{l\in L} \Gamma_k^l \cdot z_k^l\right)\right] = \left(\sum_{i\in N} O_i \cdot x_{i,k}\right) \cdot \left[\sum_{l\in L} \left(1/\Gamma_k^l \right) \cdot z_k^l\right] \\
&= \sum_{i\in N} \sum_{l\in L} O_i \cdot \left(1/\Gamma_k^l \right) \cdot x_{i,k} \cdot z_k^l
\end{aligned}$$ □

S.6. Proof of Proposition 5

Proof. For the proof of this Proposition, we can refer to Theorem 1 in Fügenschuh et al. (2015). □

S.7. Proof of Proposition 6

Proof. For a given hub link $k-m$, the set $\zeta_{k,m}$ consists of the following points: $\left(\Lambda_{k,m}^1, \ln\left(\Lambda_{k,m}^1\right)\right)$, $\left(\Lambda_{k,m}^2, \ln\left(\Lambda_{k,m}^2\right)\right)$, …, $\left(\Lambda_{k,m}^{|Q|}, \ln\left(\Lambda_{k,m}^{|Q|}\right)\right)$. It is easy to show that these points are feasible for inequalities (73) and (74). These points can actually be expressed as the intersection of inequalities from (73) and (74). However, each intersection of two arbitrary inequalities may lead to an infeasible point because such a point is cut off by some other inequalities from (73) and (74). To identify such points that are outside the convex hull, we check the following two cases.

Case 1. It represents the selection of two arbitrary inequalities from (73). For the convenience of expression, we define coefficients $\theta_{k,m}^q = \dfrac{\ln\left(\Lambda_{k,m}^q\right) - \ln\left(\Lambda_{k,m}^{q-1}\right)}{\Lambda_{k,m}^q - \Lambda_{k,m}^{q-1}}$ and $\vartheta_{k,m}^q = -\theta_{k,m}^q \cdot \Lambda_{k,m}^{q-1} + \ln\left(\Lambda_{k,m}^{q-1}\right)$ for $q = 2, ..., |Q|$. Then, equation system (S-4) is given, where $q*, q** \in \{2, ..., |Q|\}$ and $q* \neq q**$.

$$\begin{cases} lu_{k,m} = \theta_{k,m}^{q*} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} \right) + \vartheta_{k,m}^{q*} \\ lu_{k,m} = \theta_{k,m}^{q**} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} \right) + \vartheta_{k,m}^{q**} \end{cases} \tag{S-4}$$

(1) Assume $|q*-q**|=1$, without loss of generality, let $q**=q*+1$, $q*\in\{2,...,|Q|-1\}$. Then, the solution of (S-4) is $\left(\Lambda_{k,m}^{q*}, \ln\left(\Lambda_{k,m}^{q*}\right)\right)$, $q*\in\{2,...,|Q|-1\}$. Obviously, these solutions are points in $\zeta_{k,m}$.

(2) Assume $|q*-q**|>1$, without loss of generality, let $q**=q*+qq$, $qq>1$. Then, the solution of (S-4) is $\left(\frac{\vartheta_{k,m}^{q**} - \vartheta_{k,m}^{q*}}{\theta_{k,m}^{q*} - \theta_{k,m}^{q**}}, \frac{\theta_{k,m}^{q*} \cdot \vartheta_{k,m}^{q**} - \theta_{k,m}^{q**} \cdot \vartheta_{k,m}^{q*}}{\theta_{k,m}^{q*} - \theta_{k,m}^{q**}} \right)$. Obviously, it is easy to see that such a solution violates at least one inequality from (73): $lu_{k,m} \le \theta_{k,m}^{q} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} \right) + \vartheta_{k,m}^{q}$, $q\in\{2,...,|Q|\}/\{q*,q**\}$. For example, let $qq=2$, the obtained solution violates inequality $lu_{k,m} \le \theta_{k,m}^{q*+1} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} \right) + \vartheta_{k,m}^{q*+1}$. Thus, this solution does not belong to $\zeta_{k,m}$.

Case 2. It represents the selection of one arbitrary inequality from (73) and the one from (74). Then, equation system (S-5) is given, where $q\in\{2,...,|Q|\}$.

$$\begin{cases} lu_{k,m} = \theta_{k,m}^{q} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} \right) + \vartheta_{k,m}^{q} \\ lu_{k,m} = \frac{\ln\left(\Lambda_{k,m}^{|Q|}\right) - \ln\left(\Lambda_{k,m}^{1}\right)}{\Lambda_{k,m}^{|Q|} - \Lambda_{k,m}^{1}} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} - \Lambda_{k,m}^{1} \right) + \ln\left(\Lambda_{k,m}^{1}\right) \end{cases} \tag{S-5}$$

(1) When $q=1$, the solution of (S-5) is $\left(\Lambda_{k,m}^{1}, \ln\left(\Lambda_{k,m}^{1}\right)\right)$, which is a point in $\zeta_{k,m}$.

