HUB LOCATION UNDER CAPACITY CONSTRAINTS

Victoria Rodríguez; María Jesús Alvarez; L. Barcos

Abstract

Hub-and-spoke networks are employed in cargo transportation. This paper presents a model for hub location in these kinds of networks. Hubs are considered capacity-limited. For that reason, costs due to congestion in hubs are introduced into the model. A Simulated Annealing algorithm has been developed to solve the model. The algorithm includes a balanced module, the purpose of which is to reduce congestion. Each hub is modelled as an M/M/1 queuing system. The algorithm has been tested in randomly created networks. The results show that when a situation of congestion occurs, the proposed algorithm enables better solutions to be found, thereby improving the standard of service.

Keywords: hub location, cargo distribution, facility planning, Simulated Annealing hub capacity.

Introduction

Since Weber presented the findings of his studies in 1900, many researchers have centred their efforts on studying the problems of hub location. In the area of transport, many of their studies have been concerned with defining the optimum location for manufacturing plants, distribution centres, and hubs. Handler and Mirchandani (1979) offer a classification of the problems associated with location and deal with some of the solutions in depth. Hurter & Martinich (1989) study the problem of locating a new manufacturing plant bearing in mind different production theories.

In order to solve the problem of hub location, it is necessary to study the problem of route design. The traditional means of handling the task of determining the best location for a hub consists of assigning nodes to already existing hubs. In addition, when trying to optimize distribution routes, it is generally accepted that the hubs pre-exist, and the most common strategy is to study each existing origin/destination pair and then determine the best route, depending on the available hubs. Several authors have studied and compared the strategies of direct transport and transport via different terminals.

Blumenfeld, Lawrence, Diltz y Daganzo (1985) study the optimum transport strategy, considering both the direct transport of goods and also the sending of those goods via a terminal. Hall (1987) studies the strategy of direct goods transport as against the transport of goods via a terminal in a single terminal network with a lot of origins and a few destinations. Later, the same author (Hall, 1989) studies the problem in networks with several terminals. Leung, Magnanti y Singhal (1990) study the problem of point-to-point transport within a distribution network. A comprehensive study and review of these models can be seen in Barcos (2002).

In order to solve the problem of locating p hubs in an N-node network, Campbell (1994) proposed a linear model; later, Skorin-Kapov et al. (1997) presented a compact formulation of this problem. Aykin (1995) proposes an algorithm to solve the problem of locating p hubs in an N-node network. O'Kelly & Bryan (1998) develop a model called FlowLoc, which consists of introducing changes to the function which was the subject of the model proposed by Skorin-Kapov et al. (1997). Through these changes they show that the cost of transport between hubs goes down when the flow between hubs increases.

Berman et al. (1985) offer two models for solving the one-median problem in which they consider that the problem can be seen as a M/G/1 queuing system. With respect to the location of hubs, if the capacity of the hubs is considered as limited, at certain moments they may become congested and therefore the time the cargo spends in the hub may negatively affect the standard of service offered. A possible approach to the problem is to consider the hub as a M/M/1 queuing system (Rodríguez 2002).

The purpose of this paper is to consider costs due to possible congestion in hubs and to balance the cargo flows in hubs in order to improve delivery times and thus provide a more efficient service to customers.

The first part of this paper presents the problem and its characteristics; the second part offers the mathematical representation of the problem, and the third part proposes a resolution algorithm and some of the findings when it has been applied. Finally, conclusions are presented.

1. Defining the problem

An air or cargo distribution network almost always involves passenger and/or cargo exchange between the origin/destination pairs in the network. In order to optimize the distribution process, hubs are located at the points where cargo arrives from many origins; it is reorganized at this point and sent on to an intermediate or final destination. The aim of this configuration is the reduction of costs, economy of scale being achieved by transporting large quantities of cargo between hubs in such a way that neither trucks nor planes have to make long journeys with little load. The resulting network is hub-and-spoke-shaped. In this kind of distribution network, depending on the policy adopted, cargo going from i to j can be sent directly or it can be sent through one or two hubs. In practice, cargo does not usually pass through more than two hubs (Aykin 1995). The cost per unit of flow and unit of distance for cargo sent between hubs is generally lower than for cargo sent directly due to the large quantities of traffic and the centralization of operations. Cost may, therefore, be reduced as a result of the economy of scale. However, the distance travelled by the cargo when it passes through a hub is greater than the distance travelled if it is transported directly, The time it takes the cargo to get from its origin to its destination is, therefore, greater, and it becomes necessary to compensate for this increase in cost and transport-time by reducing costs through the grouping of cargo at hubs. These are the aspects that traditionally have been born in mind when studying the problem of the location of hubs. However, over the last few years as a result of increasing competition between companies in the sector, it has become necessary to reduce service times and to meet the delivery dates offered. Nowadays, this is a factor of strategic importance for these companies. Delivery time can be divided into two parts: transport time, and time spent at the hub. As the classification capability of the hub is limited, when trucks arrive at the hub and find it busy, they have to wait to be attended to; therefore the time the cargo spends at the hub will be the sum of the time spent waiting in the queue plus the service time.

