Economies of Scale in Hub & Spoke Network Design Models: We Have It All Wrong

Alf Kimms

Lehrstuhl für ABWL, insbes. Produktion und Logistik
Technische Universität Freiberg

1 Traditional Models for Hub & Spoke Network Design

The hub & spoke network design problem is a strategic logistics planning problem with applications for airlines, telecommunication companies, computer networks, postal services, and trucking companies, for example. Basically, the problem in all these applications is that for a given set $V = \{1, \ldots, n\}$ of nodes (airports, computers, post offices, depots, ...) goods must be transported between possibly every pair of nodes. Direct connections between every pair of nodes would result in $n(n-1)$ linkages which is impractically high and economically non–profitable. Consider, for instance, an airline that serves several airports worldwide. Offering non–stop flights between every pair of airports would require a huge amount of planes and crews and many empty seats on board could be observed for many connections. In such settings, it turns out to be reasonable to install one or more so–called hub locations where direct links are then available to hub nodes as indicated in figure 1 where nodes 3, 6, and 9 are assumed to be hubs. Transporting goods from, say, node 1 to node 11, can then be done via hubs 3 and 6.

Roughly speaking, the network design problem at hand can be couched as follows: Given a graph with node set V and edge set $E = V \times V$, select one or more nodes from V to become hub nodes and select some edges from E to become transportation links. For each pair of nodes $(i, j) \in V \times V$ we have a quantity $q_{ij} \in \boldsymbol{R}_{\geq 0}$ that is to be transported from node i to node j. Established models assume that the unit cost of transportation using an edge e is $c_e \in \boldsymbol{R}_{\geq 0}$ and that, if e connects two hub nodes, a discount can be gained such that the unit cost of transportation using edge e is $\alpha c_e \in \boldsymbol{R}_{\geq 0}$ with $0 \leq \alpha \leq 1$. We will question these cost assumptions in section 3 and discuss alternatives. Note that the costs may be asymmetric, i.e. $c_{ij} = c_{ji}$ may not be true.

For the hub & spoke network structure one can wish to have specific characteristics that define the design problem to be solved. Some usual and

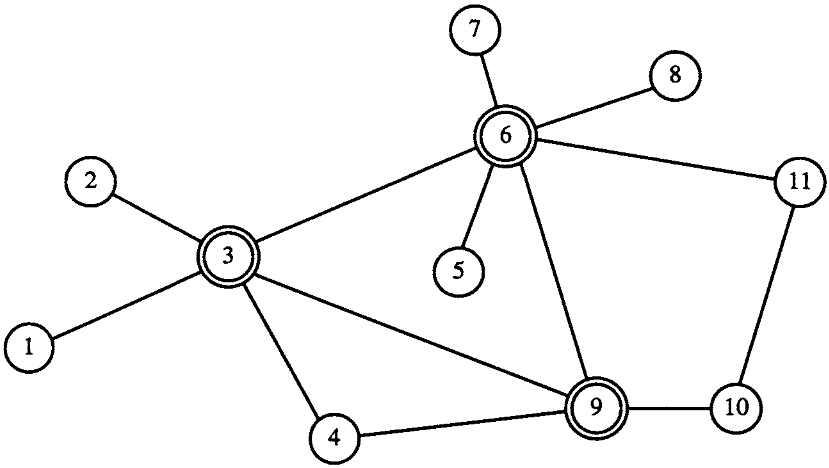

Figure 1: An Illustration of a Hub & Spoke System

basic features are the following:

- What determines the number of hubs?
  - Hub location problems with fixed hub costs: Installing a hub at node h incurs a fixed cost $f_h \in \boldsymbol{R}_{\geq 0}$. The number of hubs is a result of the planning process.
  - p–Hub median problems: The number p of hubs is predefined. Fixed hub costs are usually assumed to be the same for all nodes so that they can be ignored for the purpose of optimization.
- How are the non–hub nodes connected?
  - Single allocation: Each non–hub node is allocated to a unique hub. In a single allocation network, node 4 (figure 1) would not be allowed to have a direct link to two hubs.
  - Multiple allocation: Non–hub nodes (like node 4 in figure 1) may be connected to several hubs.
  - Direct services: Non–hub nodes may have a direct connection like nodes 10 and 11 in figure 1.

Throughout this paper, we assume that the set of hubs is fully meshed, i.e. the subgraph induced by the hub nodes is complete, and that the objective for designing the network is to minimize the sum of relevant costs.

Mathematically, typical hub & spoke network design problems can be stated as follows (see Campbell, 1994b, for several models in this area):

For the single allocation case without direct services, define a binary decision variable y_{ih} which is equal to one, if and only if node h is the hub that node i is allocated to, and zero, otherwise. A value $y_{hh} = 1$ indicates that node h is a hub. Additionally, a real–valued variable x_{ijhk} is used to model the fraction of flow that is routed from node i to node j via hubs h and k in that order.

$$\text{Min } F_1(\mathbf{x},\mathbf{y}) = \sum_{h=1}^{n} f_h y_{hh} + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\sum_{k=1}^{n}(c_{ih} + \alpha c_{hk} + c_{kj}) q_{ij} x_{ijhk} \tag{1}$$

s.t.

$$y_{ih} \le y_{hh} \qquad i,h = 1,\ldots,n \quad i \ne h \tag{2}$$

$$\sum_{h=1}^{n} y_{ih} = 1 \qquad i = 1,\ldots,n \tag{3}$$

$$\sum_{h=1}^{n}\sum_{k=1}^{n} x_{ijhk} = 1 \qquad i,j = 1,\ldots,n \tag{4}$$

$$\sum_{k=1}^{n} x_{ijhk} \le y_{ih} \qquad i,j,h = 1,\ldots,n \tag{5}$$

$$\sum_{h=1}^{n} x_{ijhk} \le y_{jk} \qquad i,j,k = 1,\ldots,n \tag{6}$$

$$x_{ijhk} \ge 0 \qquad i,j,h,k = 1,\ldots,n \tag{7}$$

$$y_{ih} \in \{0,1\} \qquad i,h = 1,\ldots,n \tag{8}$$

The objective (1) minimizes the sum of fixed costs for installing the hubs plus transportation costs. Due to (2), a non–hub node i can only be linked to a hub h, if node h is indeed a hub. (3) makes sure that every node is linked to exactly one hub. Because of (4), all quantities are shipped. And because of (5) and (6) a path from node i to node j via nodes h and k, respectively, can only be used, if h and k are hub nodes. It is noteworthy to say that it is a common (e.g., if it is assumed that costs fulfill $c_{ij} + c_{jk} \ge c_{ik}$) that while moving from a node i to a node j at most two other (hub) nodes are passed in-between. (7) and (8) define the decision variables. Note that there exists an optimal solution with x being integral.