(2) When $q=|Q|$, the solution of (S-5) is $\left(\Lambda_{k,m}^{|Q|}, \ln\left(\Lambda_{k,m}^{|Q|}\right)\right)$, which is a point in $\zeta_{k,m}$.

(3) When $q\in\{2,...,|Q|-1\}$, the solution of (S-5) violates at least one inequality from (73): $lu_{k,m} \le \theta_{k,m}^{q*} \cdot \left(\sum_{q\in Q} \Lambda_{k,m}^{q} \cdot u_{k,m}^{q} \right) + \vartheta_{k,m}^{q*}$, $q*\ne q$ and $q*\in\{2,...,|Q|\}$, which is similar to that of (2) in Case 1. Thus, the solution of (S-5) does not belong to $\zeta_{k,m}$. □

S.8. Proof of Proposition 7

Proof. Obvious from the construction. □

S.9. Proof of Proposition 8

Proof. Given any $E_{k,m} \in \left[2^{b''-1}, 2^{b''}\right]$, $b'' = 1, 2, \ldots, \left|B''_{k,m}\right| - 2$, $\forall k, m \in H, k \neq m$, the linear equation via points $\left(2^{b''-1}, \left(2^{b''-1}\right)^{\beta'}\right)$ and $\left(2^{b''}, \left(2^{b''}\right)^{\beta'}\right)$ is expressed as

$$\tilde{G}\left(E_{k,m}\right) = \frac{2^{\beta' b''} \cdot \left(1 - 2^{-\beta'}\right)}{2^{b''-1}} \cdot E_{k,m} + 2^{\beta' b''} \cdot \left(2^{1-\beta'} - 1\right), \forall k, m \in H, k \neq m .$$

Then, we can construct the following equation (S-6).

$$\frac{G\left(E_{k,m}\right) - \tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)} = 1 - \frac{\frac{2^{\beta' b''} \cdot \left(1 - 2^{-\beta'}\right)}{2^{b''-1}} \cdot E_{k,m} + 2^{\beta' b''} \cdot \left(2^{1-\beta'} - 1\right)}{\left(E_{k,m}\right)^{\beta'}}, \forall k, m \in H, k \neq m \tag{S-6}$$

Taking the first-order derivative of (S-6) on $E_{k,m}$, we obtain that

$$\frac{d\left(\frac{G\left(E_{k,m}\right) - \tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}\right)}{d\left(E_{k,m}\right)} = -\frac{\frac{2^{\beta' b''} \cdot \left(1 - 2^{-\beta'}\right)}{2^{b''-1}} \cdot \left(1 - \beta'\right) - 2^{\beta' b''} \cdot \left(2^{1-\beta'} - 1\right) \cdot \beta' \cdot \left(E_{k,m}\right)^{-1}}{\left(E_{k,m}\right)^{\beta'}}, \forall k, m \in H, k \neq m .$$

Let the first-order derivative of (S-6) be 0, we can derive that $E_{k,m} = \frac{2^{b''-1} \cdot \left(2^{1-\beta'} - 1\right) \cdot \beta'}{\left(1 - 2^{-\beta'}\right) \cdot \left(1 - \beta'\right)}$, $\forall k, m \in H, k \neq m$. Taking the second-order derivative of (S-6) on $E_{k,m}$, we can show that

$$\frac{d^2\left(\frac{G\left(E_{k,m}\right) - \tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}\right)}{d\left(E_{k,m}\right)} < 0, \forall k, m \in H, k \neq m .$$

Therefore, $\frac{G\left(E_{k,m}\right) - \tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}$ achieves the maximum value at $E_{k,m} = \frac{2^{b''-1} \cdot \left(2^{1-\beta'} - 1\right) \cdot \beta'}{\left(1 - 2^{-\beta'}\right) \cdot \left(1 - \beta'\right)}$ ($\forall k, m \in H, k \neq m$), namely

$$\sup_{E_{k,m} \in \left[2^{b''-1}, 2^{b''}\right]} \left[\frac{G\left(E_{k,m}\right) - \tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}\right] = 1 - \left(\frac{2 - 2^{\beta'}}{1 - \beta'}\right) \cdot \left[\frac{\left(2^{1-\beta'} - 1\right) \cdot \beta'}{\left(1 - 2^{-\beta'}\right) \cdot \left(1 - \beta'\right)}\right]^{-\beta'}, \forall k, m \in H, k \neq m .$$

Moreover, it is noted that the maximum value of $\frac{G\left(E_{k,m}\right) - \tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}$ is a function of β'. If we consider the situation of $\beta' \in (0,1)$, then the following equation can be obtained by taking the first-order derivative of this maximum value on β' and letting the first-order derivative be 0. As a result, it is derived that $\beta' = 0.5$.

$$\frac{d\left(1-\left(\frac{2-2^{\beta'}}{1-\beta'}\right)\cdot\left[\frac{\left(2^{1-\beta'}-1\right)\cdot\beta'}{\left(1-2^{-\beta'}\right)\cdot\left(1-\beta'\right)}\right]^{-\beta'}\right)}{d\left(\beta'\right)}=0$$