Figure 1 Distribution of the delivery time between hubs

The service given at the hub has been considered an M/M1 queuing system. $1/\mu$ is defined as the average service time and λ as the average arrival rate. The value of λ varies, depending on the number of trucks assigned to a hub. From Little's formula for calculating the average time spent waiting in queue (Chee Hock, 1996) it can see that as the value of λ increases, the time spent waiting in the queue increases more quickly. As the value of λ is determined by the number of trucks assigned to the hub, the capacity of the hub $CHub$ is defined as the maximum number of trucks that may be assigned to the hub. If more trucks are assigned to the hub, the resulting value of λ will mean that the waiting time in the queue will be so long that it will be impossible to reach the delivery times offere

(Rodríguez 2002).

$$\omega = \frac{n}{\lambda}$$

where:

n average number of customers in the queue.

ω average time spent in the queue.

λ average arrival rate.

Also, it is known that:

$$n = \frac{\lambda}{(\mu - \lambda)}$$

Therefore:

$$\omega = \frac{1}{\mu - \lambda}$$

In the figure it is possible to see that from a certain arrival rate value onwards, the waiting time increases so much that it becomes impossible to fulfil the standard of service and delivery dates. Because of this, part of the nodes in hubs with a high λ will be assigned to other hubs which receive less cargo.

Figure 2 Waiting time as a function of the arrival time

The value of $CHub$ is influenced by the rate of service and the characteristics of the network. The key point is the time spent in the hub by the cargo. If that time is too long due to congestion, the truck which should leave the hub with the reorganised cargo leaves it late or without the total load. The company should determine the value of $CHub$ according to the targeted percentage of cargo served on time (also called level of service).

2. Model

The model developed in this study is based on the model proposed by Aykin (1995) for the location of p-hubs in an N-node network with capacity and distance limitations when the service standard offered needs to be fulfilled.

$$Min \sum_{i}\sum_{k}\sum_{t}\sum_{j} W_{i,j}\left(\beta C_{i,k} + \alpha C_{k,t} + \beta C_{t,j}\right) V_{i,k,t,j} \quad (1)$$

s.t.

$$\sum_{k}\sum_{t} V_{i,k,t,j} = 1 \quad \forall \quad i,j \quad (2)$$

$$\sum_{k} Y_{k} = p \quad (3)$$

$$\sum_{i \neq k}\sum_{\substack{l \neq k \\ h \neq k}} \left(V_{i,l,h,k} + V_{k,l,h,i}\right) \leq M\left(1 - Y_{k}\right) \quad \forall \quad k \quad (4)$$

$$\sum_{\substack{h \neq k \\ h \neq t}} \left(V_{k,k,h,t} + V_{t,h,k,k} + V_{t,t,h,k} + V_{k,h,t,t}\right) + \sum_{h} \left(V_{k,h,h,t} + V_{t,h,h,k}\right) \leq M\left(2 - Y_{k} - Y_{t}\right) \quad \forall \; k \neq t \quad (5)$$

$$\sum_{k} Z_{i,k} = 1 \quad \forall \; i \quad (6)$$

$$\sum_{t}\sum_{j} V_{i,k,t,j} \leq \left(N - p + 1\right) Z_{i,k} \quad \forall \; i,k \quad (7)$$

$$\sum_{i}\sum_{k} V_{i,k,t,j} \leq \left(N - p + 1\right) Z_{j,t} \quad \forall \; t,j \quad (8)$$

$$V_{i,k,t,j}\left(t_{i,k} + t_{k,t} + t_{t,j}\right) \leq tMax \quad \forall \; i,k,t,j \quad (9)$$

$$\sum_{i}\sum_{t}\sum_{j} \left(V_{i,k,t,j} W_{i,j}\right) \leq CHub_{k} \quad \forall \; k \quad (10)$$

$$\sum_{i}\sum_{t \neq k}\sum_{j} \left(V_{j,t,k,i} W_{j,i}\right) \leq CHub_{k} \quad \forall \; k \quad (11)$$

$Z_{i,j}, V_{i,k,t,j}, Y_k \in \{0,1\} \quad \forall\ i, k, t, j$

where:

$V_{i,k,t,j}$ = 1 when the cargo going from i to j passes through hubs k and t

= 0 in any other case.