By adding the constraint

$$\sum_{h=1}^{n} y_{hh} = p$$

and assuming the special case $f_h = 0$ for all h, we would get a model for the p–hub median problem.

For representing the multiple allocation case, a one–index binary variable y_h is sufficient. The variable y_h is one, if node h is a hub, and zero, otherwise.

$$\begin{aligned} \text{Min } F_2(\mathbf{x}, \mathbf{y}) = & \sum_{h=1}^{n} f_h y_h \\ & + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{h=1}^{n} \sum_{k=1}^{n} (c_{ih} + \alpha c_{hk} + c_{kj}) q_{ij} x_{ijhk} \end{aligned} \tag{9}$$

s.t.

$$\sum_{h=1}^{n} \sum_{k=1}^{n} x_{ijhk} = 1 \qquad i, j = 1, \ldots, n \tag{10}$$

$$\sum_{k=1}^{n} x_{ijhk} + \sum_{k=1, k \neq h}^{n} x_{ijkh} \leq y_h \qquad i, j, h = 1, \ldots, n \tag{11}$$

$$x_{ijhk} \geq 0 \qquad i, j, h, k = 1, \ldots, n \tag{12}$$

$$y_h \in \{0, 1\} \qquad h = 1, \ldots, n \tag{13}$$

It should be sufficient to explain (11). It makes sure that a flow from node i to node j can pass a node h only if h is a hub node. In this case either i or j (or both) would be linked to h.

By adding the constraint

$$\sum_{h=1}^{n} y_h = p \tag{14}$$

and assuming the special case $f_h = 0$ for all h, we would get a model for the p–hub median problem.

If direct services are allowed, an additional real–valued variable x'_{ij} for the fraction of flow from node i to node j leads to the following model:

$$\begin{aligned} \text{Min } F_3(\mathbf{x}', \mathbf{x}, \mathbf{y}) = & \sum_{h=1}^{n} f_h y_h \\ & + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{h=1}^{n} \sum_{k=1}^{n} (c_{ih} + \alpha c_{hk} + c_{kj}) q_{ij} x_{ijhk} \\ & + \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} q_{ij} x'_{ij} \end{aligned} \tag{15}$$

s.t.

$$x'_{ij} + \sum_{h=1}^{n}\sum_{k=1}^{n} x_{ijhk} = 1 \qquad i,j = 1,\ldots,n \qquad (16)$$

$$\sum_{k=1}^{n} x_{ijhk} + \sum_{k=1,k\neq h}^{n} x_{ijkh} \leq y_h \qquad i,j,h = 1,\ldots,n \qquad (17)$$

$$x'_{ij} \geq 0 \qquad i,j = 1,\ldots,n \qquad (18)$$

$$x_{ijhk} \geq 0 \qquad i,j,h,k = 1,\ldots,n \qquad (19)$$

$$y_h \in \{0,1\} \qquad h = 1,\ldots,n \qquad (20)$$

Here, the objective function (15) takes into account the costs for direct services as well. Due to (16), all quantities must be shipped where some of them may be shipped directly.

And again, by adding the constraint (14) and assuming the special case $f_h = 0$ for all h, we would get a model for the p–hub median problem.

The seminal work of O'Kelly (1986,1987) has launched a series of follow–up publications. To get an overview of the work that has been done, we refer to Bryan and O'Kelly (1999), Campbell (1994a), Campbell et al. (2002), Cánovas et al. (2004), Klincewicz (1998), Mayer (2001), O'Kelly and Miller (1994) and Wagner (2005). Note that the models presented here are not intended to define the most efficient model formulations for the problems described. The models should just describe the problems formally and should show what almost — see section 2 — all authors assume. Publications where efficient model formulations are discussed are, e.g., Cánovas et al. (2005), Ernst et al. (2002), Ernst and Krishnamoorthy (1999), Kara and Tansel (2000), Skorin–Kapov et al. (1996), and Wagner (2003) just to mention a few.

In what follows, we will concentrate on the modeling of the transportation costs in the presence of economies of scale. In section 2 we will review the (surprisingly few) work that has been devoted to economies of scale in the context of hub & spoke network design. Eventually, in section 3 we will present and discuss our point of view on this matter. Section 4 will illustrate the new models from section 3 by means of a few numerical examples. A short conclusion will summarize the work in the final section.

2 Existing Approaches to Represent Economies of Scale

Surprisingly enough, although the economies of scale phenomenon is one of the main motivations for installing hub & spoke systems, the way costs are modeled has not really been questioned by many authors in this area.

At least some authors have noted that applying a discount factor α to the costs on arcs between hubs while disregarding the flow on these arcs contradicts the motivation of this discount factor. Consequently, Podnar et al. (2002) introduced flow tresholds that must be reached in order to gain the discount α. Campbell et al. (2004a, 2004b) state that "...the basic assumption in hub median models that flow costs are discounted on hub arcs to reflect high volumes leads to a possible mismatch between the abstracted model and the underlying motivations of the model". So, they consider models where so–called hub arcs, i.e. arcs that link two hubs and on which costs being discounted by a factor α, are to be selected explicitly which means that the cost for a flow between two hubs may or may not be discounted by a factor α depending on the selection. The total number of such hub arcs to be chosen is prespecified in their models, because otherwise every arc that links two hubs would be selected as a hub arc.

Three notable exceptions that address economies of scale are Bryan (1998), O'Kelly and Bryan (1998) and Klincewicz (2002). They argue — and we agree — that "By simplifying interhub travel costs and assuming that these costs are independent of flows, the current model not only miscalculates the total network cost, but also erroneously selects optimal hub locations and allocations".