Further, taking the second-order derivative of this maximum value on β', it shows that

$$\frac{d^2\left(1-\left(\frac{2-2^{\beta'}}{1-\beta'}\right)\cdot\left[\frac{\left(2^{1-\beta'}-1\right)\cdot\beta'}{\left(1-2^{-\beta'}\right)\cdot\left(1-\beta'\right)}\right]^{-\beta'}\right)}{d\left(\beta'\right)}<0,\quad \beta'\in(0,1).$$

Thus, the maximum value of $\frac{G\left(E_{k,m}\right)-\tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}$ achieves the corresponding maximum value at $\beta'=0.5$. That is, $1-\left(\frac{2-2^{\beta'}}{1-\beta'}\right)\cdot\left[\frac{\left(2^{1-\beta'}-1\right)\cdot\beta'}{\left(1-2^{-\beta'}\right)\cdot\left(1-\beta'\right)}\right]^{-\beta'}\leq 1-\left(\frac{2-2^{0.5}}{1-0.5}\right)\cdot\left[\frac{\left(2^{1-0.5}-1\right)\cdot 0.5}{\left(1-2^{-0.5}\right)\cdot\left(1-0.5\right)}\right]^{-0.5}=0.0148$ ($0<\beta'<1$).

Finally, since $G\left(E_{k,m}\right)\geq\tilde{G}\left(E_{k,m}\right)$ (see Figure S-4), $G\left(E_{k,m}\right)>0$ and $\tilde{G}\left(E_{k,m}\right)>0$, it is derived that $\varepsilon_{k,m}^2=\left|\frac{G\left(E_{k,m}\right)-\tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)}\right|=\frac{G\left(E_{k,m}\right)-\tilde{G}\left(E_{k,m}\right)}{G\left(E_{k,m}\right)},\forall k,m\in H,k\neq m$. Given $\beta'=\frac{1}{\beta+1}=\frac{1}{4+1}=0.2$ in this work, it is easy to demonstrate that $\varepsilon_{k,m}^2\leq 1-\left(\frac{2-2^{\beta'}}{1-\beta'}\right)\cdot\left[\frac{\left(2^{1-\beta'}-1\right)\cdot\beta'}{\left(1-2^{-\beta'}\right)\cdot\left(1-\beta'\right)}\right]^{-\beta'}=0.0095<1\%$, based on the above derivation. □

S.10. Proof of Proposition 9

Proof. Given any $E_{k,m}\in\left[2^{b''-1},2^{b''}\right]$, $b''=1,2,\ldots,\left|B_{k,m}''\right|-2$, $\forall k,m\in H,k\neq m$, the following equation system (S-7) can be constructed based on the HL-LA-2 reformulation.

$$\begin{cases}\sum_{i\in N}y_{k,m}^i=\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}\cdot g_{k,m,b''-1}^1+\left(2^{b''}\right)^{\frac{1}{\beta+1}}\cdot g_{k,m,b''}^1,\forall k,m\in H,k\neq m\\ g_{k,m,b''-1}^1+g_{k,m,b''}^1=1,\forall k,m\in H,k\neq m\end{cases}\qquad\text{(S-7)}$$

By solving equation system (S-7), we obtain that $g_{k,m,b''-1}^1=\frac{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\sum_{i\in N}y_{k,m}^i}{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}}$ and $g_{k,m,b''}^1=\frac{\sum_{i\in N}y_{k,m}^i-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}}{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}}$. Then, $\hat{E}_{k,m}$ can be correspondingly expressed as

$$\hat{E}_{k,m}=\frac{2^{b''-1}\cdot\left(2^{b''}\right)^{\frac{1}{\beta+1}}-2^{b''}\cdot\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}+\left(2^{b''}-2^{b''-1}\right)\cdot\sum_{i\in N}y_{k,m}^{i}}{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}},\forall k,m\in H,k\neq m\,.$$

Subsequently, it is proved that $\varepsilon_{k,m}^{3}=\left|\frac{E_{k,m}-\hat{E}_{k,m}}{E_{k,m}}\right|=\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}},\forall k,m\in H,k\neq m$. Denoting

$$\hat{E}_{k,m}-E_{k,m}=\frac{2^{b''-1}\cdot\left(2^{b''}\right)^{\frac{1}{\beta+1}}-2^{b''}\cdot\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}+\left(2^{b''}-2^{b''-1}\right)\cdot\sum_{i\in N}y_{k,m}^{i}-\left[\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}\right]\cdot\left(\sum_{i\in N}y_{k,m}^{i}\right)^{\beta+1}}{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}}$$ ($\forall k,m\in H,k\neq m$) and

taking derivative of $\hat{E}_{k,m}-E_{k,m}$ on $\sum_{i\in N}y_{k,m}^{i}$, we can derive that $\hat{E}_{k,m}-E_{k,m}$ is concave when $\sum_{i\in N}y_{k,m}^{i}\in\left[\left(2^{b''-1}\right)^{\frac{1}{\beta+1}},\left(2^{b''}\right)^{\frac{1}{\beta+1}}\right]$. This means that it has maximum value and minimum value. It is noted that its minimum value is equal to 0 when $\sum_{i\in N}y_{k,m}^{i}=\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}$ or $\sum_{i\in N}y_{k,m}^{i}=\left(2^{b''}\right)^{\frac{1}{\beta+1}}$. Thus, $\hat{E}_{k,m}-E_{k,m}\geq 0$. Again, since $E_{k,m}>0$, $\varepsilon_{k,m}^{3}=\left|\frac{E_{k,m}-\hat{E}_{k,m}}{E_{k,m}}\right|=\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}},\forall k,m\in H,k\neq m$.