$Z_{i,k}$ = 1 if the i node is assigned to hub k

= 0 in any other case.

$W_{i,j}$ cargo flow that has to be taken from i to j.

$C_{i,j}$ unit cost of cargo travelling from i to j.

$d_{i,j}$ distance between nodes i and j.

p number of hubs.

Y_k = 1 when the k node is a hub

= 0 in any other case.

F_k the fixed cost of locating a hub in k.

M a very high number.

α factor which represents the economy of scale between hubs, such that: $0 < \alpha < 1$

β factor which represents the economy of scale between a hub and a node which is not a hub, such that: $0 < \beta < 1$ and $\beta > \alpha$

$DMax$ maximum distance between any origin/destination pair.

$CHub_k$ capacity of hub k.

The set of constraints (2) ensures that there is only one route for each origin/destination pair. The constraints (3) ensure that p hubs are located in the network.

Depending on the location of the hubs, the constraints (4) and (5) set the necessary variables at 0 so as to consider these routes as routes that pass through hubs. The economy of scale is thus applied properly and irregular routes are avoided. The variables $V_{i,i,i,j}$, if i is a hub, and $V_{i,j,j,j}$, if j is a hub, and $V_{i,i,j,j}$, if i and j are hubs, all represent the same route without stops. The irregular routes $k \rightarrow t \rightarrow k \rightarrow i$ and $k \rightarrow t \rightarrow k \rightarrow t$, when $k \neq t$, are not included in the model. If k is a hub, the constraints (4) set a value of zero for all the variables associated with the routes which start or finish at k except $V_{k,k,h,i}$ and $V_{i,h,k,k}$. Furthermore, according to the constraints (5), if two k and t nodes are hubs, all the variables which represent the cargo routes which originate in k and have t as their destination or vice versa have a value of zero except $V_{k,k,t,t}$ and $V_{t,t,k,k}$.

As can be seen, here the values α, β - have been taken as the economy of scale. The economy of scale, α, achieved when transporting cargo between hubs is constant and greater than the economy of scale, β, achieved when the cargo is transported between nodes which are not hubs and their respective hub.

The set of constraints (6) simplifies the assignation of each node. This means that all cargo that enters and leaves a node passes through the hub to which it has been assigned. With the sets of constraints, (7) and (8), all the routes starting at i pass through the hub to which i has been assigned, and all the routes which finish at j pass through the hub to which j has been assigned.

The set of restrictions (9) limits the transport time between all the origin/destination pairs in the network so as to be able to meet delivery times.

The hub has two peak usage periods. The first of these is when the trucks arrive with the cargo from the hub's assigned nodes. Here, the constraints (10) ensure that the allowed nominal capacity of the hub is not exceeded by preventing cargo from entering. The second peak usage period is when trucks arrive from other hubs. Here, the constraints (11) limit the maximum cargo load entering the hub at that moment.

3. Solution algorithm

In order to solve the problem of *hub* location an algorithm has been developed. This algorithm consists of two parts: in the first part an initial solution is determined, while in the second part this solution is improved.

Initial Solution

When seeking the initial solution, all the nodes in the network are considered hubs. Therefore, each node is, effectively, assigned to itself. The total cost is calculated with this solution. In each loop the cost of the solution of closing down one of the hubs in the network is calculated, and the hub which represents the least increase in cost is closed. When an h hub is closed down, it becomes necessary to reassign all its assigned i nodes to another r hub in the set of hubs.

The assignation function is as follows:

$$A_{i,r} = \min_{k \in H} \{A_{i,k}\} \qquad \forall \ i \in K_r, \ r \in H \qquad (12)$$

Where:

K_r Set of all the nodes assigned to hub r.

$A_{i,k}$ transport cost of the cargo between i and the other nodes of the network in both directions if the node i is assigned to hub k.