As a consequence, O'Kelly and Bryan (1998) suggest to replace the interhub cost expression

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha c_{hk} q_{ij} x_{ijhk}$$

for hubs h and k in the objective function(s) by a concave function

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1-\theta\left(\frac{\sum_{i'=1}^{n}\sum_{j'=1}^{n} q_{i'j'}x_{i'j'hk}}{\sum_{i'=1}^{n}\sum_{j'=1}^{n} q_{i'j'}}\right)^{\beta}\right) c_{hk} q_{ij} x_{ijhk}$$

where the parameters $1 \geq \theta > 0$ and $\beta > 0$ are to be specified in advance. This function is monotonically increasing with the flow across the link between node h and node k. It can reasonably be used to model economies of scale: the marginal cost per unit is decreasing and the average cost per unit is decreasing as well. The latter property is what is called economies of scale.

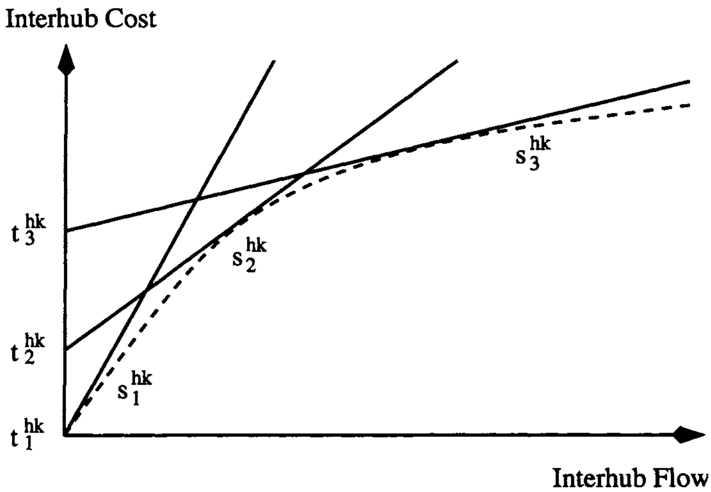

Figure 2: Interhub Cost Structure due to O'Kelly and Bryan (1998)

As illustrated in figure 2, the non–linear function

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1-\theta\left(\frac{\sum\limits_{i'=1}^{n}\sum\limits_{j'=1}^{n} q_{i'j'}x_{i'j'hk}}{\sum\limits_{i'=1}^{n}\sum\limits_{j'=1}^{n} q_{i'j'}}\right)^{\beta}\right) q_{ij}x_{ijhk}$$

can be approximated by a piecewise linear function in such a way that the lower envelope of this piecewise linear function approximates the function from above. This is exactly what Bryan (1998), O'Kelly and Bryan (1998) and Klincewicz (2002) do. Each piece r of these o linear pieces can be specified by two parameters: the intercept $t_r^{hk} \geq 0$ and the slope $s_r^{hk} > 0$. This allows to provide a linear mixed–integer formulation where the real–valued decision variable f_{rhk} is the total flow between hubs h and k to which the linear piece r is applied and the binary variable z_{rhk} indicates whether or not the piece r is applied at all. For the sake of brevity, we will provide here a model for the multiple allocation p–hub median problem without direct services. It should be easy for the reader to write down other problem variants in a similar fashion.

$$\text{Min } F_4(\mathbf{f},\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\sum_{k=1}^{n}(c_{ih}+c_{kj})q_{ij}x_{ijhk} + \sum_{r=1}^{o}\sum_{h=1}^{n}\sum_{k=1}^{n} c_{hk}(t_r^{hk} z_{rhk} + s_r^{hk} f_{rhk}) \tag{21}$$

s.t.

$$\sum_{r=1}^{o} f_{rhk} = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x_{ijhk} \qquad h,k=1,\ldots,n \tag{22}$$

$$f_{rhk} \le \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}z_{rhk} \qquad r=1,\ldots,o;\ h,k=1,\ldots,n \tag{23}$$

$$\sum_{h=1}^{n}\sum_{k=1}^{n} x_{ijhk} = 1 \qquad i,j=1,\ldots,n \tag{24}$$

$$\sum_{k=1}^{n} x_{ijhk} + \sum_{k=1,k\neq h}^{n} x_{ijkh} \le y_h \qquad i,j,h=1,\ldots,n \tag{25}$$

$$\sum_{h=1}^{n} y_h = p \tag{26}$$

$$f_{rhk} \ge 0 \qquad r=1,\ldots,o;\ h,k=1,\ldots,n \tag{27}$$

$$x_{ijhk} \ge 0 \qquad i,j,h,k=1,\ldots,n \tag{28}$$

$$y_h \in \{0,1\} \qquad h=1,\ldots,n \tag{29}$$

$$z_{rhk} \in \{0,1\} \qquad r=1,\ldots,o;\ h,k=1,\ldots,n \tag{30}$$

The objective function (21) measures interhub traffic by means of the cost approximation described above. The total flow between two hubs is represented by f as well as by x. (22) links these two decision variables in the correct manner. (23) guarantees that the flag to indicate whether or not a certain piece of the piecewise linear approximation is used is properly set. Note that due to the minimization objective, the lower envelope of the piecewise linear approximation is indeed used.

Bryan (1998) extends this model for economies of scale on all connections, not only the interhub ones. The technique with which this is done is the same piecewise linear approximation that is described above.

3 Alternative Model Formulations

From a cost accounting point of view, the model for economies of scale described in the previous section is hard to motivate. If economies of scale are due to quantity discounts, then we face a piecewise linear, monotone, quasiconcave cost function. We refer to Stadtler (2004) for a discussion of modeling quantity discounts. In a hub & spoke network design setting, however, economies of scale due to quantity discounts may appear only if the transportation is done by a third party. We will not discuss this case here.