Then, we can further construct the following equation (S-8).

$$\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}}=\frac{\left[2^{b''-1}\cdot\left(2^{b''}\right)^{\frac{1}{\beta+1}}-2^{b''}\cdot\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}\right]\cdot\left(\sum_{i\in N}y_{k,m}^{i}\right)^{-(\beta+1)}+\left(2^{b''}-2^{b''-1}\right)\cdot\left(\sum_{i\in N}y_{k,m}^{i}\right)^{-\beta}}{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}} -1,\forall k,m\in H,k\neq m \tag{S-8}$$

Taking the first-order derivative of (S-8) on $\sum_{i\in N}y_{k,m}^{i}$, we obtain that

$$\frac{d\left(\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}}\right)}{d\left(\sum_{i\in N}y_{k,m}^{i}\right)}=\frac{-(\beta+1)\cdot\left[2^{b''-1}\cdot\left(2^{b''}\right)^{\frac{1}{\beta+1}}-2^{b''}\cdot\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}\right]\cdot\left(\sum_{i\in N}y_{k,m}^{i}\right)^{-(\beta+2)}-\beta\cdot\left(2^{b''}-2^{b''-1}\right)\cdot\left(\sum_{i\in N}y_{k,m}^{i}\right)^{-(\beta+1)}}{\left(2^{b''}\right)^{\frac{1}{\beta+1}}-\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}},$$
$$\forall k,m\in H,k\neq m$$

Let the first-order derivative of (S-8) be 0, we can derive that $\sum_{i\in N}y_{k,m}^{i}=\frac{(\beta+1)\cdot\left[2^{b''}\cdot\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}-2^{b''-1}\cdot\left(2^{b''}\right)^{\frac{1}{\beta+1}}\right]}{\beta\cdot\left(2^{b''}-2^{b''-1}\right)}$, $\forall k,m\in H,k\neq m$. Taking the second-order derivative of (S-8) on $\sum_{i\in N}y_{k,m}^{i}$, we can demonstrate that

$$\frac{d^2\left(\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}}\right)}{d\left(\sum_{i\in N} y_{k,m}^i\right)}<0, \forall k,m\in H, k\neq m .$$

Therefore, $\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}}$ achieves the maximum value at $\sum_{i\in N} y_{k,m}^i=\frac{(\beta+1)\cdot\left[2^{b''}\cdot\left(2^{b''-1}\right)^{\frac{1}{\beta+1}}-2^{b''-1}\cdot\left(2^{b''}\right)^{\frac{1}{\beta+1}}\right]}{\beta\cdot\left(2^{b''}-2^{b''-1}\right)}$

($\forall k,m\in H, k\neq m$), namely

$$\sup_{\sum_{i\in N} y_{k,m}^i\in\left[\left(2^{b''-1}\right)^{\frac{1}{\beta+1}},\left(2^{b''}\right)^{\frac{1}{\beta+1}}\right]}\left[\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}}\right]=\frac{\beta^\beta}{(1+\beta)^{(1+\beta)}}\cdot\frac{1}{\left(2-2^{\frac{1}{\beta+1}}\right)^\beta\cdot\left(2^{\frac{1}{\beta+1}}-1\right)}-1, \forall k,m\in H, k\neq m .$$

Apparently, it is observed that the maximum value of $\frac{\hat{E}_{k,m}-E_{k,m}}{E_{k,m}}$ is uncorrelated to b''. Given $\beta=4$ and $\beta'=\frac{1}{\beta+1}$, we can obtain that $\varepsilon_{k,m}^3\le\frac{\beta^\beta}{(1+\beta)^{(1+\beta)}}\cdot\frac{1}{\left(2-2^{\beta'}\right)^\beta\cdot\left(2^{\beta'}-1\right)}-1=0.0489<5\%$. □