The value of $A_{i,k}$ is calculated in the following way:

$$A_{i,k} = \sum_t \sum_j \left(W_{i,j} + W_{j,i}\right)\left(\beta C_{i,k} + \alpha C_{k,t} + \beta C_{t,j}\right) \qquad (13)$$

The result of each loop is the closure of a hub. Therefore, if it is necessary to locate p hubs in a network of N-nodes, it is necessary to do the loop $(N - p)$ times in the initial stage.

Improvement of the initial solution

The second stage of the model is the improvement stage. Here, through the use of Simulated Annealing, we are seeking to improve the initial solution previously obtained. In this stage, exchanges are made between those nodes that are *hubs* and those that are not. To do so, we try switching the location of each of the ***h*** hubs with one of the ***n*** nodes assigned to the hub. Finally, the most money-saving or least money-losing changeover is selected in accordance with the following equation for estimating savings:

$$U_{t,r} = \max_{k \in H} \{R_{k,r}\} \qquad (14)$$

Where:

$R_{k,r}$ estimated saving when substituting hub ***k*** for node ***r***.

To calculate the values for $R_{k,r}$, it is necessary to calculate the amount of cargo which goes from hub **k** to hub **l**:

$$I_{k,l} = \sum_{\{i / Z_{i,k} = 1\}} \sum_{\{j / Z_{j,l} = 1\}} \left(W_{i,j}\right) \qquad (15)$$

Where:

$I_{k,l}$ cargo which goes from hub ***k*** to hub ***l***.

The saving achieved when hub k is replaced by the node, r, is:

$$\begin{aligned} R_{k,r} = & \sum_{\{i / Z_{i,r} = 1\}} O_i \left(C_{i,r} - C_{i,k}\right) + \sum_{\{i / Z_{i,r} = 1\}} D_i \left(C_{r,i} - C_{k,i}\right) \\ & + \sum_{\substack{q \in K \\ q \neq r}} \left(I_{r,q}\right) \alpha \left(C_{r,q} - C_{k,q}\right) + \sum_{\substack{q \in K \\ q \neq r}} \left(I_{q,r}\right) \alpha \left(C_{q,r} - C_{q,k}\right) \end{aligned} \qquad (16)$$

Where:

O_i total cargo originating at node ***i***.

D_i total cargo arriving at node ***i***.

$Z_{i,k}$ = 1 if node i is assigned to hub k.

= 0 otherwise.

Since $U_{t,r}$ is an estimate of the saving involved and is its lower limit, in some cases the result could be a negative value. However, on updating the assignation of nodes to hubs, the switching of node r to a hub is a good choice. Because of this, the algorithm calculates the probability of acceptance. The exchange between r and t is calculated in the following function:

$$P(\text{acept at } \Delta) = \begin{cases} e^{-\Delta/T}, & \text{if } U_{r,t} < 0 \\ 1, & \text{if } U_{r,t} \geq 0 \end{cases} \quad (17)$$

$$\Delta = U_{t,r} / (\text{value of the objective function for the best known solution}) \quad (18)$$

As $U_{r,t}$ is the lower limit of the saving involved when t substitutes r, if $U_{r,t} \geq 0$, it is always recommendable to make this change, since the probability of accepting the change is 1; while, if $U_{r,t} \leq 0$, the changeover may either improve or worsen the situation. This is why it is necessary to calculate the probability of accepting the change. Once this value has been calculated, a random number ω is generated which is compared with the value of the probability of accepting the change. If $\omega < P(\text{accept at } \Delta)$, the switchover between r and t will be made.

The cooling function is:

$$T_n = T_{n-1} v \quad (19)$$

Where:

T_n temperature value for the loop n.

ν cooling parameter.

The FLAG parameter has been brought into the algorithm at this stage so that the algorithm does not carry out the improvement cycle ad infinitum. This parameter stops the algorithm if no change has been made after all the nodes have been evaluated.

Reduction of congestion

The aim of the second modification to the assignation function is to balance the flow of cargo in the network. First the quantity of cargo that passes through each hub is calculated using equations (10) y (11). Next, the hubs where the nominal capacity is exceeded and also those where the amount of cargo is not below their nominal capacity are identified. The necessary reassignations are then made to reduce the cargo going through those hubs where the nominal capacity is being exceeded. In order to reduce congestion in an h hub, the $(Z_{i,h} = 1)$ i node nearest to a non-congested k hub is selected from all the h hub's assigned nodes, and the i assignation is changed using $(Z_{i,h} = 0\,;\; Z_{i,k} = 1)$.