Consequently we will try to derive an alternative model and reveal other sources for economies of scale. For doing this, let us consider two nodes i and h. These nodes can be hubs or not. And indeed we will see that economies of scale occur not only on interhub links but on all links. In general, if we provide a service from node i to node h, i.e. we have a positive flow from i to h, we face a (flow independent) fixed cost $c^f_{ih} \in \boldsymbol{R}_{\geq 0}$ and a unit (handling) cost $c^v_{ih} \in \boldsymbol{R}_{\geq 0}$. If f_{ih} is the flow between i and h then

$$\begin{array}{rl} c^f_{ih} + c^v_{ih} f_{ih}, & \text{if } f_{ih} > 0 \\ 0, & \text{otherwise} \end{array}$$

is the total cost for the service. Imagine, for example, the airline situation for passenger transport. If there is a flight from airport i to airport h, a (large) fixed cost c^f_{ih} for using the plane, employing the cabin crew etc. is incurred. In addition to that, a (relatively small) unit handling cost c^v_{ih} per passenger is incurred (mainly for serving an additional meal and a few drinks on board). Note that this simple cost model describes economies of scale already: The unit total cost is

$$\begin{array}{rl} \dfrac{c^f_{ih}}{f_{ih}} + c^v_{ih} & \text{if } f_{ih} > 0 \\ 0, & \text{otherwise} \end{array}$$

which means that the unit total cost decreases if the flow increases. Note that such a model, a model with fixed costs on active links, has already been provided by Campbell (1994b) who did not mention the economies of scale aspect. Another work with fixed costs on active arcs is the one by Garfinkel et al. (1996) which is confined to a very special situation (no handling costs, two hubs), but again, economies of scale have not been discussed. Gavish (1992) considers fixed costs on links for a computer network design problem.

It is important to note that flow from a node i to a node j may be routed through several — possibly more than two — other hub nodes to save fixed costs. We will just provide a model formulation for the multiple allocation p–hub median problem with direct services where the decision variable f_{ij} denotes the direct flow between nodes i and j and the binary decision variable z_{ij} indicates whether or not there is a positive flow between nodes i and j. The variable x_{ijhk} denotes the fraction of flow that originates from i and is sent to node j using the link between h and k:

$$\text{Min } F_5(\mathbf{f},\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i=1}^{n}\sum_{j=1}^{n}(c_{ij}^{f} z_{ij} + c_{ij}^{v} f_{ij}) \tag{31}$$

s.t.

$$\sum_{k=1}^{n} x_{ijik} = 1 \qquad i,j = 1,\dots,n \tag{32}$$

$$\sum_{k=1,k\neq h}^{n} x_{ijkh} = \sum_{k=1,k\neq h}^{n} x_{ijhk} \qquad i,j,h = 1,\dots,n \quad h \neq i,j \tag{33}$$

$$\sum_{h=1}^{n} x_{ijhi} = 0 \qquad i,j = 1,\dots,n \quad i \neq j \tag{34}$$

$$\sum_{k=1}^{n} x_{ijhk} \le y_h \qquad i,j,h = 1,\dots,n \quad h \neq i,j \tag{35}$$

$$\sum_{k=1}^{n} x_{ijkh} \le y_h \qquad i,j,h = 1,\dots,n \quad h \neq i,j \tag{36}$$

$$\sum_{h=1}^{n} y_h = p \tag{37}$$

$$f_{ij} = \sum_{h=1}^{n}\sum_{k=1}^{n} q_{hk} x_{hkij} \qquad i,j = 1,\dots,n \tag{38}$$

$$f_{ij} \le \sum_{h=1}^{n}\sum_{k=1}^{n} q_{hk} z_{ij} \qquad i,j = 1,\dots,n \tag{39}$$

$$f_{ij} \ge 0 \qquad i,j = 1,\dots,n \tag{40}$$

$$x_{ijhk} \ge 0 \qquad i,j,h,k = 1,\dots,n \tag{41}$$

$$y_h \in \{0,1\} \qquad h = 1,\dots,n \tag{42}$$

$$z_{ij} \in \{0,1\} \qquad i,j = 1,\dots,n \tag{43}$$

The objective function (31) is defined as described above. (32) and (33) are flow constraints. For each pair of nodes i and j the complete quantity must leave node i and reach node j. Each flow that reaches a node in–between must leave that node completely. (34) prevents short cycles. (35) and (36) make sure that only hubs nodes can be used in–between to go from a node i to a node j. (38) calculates the total flow over a link from node i to node j. Of course, because of this equality constraint, the variable f could be completely eliminated from the model by substitution. (39) makes sure that a positive flow on a link between i and j is correctly indicated. It is noteworthy to mention that there always exists an optimum solution with all x–variables being integer valued.

Note that if we replace $\sum_h \sum_k q_{hk}$ by some number L_{ij} in (39), we could easily model a constrained capacity on that link. Then, an optimum solution with integral x–values may not exist.

Should direct services not be allowed, the constraint

$$x_{ijij} \leq y_i + y_j \qquad i, j = 1, \ldots, n$$

would forbid such service.

The fixed cost c^f_{ij} is charged if the directed link between node i and node j is active. Using that link in opposite direction incurs an additional fixed cost c^f_{ji}. If we consider two nodes i and j where $i \leq j$ and if we would like to model a situation where a fixed cost c^f_{ij} is charged no matter in what direction the link between these two nodes is used, then we could simply add the constraint

$$z_{ij} = z_{ji} \qquad \begin{array}{l} i = 1, \ldots, n-1 \\ j = i+1, \ldots, n \end{array}$$

and set $c^f_{ji} = 0$ for all $i < j$. Of course, we could eliminate almost half of the z–variables instead to keep only those variables z_{ij} with $i \leq j$ which requires a slight modification of the model formulation especially in (31) and (39).

There may be situations where the just presented model is still too simple. Imagine, for instance, a trucking company. If they provide a transportation service from town i to town h, they may have a fixed cost c^f_{ij}, they may have a unit handling cost c^v_{ij}, but they may also figure out that the number of trucks to serve that link depends on the flow, because each truck has a limited capacity Q, where each truck incurs a fixed cost $c^t_{ij} \in \boldsymbol{R}_{\geq 0}$ as well. If f_{ih} is the flow on that connection, the total cost adds to

$$\begin{array}{rl} c^f_{ih} + c^t_{ih} \left\lceil \dfrac{f_{ih}}{Q} \right\rceil + c^v_{ih} f_{ih}, & \text{if } f_{ih} > 0 \\ 0, & \text{otherwise} \end{array}$$

and figure 3 illustrates the situation. Similar situations have been mentioned by Ebner (1997) and Wlček (1998).