S.11. Proof of Proposition 10

Proof. First, based on Remark 4, it is obtained that $\left(\sum_{q\in Q}\Lambda_{k,m}^q\cdot u_{k,m}^q\right)^\beta=\sum_{q\in Q}\left(\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q$. Second, considering that $\sum_{q\in Q}u_{k,m}^q=1$ and $u_{k,m}^q\in\{0,1\}$, $\sum_{q\in Q}\left(\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q\in\left\{\left(\Lambda_{k,m}^1\right)^\beta,\left(\Lambda_{k,m}^2\right)^\beta,...,\left(\Lambda_{k,m}^{|Q|}\right)^\beta\right\}$, $\forall k,m\in H, k\neq m$. Furthermore, it is easy to demonstrate that $1\Big/\left[\sum_{q\in Q}\left(\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q\right]\in\left\{\left(1/\Lambda_{k,m}^1\right)^\beta,\left(1/\Lambda_{k,m}^2\right)^\beta,...,\left(1/\Lambda_{k,m}^{|Q|}\right)^\beta\right\}$. Meanwhile, it is worth noting that $\sum_{q\in Q}\left(1/\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q\in\left\{\left(1/\Lambda_{k,m}^1\right)^\beta,\left(1/\Lambda_{k,m}^2\right)^\beta,...,\left(1/\Lambda_{k,m}^{|Q|}\right)^\beta\right\}$, $\forall k,m\in H, k\neq m$. This result suggests that $1\Big/\left[\left(\sum_{q\in Q}\Lambda_{k,m}^q\cdot u_{k,m}^q\right)^\beta\right]$ and $\sum_{q\in Q}\left(1/\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q$ have the same set that they take the corresponding value. Finally, it is proven that $1\Big/\left[\left(\sum_{q\in Q}\Lambda_{k,m}^q\cdot u_{k,m}^q\right)^\beta\right]=\sum_{q\in Q}\left(1/\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q$, $\forall k,m\in H, k\neq m$. This is because that $\sum_{q\in Q}\left(1/\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q=\left(1/\Lambda_{k,m}^{q^*}\right)^\beta$ and $1\Big/\left[\left(\sum_{q\in Q}\Lambda_{k,m}^q\cdot u_{k,m}^q\right)^\beta\right]=\left(1/\Lambda_{k,m}^{q^*}\right)^\beta$ when $u_{k,m}^{q^*}=1$ ($q^*\in Q$). Then, we can give that

$$\begin{aligned}E_{k,m}\Big/\left[\left(\sum_{q\in Q}\Lambda_{k,m}^q\cdot u_{k,m}^q\right)^\beta\right]&=E_{k,m}\cdot\left[1\Big/\left(\sum_{q\in Q}\Lambda_{k,m}^q\cdot u_{k,m}^q\right)^\beta\right]=E_{k,m}\cdot\left[\sum_{q\in Q}\left(1/\Lambda_{k,m}^q\right)^\beta\cdot u_{k,m}^q\right]\\&=\sum_{q\in Q}\left(1/\Lambda_{k,m}^q\right)^\beta\cdot E_{k,m}\cdot u_{k,m}^q, \forall k,m\in H, k\neq m\end{aligned}$$ □

S.12. Proof of Proposition 11

Proof. Note that the term $\sum_{k\in H}\sum_{\substack{m\in H\\k\neq m}} c_{k,m}^{1}\cdot\Xi_{k,m}\cdot\left[\sum_{i\in N} y_{k,m}^{i}+\sum_{q\in Q}\phi\cdot\left(\frac{1}{\Lambda_{k,m}^{q}}\right)^{\beta}\cdot\Omega_{k,m}^{q}\right]$ subjected to additional constraints in the HL-LA-2 reformulation represents the congestion cost for hub links, and it can be equivalently expressed as the term $\sum_{k\in H}\sum_{\substack{m\in H\\k\neq m}} c_{k,m}^{1}\cdot\Xi_{k,m}\cdot\left[\sum_{i\in N} y_{k,m}^{i}+\sum_{q\in Q}\phi\cdot\left(\frac{1}{\Lambda_{k,m}^{q}}\right)^{\beta}\cdot\hat{E}_{k,m}\cdot u_{k,m}^{q}\right]$. In the original model [M1], the congestion cost for hub links can be expressed as the term $\sum_{k\in H}\sum_{\substack{m\in H\\k\neq m}} c_{k,m}^{1}\cdot\Xi_{k,m}\cdot\left[\sum_{i\in N} y_{k,m}^{i}+\sum_{q\in Q}\phi\cdot\left(\frac{1}{\Lambda_{k,m}^{q}}\right)^{\beta}\cdot E_{k,m}\cdot u_{k,m}^{q}\right]$. Since $\hat{E}_{k,m}-E_{k,m}\geq 0$, which has been proved in Proposition 9, $\sum_{q\in Q}\phi\cdot\left(\frac{1}{\Lambda_{k,m}^{q}}\right)^{\beta}\cdot\hat{E}_{k,m}\cdot u_{k,m}^{q}\geq\sum_{q\in Q}\phi\cdot\left(\frac{1}{\Lambda_{k,m}^{q}}\right)^{\beta}\cdot E_{k,m}\cdot u_{k,m}^{q}$. Therefore, the solution of the HL-LA-2 reformulation provides the upper bound of the congestion cost for hub links, compared to the original model [M1]. □