As can be seen in the flow chart, this heuristic is based on Aykin's proposal (1995), but certain modifications have been made so that the algorithm can go beyond the local optimums and accept solutions which are worse than the best solution achieved to date, in accordance with the function of probability and the value of the number produced randomly. In addition, the balance module has been introduced in order to reduce congestion at the hubs.

Objective function

In real distribution networks, meeting delivery dates between all the origin/destination pairs would push up costs too much; so in this study, this initial problem was eased by introducing a surcharge in the cost function (i.e. the objective function) for those origin/destination pairs which were not being delivered to on time.

The objective function for this problem can be written in the following way:

$$\text{Min} \sum_{i}\sum_{k}\sum_{t}\sum_{j}\left(W_{i,j}+S_C WT_{i,j}\right)\left(\beta C_{i,k} + \alpha C_{k,t} + \beta C_{t,j}\right)\, V_{i,k,t,j} \qquad (20)$$

In both cases:

$$S_C = \begin{cases} 0 & if \ \left(t_{i,k} + t_{k,t} + t_{t,j} \right) \leq tMax \\ S_C & if \ \left(t_{i,k} + t_{k,t} + t_{t,j} \right) > tMax \end{cases} \quad (21)$$

where:

$WT_{i,j}$ amount of load from i to j which is not delivered on time.

S_C surcharge due to non-delivery on time.

Figure 3 Flowchart for the algorithm

4. Computational results

The model was evaluated in 10 randomly created networks of 52 nodes each . The networks were created in an area of 1200*1200 Km2 with 6 centres, where most of the nodes in the network were located. The load matrices and the distance were also generated at random. The cost matrix was based on the distance matrix. In each case, 4 situations were simulated. The problem of locating 3 hubs was solved for each situation. The capacity of any hub was measured in terms of how many trucks were assigned to it. A homogenous fleet was assumed as was an average speed. The delivery time was 24 hours; therefore in the tests it was assumed that 10 hours were available for transporting and handling the cargo at the hub. The remaining time went to delivery and local collection. It was assumed that the time the cargo spent at the hub was 1 hour when its capacity had not been exceeded; if the capacity was exceeded, this time was increased. The conditions for the tests carried out can be seen in the table below,

Table 1 Characteristics of the hubs

It is considered that the cargo that is not delivered within the service period has an additional cost of 20% over the cost of delivering the cargo within the delivery time.

To see the results of reassigning in order to balance out the cargo traffic between hubs, the results obtained with Aykin's algorithm (1995) were compared with the results obtained from the same problems introducing the task of balancing the assigning of nodes. The results are compared using equation (20).

$$\%\ \text{improvement PSIT} = \left(\frac{PSIT_b - PSIT_0}{PSIT_0} \right) \qquad (22)$$

where:

$PSIT_b$ number of pairs served in time balancing the nodes assignation.

$PSIT_0$ number of pairs served in time without balancing the nodes assignation.

$$\%\ \text{improvement LSIT} = \left(\frac{LSIT_b - LSIT_0}{LSIT_0} \right) \quad (23)$$

Where:

$LSIT_b$ load served in time balancing the nodes assignation.

$LSIT_0$ load served in time without balancing the nodes assignation.

Table 2 Results for the problems when w=1.8

Table 3 Results for the problem when w=3.8

It can be seen that when the additional waiting time due to congestion is short (0.8 hours), there are several cases where the number of origin/destination pairs and the amount of cargo that is not delivered on time increases when the assignment of nodes is balanced out. This is because when the changes necessary in order not to exceed the hub capacity are made, the reduction (in the amount of cargo and origin/destination pairs which are not delivered to on time due to congestion) is lower than the increase in late deliveries due to the longer distance resulting from the reassignation of nodes.

For the same reason when the capacity of a hub is lower, the figures representing disimprovements are greater because more changes need to be made to meet capacity restrictions. Alternatively, when the additional waiting time at the hub due to congestion is long (2.4 hours), balancing the assignation of node offers very satisfactory results because the increase in the distance the cargo travels is compensated for by the decrease in the time spent waiting at the hub.