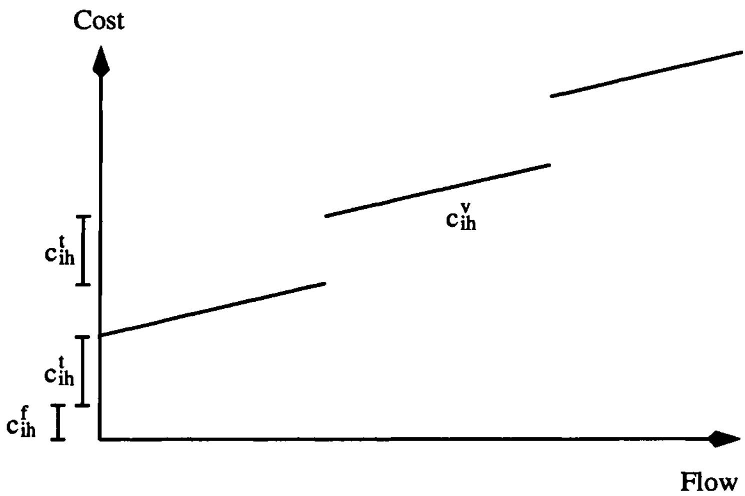

Figure 3: Cost Structure with Fixed Costs

The corresponding model formulation with an integer–valued decision variable t_{ih} for the number of vehicles needed looks as follows:

$$\text{Min } F_6(\mathbf{f},\mathbf{t},\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i=1}^{n}\sum_{j=1}^{n}(c_{ij}^{f}z_{ij} + c_{ij}^{t}t_{ij} + c_{ij}^{v}f_{ij}) \tag{44}$$

s.t.

$$\sum_{k=1}^{n} x_{ijik} = 1 \qquad i,j = 1,\ldots,n \tag{45}$$

$$\sum_{k=1,k\neq h}^{n} x_{ijkh} = \sum_{k=1,k\neq h}^{n} x_{ijhk} \qquad i,j,h = 1,\ldots,n \quad h \neq i,j \tag{46}$$

$$\sum_{h=1}^{n} x_{ijhi} = 0 \qquad i,j = 1,\ldots,n \quad i \neq j \tag{47}$$

$$\sum_{k=1}^{n} x_{ijhk} \leq y_h \qquad i,j,h = 1,\ldots,n \quad h \neq i,j \tag{48}$$

$$\sum_{k=1}^{n} x_{ijkh} \leq y_h \qquad i,j,h = 1,\ldots,n \quad h \neq i,j \tag{49}$$

$$\sum_{h=1}^{n} y_h = p \tag{50}$$

$$f_{ij} = \sum_{h=1}^{n}\sum_{k=1}^{n} q_{hk} x_{hkij} \qquad i,j = 1,\ldots,n \tag{51}$$

$$f_{ij} \leq \sum_{h=1}^{n}\sum_{k=1}^{n} q_{hk} z_{ij} \qquad i,j = 1,\ldots,n \tag{52}$$

$$f_{ij} \leq Q \cdot t_{ij} \qquad i,j = 1,\ldots,n \tag{53}$$

$$f_{ij} \geq 0 \qquad i,j = 1,\ldots,n \tag{54}$$

$$t_{ij} \in N_0 \qquad i,j = 1,\ldots,n \tag{55}$$

$$x_{ijhk} \geq 0 \qquad i,j,h,k = 1,\ldots,n \tag{56}$$

$$y_h \in \{0,1\} \qquad h = 1,\ldots,n \tag{57}$$

$$z_{ij} \in \{0,1\} \qquad i,j = 1,\ldots,n \tag{58}$$

The objective function (44) has already been defined above. The main point that is new here is restriction (53). Because of this constraint, the capacity offered on a certain link must be sufficient to transport the calculated flow on that link. Because of the minimization objective, the offered capacity will be as small as possible. It should be remarked that an optimum solution with all x–variables being integral may not exist.

Additional constraints of the form

$$t_{ij} \leq T_{ij} \qquad i,j = 1,\ldots,n$$

could be used to model a restricted number of T_{ij} vehicles on a particular link. In a similar fashion

$$\sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij} \leq T$$

could constrain the total number of vehicles to a limit T in the whole network.

This model can be adapted to even more general situations. Imagine situations where the type of transportation vehicle can also be decided. For instance, if one can use large trucks instead of small ones, trains instead of trucks, planes instead of trucks or trains and so on. Or imagine the telecommunication industry where the bandwidth of connected nodes can be decided upon. Let us assume that each link has a set of M so–called modes in which that link can be established. For a particular mode m, let Q^m be the capacity of a vehicle that corresponds to that mode and let c_{ih}^{fm},

c_{ih}^{tm} and c_{ih}^{vm} be the cost coefficients for this mode for a given link (i, h). Note that several modes may cause economies of scale if $c_{ih}^{tm}/Q^m > c_{ih}^{tm'}/Q^{m'}$ for modes m and m'. But this is not a sufficient criterion and depends especially on the values of c_{ih}^{fm} and c_{ih}^{vm}. The adapted decision variables z_{ih}^m, t_{ih}^m and f_{ih}^m have a straightforward interpretation:

$$\text{Min } F_7(\mathbf{f}, \mathbf{t}, \mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(c_{ij}^f z_{ij} + \left(\sum_{m=1}^{M} (c_{ij}^{fm} z_{ij}^m + c_{ij}^{tm} t_{ij}^m + c_{ij}^{vm} f_{ij}^m \right) \right) \tag{59}$$

s.t.

$$\sum_{k=1}^{n} x_{ijk} = 1 \qquad i, j = 1, \ldots, n \tag{60}$$

$$\sum_{k=1, k \neq h}^{n} x_{ijkh} = \sum_{k=1, k \neq h}^{n} x_{ijhk} \qquad i, j, h = 1, \ldots, n \quad h \neq i, j \tag{61}$$

$$\sum_{h=1}^{n} x_{ijhi} = 0 \qquad i, j = 1, \ldots, n \quad i \neq j \tag{62}$$

$$\sum_{k=1}^{n} x_{ijhk} \leq y_h \qquad i, j, h = 1, \ldots, n \quad h \neq i, j \tag{63}$$

$$\sum_{k=1}^{n} x_{ijkh} \leq y_h \qquad i, j, h = 1, \ldots, n \quad h \neq i, j \tag{64}$$

$$\sum_{h=1}^{n} y_h = p \tag{65}$$

$$\sum_{m=1}^{M} f_{ij}^m = \sum_{h=1}^{n} \sum_{k=1}^{n} q_{hk} x_{hkij} \qquad i, j = 1, \ldots, n \tag{66}$$