S.13. The illustrative examples of approximating nonlinear functions

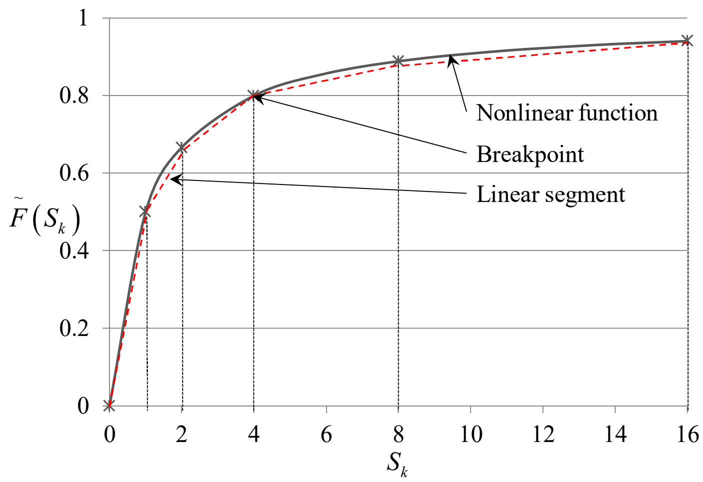

Figure S-1. An illustrative example of approximating nonlinear function $F(S_k)$.

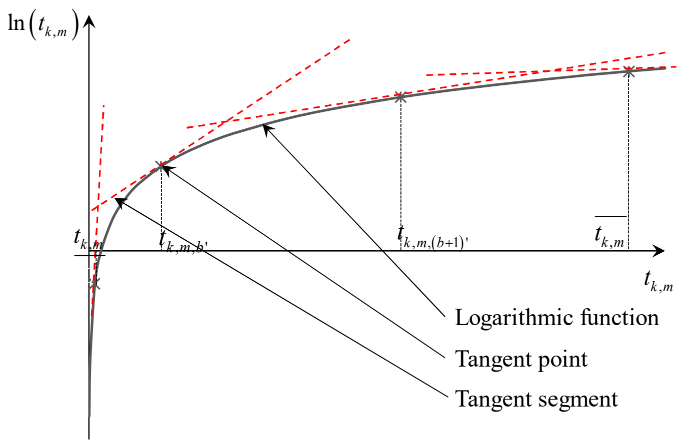

Figure S-2. An illustrative example of showing piecewise tangent approximation for logarithmic function $\ln\left(t_{k,m}\right)$.

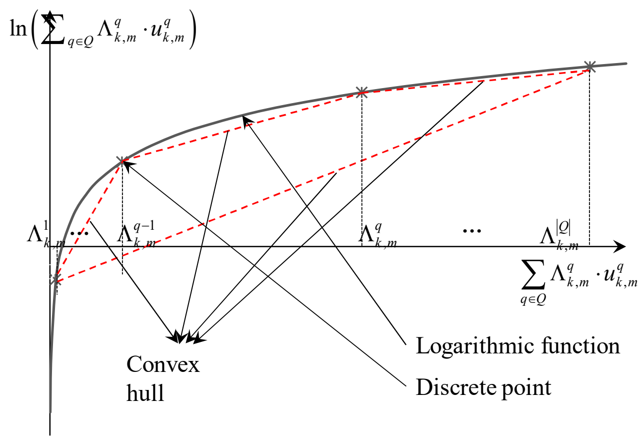

Figure S-3. A description of the convex hull of the set of points in logarithmic function $\ln\left(\sum_{q\in Q}\Lambda_{k,m}^{q}\cdot u_{k,m}^{q}\right)$.

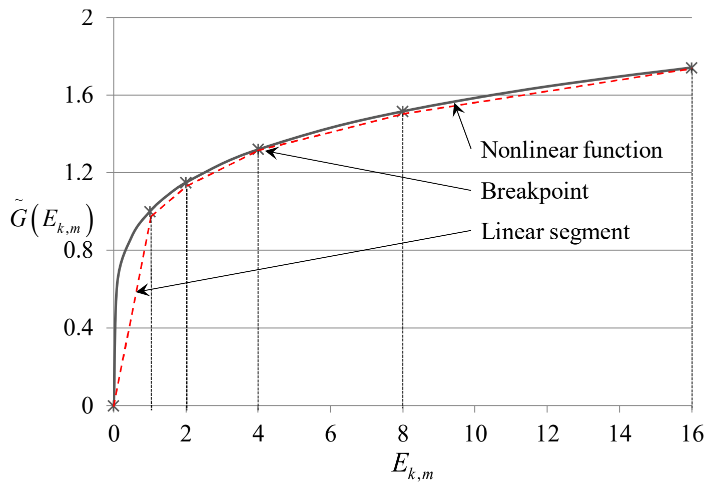

Figure S-4. An illustrative example of approximating nonlinear function $G\left(E_{k,m}\right)$ ($\beta = 4$).

S.14. Parameter descriptions

Table S-2. Parameter descriptions of computational experiments.