In order to study the consequences of balancing out the assignation of nodes, the total cost of the solution was broken down into three items: the first refers to the cost of transport, whilst the second and third reflect the cost of not meeting the delivery date. The delivery date accorded is not met either because the journey is very long or because the cargo spends a lot of time at the hub. Bearing in mind these factors, the total cost would look like this:

$$TC = TrC + SCD + SCC \qquad (24)$$

Where:

TC	total cost
TrC	travel cost
SCD	surcharge due to distance
SCC	surcharge due to congestion

The table shows the result of comparing the costs of balancing the assignation and not balancing it:

$$\%\ \text{improvement cost} = \left(\frac{Cost_0 - Cost_b}{Cost_0} \right) \qquad (25)$$

where:

$Cost_b$ cost of the solution without balancing the nodes assignation.

$Cost_0$ cost of the solution balancing the nodes assignation.

These values were calculated for each of the problems. The table shows the averages for the ten problems in each case.

Table 4 Change in the components of the cost controlling the overload in the hubs

It can be seen from the results that in all cases the cost of late deliveries due to distance increases on balancing the assignation of nodes, while the cost of late delivery through congestion is reduced. It can also be seen that on average the total cost of the solution improves.

Conclusions

The contribution of this study is to take the standard of service into account when locating hubs and to consider that each hub has limited capacity. The standard of service is measured according to the amount of cargo to be served within the delivery time and to the number of origin/destination pairs to be served on time. Deliveries are considered not to be made on time due, mainly, to two factors: firstly, that the cargo has to travel long distances; and, secondly, that the cargo remains at the hub for too long. The latter issue can be controlled by reducing the waiting time due to congestion at the hub; in this way the model allows the problem of hub location to be solved by reducing congestion at the hubs.

Distribution policy is strict and restrictive in cases where all the nodes in the network send and receive cargo via a hub. This model introduces the delivery time restriction in addition to the other restrictions already contemplated in previous studies.

The objective function considered here consists of three factors; the first reflects the cost of transport between each origin/destination pair; the second and third, factors concern the cost arising from not delivering the goods within the agreed time limit. The second factor reflects the surcharge generated when cargo makes too-long-a journey, while the third factor reflects the cost created by cargo which is delivered late because it spent too much time at the hub. The methodology used to improve the initial situation is Simulated Annealing. The cooling function used is the traditional one.

When the waiting time at the hub is long (in other words, when a situation of congestion occurs), the proposed model enables the global results to be improved because the percentage of cargo delivered on time is increased. This methodology is useful for cases where congestion is a problem at hubs.

When additional waiting time at a hub is not excessive, the modifications made to the problem to reduce congestion at the hub may cause a drop in the standard of service in the network. This is because the drop in late deliveries through congestion does not offset the rise in late deliveries resulting from the greater distances travelled. In contrast, when the waiting time is long, better results are obtained.

Acknowledgements

The research on which this paper is based was funded by Tecnun (Universidad de Navarra) and the Diputación Foral de Guipuzcoa.

References

Aykin, T., 1995. Networking Policies for Hub-and-Spoke Systems with Applicaion to the Air Transportatios System. *Transportation Science* 29, pp. 201-221

Barcos, L., 2001. Optimización de Rutas de Transporte de Carga de Muchos Orígenes a Muchos Destinos Mediante Colonias de Hormigas, Tesis doctoral, Tecnun, Universidad de Navarra. San Sebastián

Berman, O., Larson, R., Chiu, S.S., 1985. Optimal server location on a network operating as an M/G/1 queue. Operations Research 33(4) pp.746-771

Blumenfeld, D.E, Lawrence, D.B., Dilzt, J.D., Daganzo, C., 1985. Analyzing Trade-offs, Inventory and Productions Costs on Freight Networks. Transportation Research 19B pp. 361-380

Campbell, J.F., 1994. Integer Programming Formulations of Discrete Hub Location Problems. European Journal of Operational Research 72, pp.387-405

Chee Hock, Ng., 1996. Queueing Modelling Fundamentals, John Wiley & Sons, West Sussex, England.

Hall, R.W., 1987. Direct versus terminal freight routing on a network with concave costs. Transportation Research B, 21B, pp. 287-298.

Hall, R.W., 1989. Configuration of an overnight package air network. Transportation Research A, 23A(2) pp. 139-149.

Hander, G.Y., Mirchandani, P.B., 1979. Location on Networks. The Massachusetts Institute of Technology. Cambridge, USA.

Hurter, A.P., Martinich, J.S., 1989. Facility Location and the Theory of Production. Kluwer Academic Publishers, New York, USA.