$$\sum_{m=1}^{M} f_{ij}^m \leq \sum_{h=1}^{n} \sum_{k=1}^{n} q_{hk} z_{ij} \qquad i, j = 1, \ldots, n \tag{67}$$

$$f_{ij}^m \leq \sum_{h=1}^{n} \sum_{k=1}^{n} q_{hk} z_{ij}^m \qquad m = 1, \ldots, M \quad i, j = 1, \ldots, n \tag{68}$$

$$f_{ij}^m \leq Q^m \cdot t_{ij}^m \qquad m = 1, \ldots, M \quad i, j = 1, \ldots, n \tag{69}$$

$$f_{ij}^m \geq 0 \qquad m = 1, \ldots, M \quad i, j = 1, \ldots, n \tag{70}$$

$$t_{ij}^m \in N_0 \qquad \begin{array}{l} m = 1, \ldots, M \\ i, j = 1, \ldots, n \end{array} \tag{71}$$

$$x_{ijhk} \geq 0 \qquad i, j, h, k = 1, \ldots, n \tag{72}$$

$$y_h \in \{0, 1\} \qquad h = 1, \ldots, n \tag{73}$$

$$z_{ij} \in \{0, 1\} \qquad i, j = 1, \ldots, n \tag{74}$$

$$z_{ij}^m \in \{0, 1\} \qquad \begin{array}{l} m = 1, \ldots, M \\ i, j = 1, \ldots, n \end{array} \tag{75}$$

Again, the objective function (59) has already been explained above. The interesting aspect here is (66). One should note that due to this formulation, the mode for a link between i and h need not be unique. Indeed, one could use a mix of different modes. If a certain mode m should be not available on a link from node i to node j, then one can simply set $z_{ij}^m = 0$ so that due to (68) this mode will not be active on that link. Eventually, one should note that an optimum solution with all x–variables being integer–valued may not exist.

An extended model where at most one mode per link can be chosen is easy to formulate by adding:

$$\sum_{m=1}^{M} z_{ij}^m \leq 1 \qquad i, j = 1, \ldots, n$$

4 Numerical Examples

To illustrate the above models, we provide here the optimum results for three random examples computed with the commercial mathematical programming software package AMPL/CPLEX. Every example consists of $n = 7$ nodes. The number of hubs is always $p = 3$. The quantities q_{ij} to be transported in all examples are given by table 1, and the fixed cost coefficients c_{ij}^f are provided in table 2.

Example 1 corresponds to the model formulation (31)–(43). In addition to the parameters already introduced, the variable cost coefficients c_{ij}^v as defined in table 3 are used.

Figure 4 shows an optimum solution for example 1 (what is shown are the positive flows f_{ij}). The optimum objective function value is 2403. It is remarkable to note that in this solution some quantities must indeed be transported via more than two hubs. The five units to be shipped from node $i = 5$ to node $j = 2$, for example, flow through hubs 1, 3, and 4 in that order.

Table 1: Transportation Quantities q_{ij} for the Examples

q_{ij}	1	2	3	4	5	6	7
1	8	3	7	7	4	3	4
2	4	1	5	5	6	5	7
3	5	4	2	4	9	4	8
4	8	8	1	8	8	7	5
5	5	5	4	2	5	8	1
6	8	2	8	1	6	1	2
7	7	3	9	3	2	2	3

Table 2: Fixed Cost Coefficients c^f_{ij} for the Examples

c^f_{ij}	1	2	3	4	5	6	7
1	60	176	50	175	130	134	70
2	80	184	40	120	190	60	64
3	90	132	46	149	168	57	59
4	200	154	47	138	147	180	80
5	150	156	48	168	123	147	190
6	140	189	168	147	150	153	130
7	160	147	137	159	130	164	180

Table 3: Variable Cost Coefficients c^v_{ij} for the Examples

c^v_{ij}	1	2	3	4	5	6	7
1	7	2	3	8	1	2	4
2	2	3	3	6	5	8	2
3	7	4	5	1	3	9	9
4	2	1	3	5	7	4	4
5	2	5	9	5	8	4	2
6	7	5	5	3	1	1	2
7	3	3	1	5	9	8	7

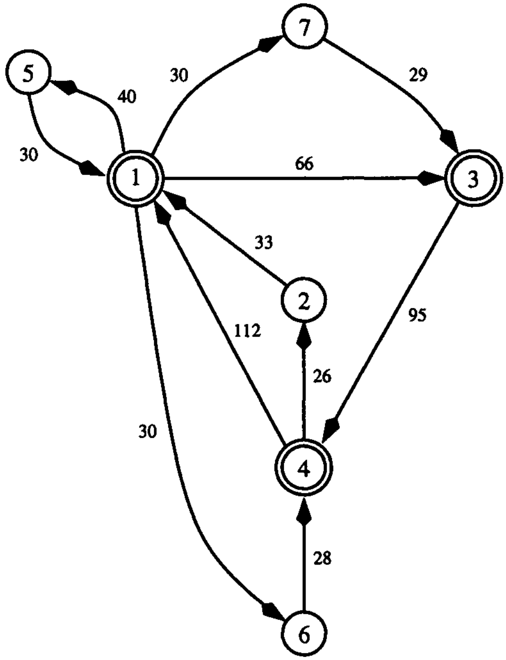

Figure 4: An Optimum Solution for Example 1

Example 2 corresponds to the model formulation (44)–(58). We use $Q = 10$. The cost coefficients c_{ij}^{f} and c_{ij}^{v} are the same as in example 1. The cost coefficients c_{ij}^{t} can be found in table 4.

Table 4: Cost Coefficients c_{ij}^{t} for the Examples

c_{ij}^{t}	1	2	3	4	5	6	7
1	45	49	35	18	36	14	90
2	42	62	58	46	35	58	75
3	38	15	42	58	64	75	31
4	10	20	30	57	18	45	46
5	75	15	46	35	26	14	70
6	50	60	10	80	66	55	40
7	80	60	20	25	16	34	71

Figure 5 shows an optimum solution for example 2 (again, the positive flows f_{ij} are shown). The optimum objective function value is 4327. Note that just by introducing a fixed cost c_{ij}^{t} per vehicle to be used on a link, the solution looks quite different than the one for example 1.