Data classes	Parameters	Generation descriptions	
		Original network	Expansion network
CAB	$d_{i,j}$ ($\forall i, j$)	Obtained from OR-Library (2015) directly	Obtained from OR-Library (2015) directly
	$w_{i,j}$ ($\forall i, j$)	$\lceil flow_{i,j}/1000 \rceil$, $flow_{i,j}$ is obtained from OR-Library (2015) directly	$3 \cdot \lceil flow_{i,j}/1000 \rceil$
	f_k ($\forall k$)	$3800 \cdot O_k$	--
	$\lvert N \rvert$	$\{10, 15, 20, 25\}$	$\{10, 15, 20, 25\}$
AP	$d_{i,j}$ ($\forall i, j$)	$dis_{i,j}/100$, $dis_{i,j}$ represents Euclidean distance, and it is calculated using coordinates that are obtained from OR-Library (2015)	$dis_{i,j}/100$
	$w_{i,j}$ ($\forall i, j$)	$\lceil 100 \cdot flow_{i,j} \rceil$, $flow_{i,j}$ is obtained from OR-Library (2015) directly	$3 \cdot \lceil 100 \cdot flow_{i,j} \rceil$
	f_k ($\forall k$)	$1200 \cdot O_k$	--
	$\lvert N \rvert$	$\{10, 15, 20, 25, 40, 50, 75\}$	$\{10, 15, 20, 25, 40, 50, 75\}$
CAB and AP	χ	1	1
	δ	1	1
	α	0.6	$\{0.2, 0.4, 0.6, 0.8\}$
	p (or $\lvert H \rvert$)	--	Given based on the hub location of the original network
	ϕ, β	--	0.15, 4
	$\Xi_{k,m}$ ($\forall k, m$, $k \neq m$)	--	0.5
	c_k^0 ($\forall k$)	--	$\begin{cases} 3000, & low \\ 150000, & high \end{cases}$
	$c_{k,m}^1$ ($\forall k, m$, $k \neq m$)	--	$\begin{cases} 100, & low \\ 4000, & high \end{cases}$

$\Gamma_k^l\ (\forall k)$	--	$\begin{cases} l \cdot \sum_i O_i \cdot x_{i,k}^*, & l = 1,2,3 \\ 5 \cdot \sum_i O_i \cdot x_{i,k}^*, & l = 4 \end{cases}$, $x_{i,k}^*$ is achieved by generating original network
$\Lambda_{k,m}^q\ (\forall k, m, k \neq m)$	--	$\begin{cases} q \cdot \sum_i y_{k,m}^{i*}, & q = 1,2,3 \\ 5 \cdot \sum_i y_{k,m}^{i*}, & q = 4 \end{cases}$, $y_{k,m}^{i*}$ is achieved by generating original network
$\Phi_k^l\ (\forall k)$	--	$400 \cdot \left(\Gamma_k^l - \Gamma_k^1\right),\ l = 1,2,3,4$
$\Psi_{k,m}^q\ (\forall k, m, k \neq m)$	--	$800 \cdot \left(\Lambda_{k,m}^q - \Lambda_{k,m}^1\right),\ q = 1,2,3,4$

Table S-3. Settings of extra parameters for reformulations.

Reformulations	Extra parameters	Settings
H-LA	$\lvert B_k \rvert\ (\forall k)$	$\begin{cases} 12, & O_k \big/ \max_l \left\{\Gamma_k^l\right\} \geq 0.5 \\ 16, & O_k \big/ \max_l \left\{\Gamma_k^l\right\} < 0.5 \end{cases}$
	$\mathring{S}_{k,b}\ (\forall k)$	$\begin{cases} 0, & b = 1 \\ \dfrac{r_k^0}{1-r_k^0} + \dfrac{(b-2)}{4} \cdot \left(1 - \dfrac{r_k^0}{1-r_k^0}\right), & r_k^0 = O_k \big/ \max_l \left\{\Gamma_k^l\right\} < 0.5, b = 2,3,\ldots,6 \\ 2^{b-6}, & r_k^0 = O_k \big/ \max_l \left\{\Gamma_k^l\right\} < 0.5, b = 7,8,\ldots,\lvert B_k \rvert \\ 2^{b-2}, & r_k^0 = O_k \big/ \max_l \left\{\Gamma_k^l\right\} \geq 0.5, b = 2,3,\ldots,\lvert B_k \rvert \end{cases}$
HL-LA-1	$\lvert B' \rvert$	20
HL-LA-2	$\lvert B''_{k,m} \rvert$ ($\forall k, m$, $k \neq m$)	$\left\lceil \log_2 \left(\left(\max_q \left\{\Lambda_{k,m}^q\right\} \right)^{\beta+1} \right) \right\rceil + 2$

References

Alkaabneh, F., Diabat, A., Elhedhli, S., 2019. A Lagrangian heuristic and GRASP for the hub-and-spoke network system with economies-of-scale and congestion. *Transportation Research Part C: Emerging Technologies,* 102, 249-273.

Ashraf, M. H., Chen, Y., Yalcin, M. G., 2022. Minding Braess Paradox amid third-party logistics hub capacity expansion triggered by demand surge. *International Journal of Production Economics,* 248, 108454.

Azizi, N., Vidyarthi, N., Chauhan, S. S., 2018. Modelling and analysis of hub-and-spoke networks under stochastic demand and congestion. *Annals of Operations Research,* 264, 1-40.