Leung, J.M.Y., Magnanti, T.L., Singhal,V., 1990. Routing in Point-to-Point Delivery Systems: Formulations and Solution Heuristic. Transportation Science 24 pp. 245-260

O`Kelly, M.E., Bryan, D.L., 1998. Hub Location with Flow Economies of Scale. Transportation Research B(32) pp.605-616

Rodriguez, V., 2002. Metodología para determinar la ubicación de hubs con capacidad limitada. Tesis doctoral, Tecnun, Universidad de Navarra. San Sebastián.

Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M.E. 1997 Tight Linear Programming Relaxations of Uncapacitated p-Hub Median Problems, European Journal of Operational Research 94 pp. 582-593

TABLES

	Under capacity			Over capacity		
Capacity (trucks)	Rate of service (trucks/hour)	Rate of arrives (truck/hour)	Time in the hub (hours)	Rate of service (trucks/hour)	Rate of arrives (trucks/hour)	Time in the hub (hours)
10	13	12	1	13	12,44	1,8
	13	12	1	13	12,7	3,4
15	19	18	1	19	18,44	1,8
	19	18	1	19	18,7	3,4

Table 5 Characteristics of the hubs

	% improvement			
	Capacity 10		**Capacity 15**	
	PSIT	**LSIT**	**PSIT**	**LSIT**
P01	2,8	2,5	5,0	5,1
P02	11,3	12,1	13,9	14,5
P03	-0,8	-1,1	0,1	-0,1
P04	-8,4	-8,3	-2,5	-2,9
P05	-2,1	-2,3	6,1	6,7
P06	0,6	-0,2	3,8	3,4
P07	-1,0	-1,0	0,8	0,9
P08	-2,2	-2,0	6,7	6,9
P09	-2,4	-2,1	3,6	3,6
P10	-1,7	-2,0	5,9	5,7

Table 6 Results for the problems when w=1.8

	% improvement			
	Capacity 10		**Capacity 15**	
	PSIT	**LSIT**	**PSIT**	**LSIT**
P01	2,5	3,1	23,2	23,8
P02	46,3	46,8	46,0	46,5
P03	6,3	6,5	16,8	16,8
P04	5,3	5,5	38,2	39,2
P05	10,1	9,8	27,1	28,3
P06	2,2	1,7	10,3	10,0
P07	8,0	8,6	17,4	17,9
P08	4,4	4,1	12,3	12,4
P09	0,8	1,1	10,4	11,1
P10	1,3	1,2	16,5	17,0

Table 7 Results for the problem when w=3.8

	% mean improvement			
	Capacity	Cost due to distance	Cost due to congestion	Total cost
$\Delta\omega = 2.4$	10	-1,52	2,05	0,53
	15	-1,53	4,69	3,16
$\Delta\omega = 0.8$	10	-1,43	0,49	0,49
	15	-1,18	1,16	1,16

Table 8 Change in the components of the cost controlling the overload in the hubs