Example 3 corresponds to the model formulation (59)–(75). $M = 2$ modes are used with $Q^1 = 10$ and $Q^2 = 20$. The fixed cost coefficients c_{ij}^{f} are defined in table 2. The mode dependent fixed cost coefficients c_{ij}^{fm} are provided in table 5. For mode $m = 1$ we use the cost coefficients c_{ij}^{v1} and c_{ij}^{t1} like they were defined in tables 3 and 4, respectively. For mode $m = 2$ the values of the parameters c_{ij}^{v2} and c_{ij}^{t2} are specified in tables 6 and 7, respectively.

Figure 6 shows an optimum solution for example 3 (here, the flows f_{ij}^{1}/f_{ij}^{2} are shown). The optimum objective function value is 3736. Note again that just by introducing new aspects (a second mode) and keeping all other parameters as they were before in example 2, the solution looks completely different. It should be noted in example 3 that one link indeed uses a mix of modes (see the link from node 2 to node 1). Some nodes (see, e.g., node 6) have different ingoing and outgoing modes. In practice this could mean, for instance, that small trucks are used to go to node 6 and large trucks are required to come from node 6. As a consequence one would face a number of empty moves of the vehicles. Additional constraints and terms in the objective function that take into account such empty moves may become relevant in some applications. This example also reveals the

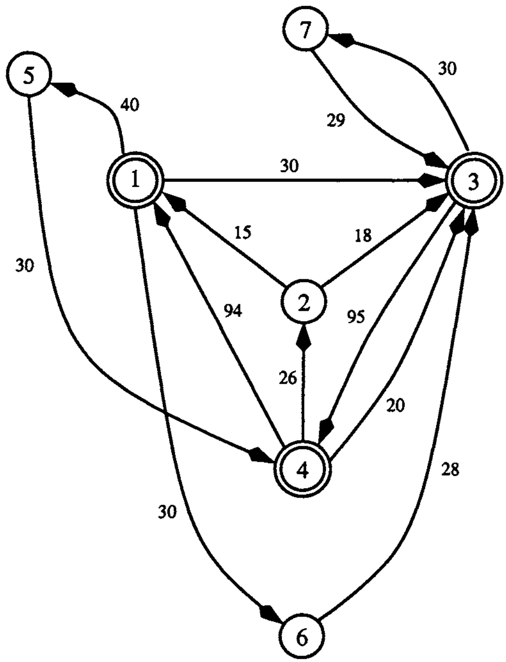

Figure 5: An Optimum Solution for Example 2

Table 5: Mode Dependent Fixed Cost Coefficients c_{ij}^{f1}/c_{ij}^{f2} for the Examples

c_{ij}^{f1}/c_{ij}^{f2}	1	2	3	4	5	6	7
1	3/9	5/5	4/7	8/8	3/4	1/6	5/5
2	6/6	4/2	7/3	8/3	9/4	6/8	4/7
3	6/3	8/1	5/4	5/5	3/6	6/8	4/6
4	3/3	5/2	7/4	9/7	3/9	5/9	4/5
5	3/1	5/4	7/7	8/8	9/5	5/4	2/3
6	4/5	8/7	9/8	3/4	4/2	4/1	7/3
7	8/2	5/4	4/5	6/7	2/8	1/9	3/3

Table 6: Variable Cost Coefficients c_{ij}^{v2} for the Examples

c_{ij}^{v2}	1	2	3	4	5	6	7
1	8	7	5	6	1	3	2
2	2	3	2	2	7	9	6
3	5	3	5	1	3	1	7
4	2	6	3	1	5	6	4
5	5	6	3	2	1	4	7
6	5	6	3	2	4	5	7
7	1	4	5	6	3	2	1

Table 7: Cost Coefficients c_{ij}^{t2} for the Examples

c_{ij}^{t2}	1	2	3	4	5	6	7
1	90	60	50	30	40	20	100
2	50	70	75	60	55	69	140
3	70	22	60	90	100	133	50
4	18	33	53	67	19	80	70
5	80	17	66	55	50	27	110
6	99	69	15	140	77	99	50
7	90	100	35	30	30	60	80

phenomenon of split quantities. It is a little hard to see in figure 6, but if we would examine in more detail which quantities q_{ij} go what what ways — which can easily be done by inspecting the values of the decision variables x — it would turn out in this example that the five units to be shipped from node $i = 2$ to node $j = 3$ are split: Four units go directly from node 2 to node 3, but one unit goes from node 2 to node 1 and from there it goes to node 3.

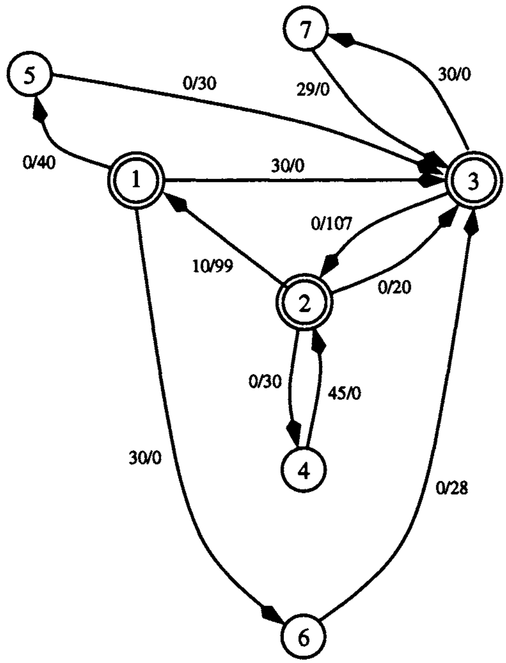

Figure 6: An Optimum Solution for Example 3

5 Conclusion

This paper is devoted to the modeling of economies of scale within hub & spoke network design problems. First we have shown how economies of scale are modeled in traditional hub & spoke network design models and it turned out that all flows have a constant unit cost. Interhub connections allow to reduce the unit cost of flow by a factor α. Hence, all traditional models do not represent flow dependent unit costs and therefore do not reflect the issue of economies of scale in a correct manner. This is rather surprising, because economies of scale are a key driving force for installing hub & spoke systems. Given that (almost) all researchers up to today have used this (incorrect) way of modeling economies of scale, it is time to question this, because obviously wrong models lead to wrong decisions no matter how good the procedures to solve these models are.