Bärmann, A., Martin, A., Schülldorf, H., 2017. A decomposition method for multiperiod railway network expansion—with a case study for Germany. *Transportation Science,* 51(4), 1102-1121.

Bhatt, S. D., Jayaswal, S., Sinha, A., Vidyarthi, N., 2021. Alternate second order conic program reformulations for hub location under stochastic demand and congestion. *Annals of Operations Research,* 304(1), 481-527.

De Camargo, R. S., Demirandajr., G., Ferreira, R. P. M., 2011. A hybrid Outer-Approximation/Benders Decomposition algorithm for the single allocation hub location problem under congestion. *Operations Research Letters,* 39(5), 329-337

De Camargo, R. S., Mirandajr., G., Ferreira, R. P. M., Luna, H. P., 2009. Multiple allocation hub-and-spoke network design under hub congestion. *Computers & Operations Research,* 36(12), 3097-3106.

Elhedhli, S., Hu, F. X., 2005. Hub-and-spoke network design with congestion. *Computers & Operations Research,* 32(6), 1615-1632.

Elhedhli, S., Wu, H., 2010. A Lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion. *INFORMS Journal on Computing,* 22(2), 282-296.

Hu, L., Zhu, J. X., Wang, Y., Lee, L. H., 2018. Joint design of fleet size, hub locations, and hub capacities for third-party logistics networks with road congestion constraints. *Transportation Research Part E: Logistics and Transportation Review,* 118, 568-588.

Karimi-Mamaghan, M., Mohammadi, M., Pirayesh, A., Karimi-Mamaghan, A. M., Irani, H., 2020. Hub-and-spoke network design under congestion: A learning based metaheuristic. *Transportation Research Part E: Logistics and Transportation Review,* 142, 102069.

Kian, R., Kargar, K., 2016. Comparison of the formulations for a hub-and-spoke network design problem under congestion. *Computers & Industrial Engineering,* 101, 504-512.

Liu, H., Wang, D. Z. W., 2015. Global optimization method for network design problem with stochastic user equilibrium. *Transportation Research Part B: Methodological,* 72, 20-39.

Magnanti, T. L., Wong, R. T., 1984. Network design and transportation planning: Models and algorithms. *Transportation Science,* 18(1), 1-55.

Marianov, V., Serra, D., 2003. Location models for airline hubs behaving as M/D/c queues. *Computers & Operations Research,* 30(7), 983-1003.

Marín, Á., Jaramillo, P., 2008. Urban rapid transit network capacity expansion. *European Journal of Operational Research,* 191(1), 45-60.

Miranda, G., Luna, H. P., Decamargo, R. S., Pinto, L. R., 2011. Tree network design avoiding congestion. *Applied Mathematical Modelling,* 35(9), 4175-4188.

Mohammadi, M., Jula, P., Tavakkoli-Moghaddam, R., 2019. Reliable single-allocation hub location problem with disruptions. *Transportation Research Part E: Logistics and Transportation Review,* 123, 90-120.

Najy, W., Diabat, A., 2020. Benders decomposition for multiple-allocation hub-and-spoke network design with economies of scale and node congestion. *Transportation Research Part B: Methodological,* 133, 62-84.

OR-Library, 2015. Hub Location. http://people.brunel.ac.uk/~mastjjb/jeb/orlib/phubinfo.html.

Paraskevopoulos, D. C., Gürel, S., Bektas, T., 2016. The congested multicommodity network design problem. *Transportation Research Part E: Logistics and Transportation Review,* 85, 166-187.

Rahimi, Y., Tavakkoli-Moghaddam, R., Mohammadi, M., Sadeghi, M., 2016. Multi-objective hub network design under uncertainty considering congestion: An M/M/c/K queue system. *Applied Mathematical Modelling,* 40(5), 4179-4198.

Şafak, Ö., Çavuş, Ö., Aktürk, M. S., 2022. A two-stage decision dependent stochastic approach for airline flight network expansion. *Transportation Research Part B: Methodological,* 158, 78-101.

Sun, Y., Schonfeld, P., 2015. Stochastic capacity expansion models for airport facilities. *Transportation Research Part B: Methodological,* 80, 1-18.

Taghavi, M., Huang, K., 2016. A multi-stage stochastic programming approach for network capacity expansion with multiple sources of capacity. *Naval Research Logistics,* 63(8), 600-614.

Zhalechian, M., Tavakkoli-Moghaddam, R., Rahimi, Y., 2017. A self-adaptive evolutionary algorithm for a fuzzy multi-objective hub location problem: An integration of responsiveness and social responsibility. *Engineering Applications of Artificial Intelligence,* 62, 1-16.

Zhang, X., Liu, X., 2022. A two-stage robust model for express service network design with surging demand. *European Journal of Operational Research,* 299(1), 154-167.

Zhao, S., Haskell, W. B., Cardin, M.-A., 2018. Decision rule-based method for flexible multi-facility capacity expansion problem. *IISE Transactions,* 50(7), 553-569.