FIGURES

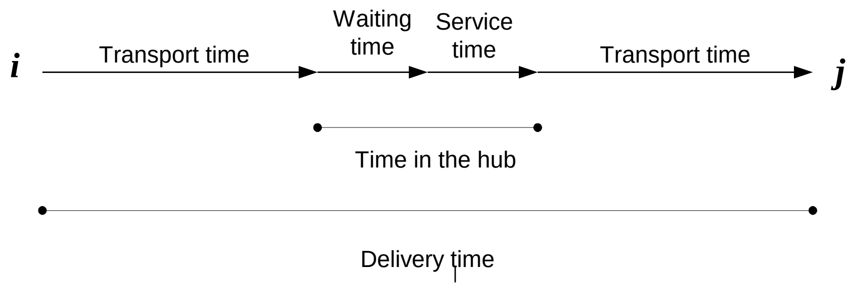

Figure 4 Distribution of the delivery time between hubs

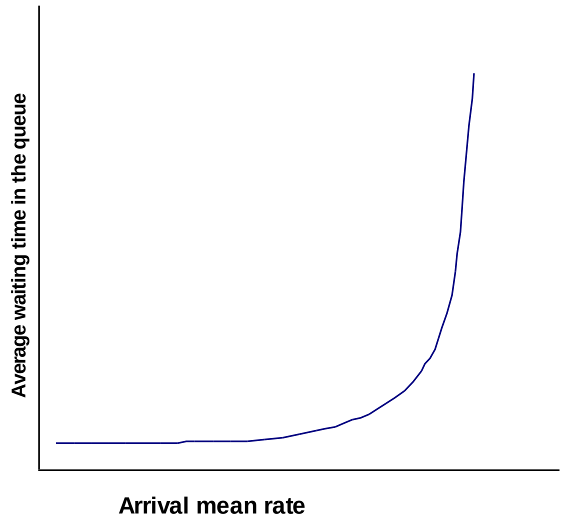

Figure 5 Waiting time as a function of the arrival time

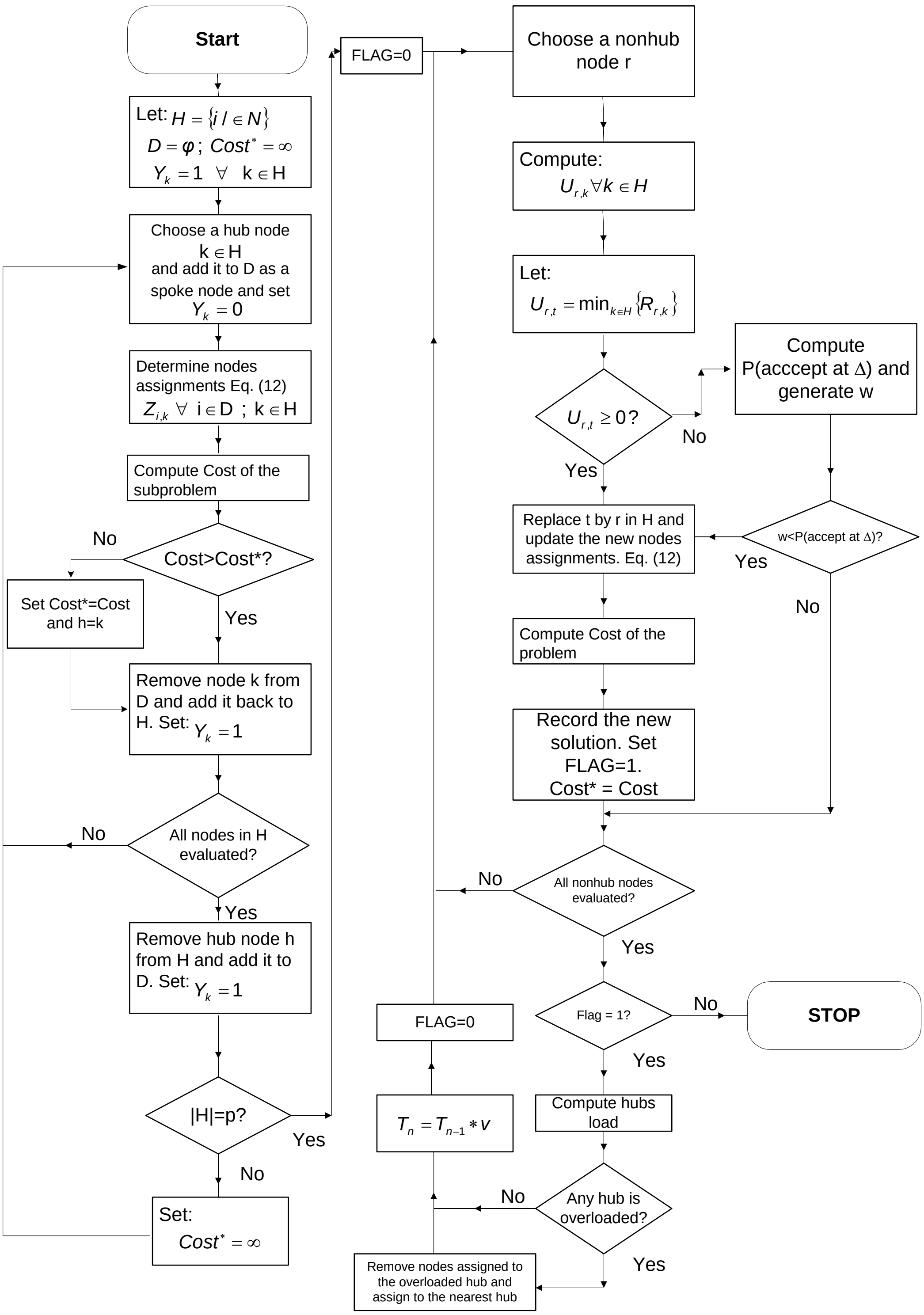

Figure 6 Flowchart for the algorithm