It seems that only very few authors, namely O'Kelly and Bryan (1998) and Klincewicz (2002), have noticed this serious defect before. To be fair, Podnar et al. (2002) and Campbell et al. (2004a, 2004b) also question traditional models and discuss approaches where the discount factor α is applied only under certain, flow dependent circumstances, but they still rely on that factor α which is not a convincing approach for modeling economies of scale. O'Kelly and Bryan (1998) and Klincewicz (2002) have suggested a non–linear, concave cost function to model economies of scale between interhub connections in a much more appropriate way. Using a piecewise linear approximation, these authors provide a linear mixed–integer model formulation (and solution procedures not discussed here) where the unit cost is flow dependent. However, the motivation for using this non–linear, concave cost function for modeling economies of scale is somewhat weak in our opinion, because there is no cost accounting argument that justifies such a cost function.

Given this, we contribute the following: First of all, we allow that economies of scale do not only occur on interhub connections only, but they can occur on all kinds of connections, a point that has been made by Bryan (1998) already before. So, we propose models to reflect this. Furthermore, we give an economic explanation for the occurrence of economies of scale which allows us to derive a cost function that is less artificial and can be explained better than the one proposed in the literature. The sources for economies of scale mentioned in this paper are (i) quantity discounts (if a third party is employed), (ii) fixed costs, and (iii) multiple modes. Models and examples are provided in this paper to illustrate the latter two aspects.

Future work should be dedicated to develop and test exact and heuristic solution procedures for these models. Also, an emphasis should be on investigating real–world applications to explain by cost accounting arguments which costs really matter and how these costs are calculated to get much more reliable solutions.

Acknowledgement

We thank James Campbell, Sergio García Quilles, John Klincewicz, Jadranka Skorin–Kapov, and Bernd Wagner for their comments on a draft version of this paper. Maria Merker has implemented the models from section 3 in AMPL.

References

[1] Bryan DL (1998) Extensions to the hub location problem: Formulations and numerical examples. Geographical Analysis 30:315–330

[2] Bryan DL, O'Kelly ME (1999) Hub-and-spoke networks in air transportation: An analytical review. Journal of Regional Science 39:275–295

[3] Campbell JF (1994a) A survey of network hub location. Studies in Locational Analysis 6:31–49

[4] Campbell JF (1994b) Integer programming formulations of discrete hub location problems. European Journal of Operational Research 72:387–405

[5] Campbell JF, Ernst AT, Krishnamoorthy M (2002) Hub location problems. In: Drezner Z, Hamacher HW (eds.) Facility Location – Applications and Theory. Springer, Berlin, pp. 373–407

[6] Campbell JF, Ernst AT, Krishnamoorthy M (2004a) Hub arc location problems: Part I – introduction and results. Management Science, to appear

[7] Campbell JF, Ernst AT, Krishnamoorthy M (2004b) Hub arc location problems: Part II – formulations and optimal algorithms. Management Science, to appear

[8] Cánovas L, García S, Marín A (2004) Localización de Concentradores. In: Pelegrín B (ed.) Avances en Localización de Servicios y sus Aplicaciones. Servicio de Publicaciones de la Universidad de Murcia, pp. 35–58

[9] Cánovas L, Landete M, Marín A (2005) New formulations for the uncapacitated multiple allocation hub location problem. European Journal of Operational Research, to appear

[10] Ebner G (1997) Controlling komplexer Logistiknetzwerke. GVB Schriftenreihe 34, Nürnberg

[11] Ernst AT, Hamacher H, Jiang H, Krishnamoorthy M, Woeginger G (2002) Uncapacitated single and multiple allocation p-hub center problems. Working Paper, CSIRO Mathematical and Information Sciences, Clayton

[12] Ernst AT, Krishnamoorthy M (1999) Solution algorithms for the capacitated single allocation hub location problem. Annals of Operations Research 86:141–159

[13] Garfinkel RS, Sundararaghavan PS, Noon C, Smith DR (1996) Optimal use of hub facilities: A two–hub model with fixed arc costs. Sociedad de Estadística e Investigación Operativa, Top 4:331–343

[14] Gavish B (1992) Topological design of computer communication networks – the overall design problem. European Journal of Operational Research 58:149–172

[15] Kara BY, Tansel BÇ (2000) On the single–assignment p–hub center problem. European Journal of Operational Research 125:648–655

[16] Klincewicz JG (1998) Hub location in backbone/tributary network design: A review. Location Science 6:307–335

[17] Klincewicz JG (2002) Enumeration and search procedures for a hub location problem with economies of scale. Annals of Operations Research 110:107–122

[18] Mayer G (2001) Strategische Logistikplanung von Hub&Spoke–Systemen. Deutscher Universitäts–Verlag, Wiesbaden

[19] O'Kelly ME (1986) The location of interacting hub facilities. Transportation Science 20:92–106

[20] O'Kelly ME (1987) A quadratic integer program for the location of interacting hub facilities. European Journal of Operational Research 32:393–404

[21] O'Kelly ME, Bryan DL (1998) Hub location with flow economies of scale. Transportation Research Part B 32:605–616

[22] O'Kelly ME, Miller HJ (1994) The hub network design problem – a review and synthesis. Journal of Transport Geography 2:31–40

[23] Podnar H, Skorin–Kapov J, Skorin–Kapov D (2002) Network cost minimization using threshold based discounting. European Journal of Operational Research 137:371–386

[24] Skorin–Kapov D, Skorin–Kapov J, O'Kelly M (1996) Tight linear programming relaxations of uncapacitated p–hub median problems. European Journal of Operational Research 94:582–593

[25] Stadtler H (2004) A general quantity discount and supplier selection mixed integer programming model. Working Paper, University of Hamburg

[26] Wagner B (2003) An exact solution procedure for a cluster hub location problem. Working Paper, TU Darmstadt

[27] Wagner B (2005) Design von Hub&Spoke-Netzwerken in der Logistik – modellbasierte Lösungsansätze. Ph.D. dissertation, TU Darmstadt, in preparation

[28] Wlček H (1998) Gestaltung der Güterverkehrsnetze von Sammelgutspeditionen. GVB Schriftenreihe 37, Nürnberg