

## APPENDIX B. Some details on the application of each heuristic to the city model.

### Dubois et al (1979)

The following pseudo-code describes the heuristic:

(0) Define the value of the tolerance  $\sigma$

(1) Define the set  $M$  that contains all the minimum-length (i.e. sum of the distances of each arc) routes that cannot be extended preserving the length condition.

(2) Build the set  $M^*$ , which starts empty:

For each route  $L \in M$ , denote  $x$  its initial node and  $y$  its final node.

For each  $z$  not in  $L$ , define  $L_z$  the route that starts in  $x$ , goes to  $z$  and then to  $y$  always following minimum-length paths.

If the length of  $L_z$  exceeds the length of  $L$  by a fraction smaller than an exogenously fixed tolerance  $\sigma$ , then add  $L_z$  to  $M^*$  and stop searching for  $L$ . Start searching for  $L_z$ , but always comparing lengths with respect to  $L$ .

(3) Define  $\bar{M}$  as the union of  $M^*$  and the paths in  $M$  connecting origins and destinations that are not connected by paths in  $M^*$ . The routes in this set are the candidates.

(4) Arrange the routes in  $\bar{M}$  according to some rule.

(5) Define  $M'$  as an empty set. Add the paths in  $\bar{M}$  to  $M'$  until the city is fully connected (admitting transfers).

(6) Calculate the portion of trips that need one or more transfers. If this portion is "small", go to step (8). Otherwise, go to (7).

(7) Add the route in  $\bar{M}$  that minimizes the number of required transfers. Go to (6).

(8) Stop if the difference between the average travel time in the system and the average time in a system where all the trips go through the minimum-length path is "small". Otherwise, go to (9).

(9) Add the route in  $\bar{M}$  that decreases the most the average travel time. Go to (8).

In our city, the set  $M$  is composed by all the routes connecting two peripheries following an optimal path (crossing the CBD or through the subcenters ring, depending how far are the respective zones). First we assume  $\sigma < 1/3$ . We check all the possible routes  $L_z$ ; we verify that  $M^*$  happens to be composed only by the routes that go from one periphery to a periphery that is 3 zones away through the subcenters ring (instead of crossing the CBD). So the candidate set  $\bar{M}$  is composed by paths going through the subcenters ring that reach peripheries that are 1, 2 or 3 zones away, and paths that go the opposite subcenter (i.e., the one that is 4 zones away) crossing the CBD. To arrange  $\bar{M}$  (step 4), the authors propose criteria such as the most used or the least costly routes. We adapt the last one, arranging the set from the shortest to the largest route. In step (6), if we only added the shortest routes in  $\bar{M}$  (i.e., the routes that go the neighbor zones) we would connect almost the whole graph: only the CBD would remain unconnected. For this not to happen we add the routes that go from each periphery to

its opposite crossing the CBD. The graph is now connected, but all the trips that go to two or more zones away (with the exception of the opposite zone) require a transfer, so the number of transfers is at least<sup>1</sup>  $(a\gamma + b\tilde{\gamma})(n - 4)Y$ . If  $\gamma < 0.1$ , we skip step (7); otherwise we add the route that go from one periphery to the periphery that is three zones away and we eliminate all the transfers. As the travel time cannot be reduced adding more routes, the final line structure has been reached.

Afterwards, we solve for  $\sigma > 1/3$ . Doing so, the line that goes from a periphery to the opposite subcenter presents a small variation: after going to the CBD, it goes to one neighbor of the opposite subcenter and then finishes in the opposite subcenter. Notice that in this case the number of transfers will always be smaller than  $\frac{n-5}{n}(a\gamma + b\tilde{\gamma})$ , and when  $n = 8$  this is smaller than  $\frac{2}{7}$ ; therefore, for  $\sigma > 1/3$  a single structure is obtained. It is worth commenting that in the case  $\sigma < 1/3$  with  $\gamma > 0.1$  passengers whose destination is located three zones away from their origin could go either through the subcenters ring, without making transfers, or take a shorter trip changing buses in the CBD; as their choice depends on the frequencies finding the optimal frequency requires iterations. We first assign all these passengers with no transfers and calculate the resulting optimal frequencies; then we verify whether the choice with no transfers is in fact the min user cost. If not, we re-calculate considering transfers. As expected the results depend on the parameters.

### **Ceder & Wilson (1986)**

This algorithm builds routes that depart from a terminal (the first origin), searching for trees (i.e., connected graphs with no cycles). To do so, select any terminal as the first one and consider the following pseudo-code:

- (1) Search for all the nodes that you can reach from the current origin; if no node is reachable, go to (2), otherwise, for each reachable node, if it has not been connected, and if the total length of the path does not exceed the length of the shortest path by a percentage  $\sigma$ , add it to the tree. If there is no such node, go to (2). Otherwise, select any of these nodes as the new origin for the same terminal and repeat (1).
- (2) Select as the new origin the last node added to the tree that has not been an origin yet. If there is no such node, go to (3).
- (3) If there are no more terminals, end. Otherwise, select a new terminal, define it as the origin and go to (1).

In our scheme, routes start from a periphery (predefined as terminals) and necessarily goes to the own subcenter. Then it is possible to go to neighbor subcenters or to the

---

<sup>1</sup> The routes that minimize the number of transfers are those in which the user goes to the CBD first and take the second bus there. Other routes require more transfers.

CBD. We start exploring the routes through the CBD such that, depending on the value of  $\sigma$ , we will reach  $H = 3, 5$  or  $7$  subcenters. Doing so,  $H = 3$  if  $\sigma < 0.2517 (\approx \frac{1}{4})$  and  $H = 5$  otherwise (recall that  $H$  is the number of foreign subcenters that are reached crossing the CBD). Once the routes through the CBD have been explored, the rest of the routes tour the subcenter ring until the whole graph is covered. Afterwards it is impossible to extend any route without reaching a previously built route that started from the same terminal.

### **Borndörfer et al (2007)**

The optimization problem, in this scheme, turns into

$$\text{Minimize } \sum_{R \in U} p_{tr} t_R y_R + \sum_{l \in L} c_0 B_l$$

s.t.

$$\sum_{R: O(R)=s, D(R)=t} y_R = (OD)_{st} \quad \forall \text{ nodes } s, t \quad (1')$$

$$\sum_{R: a \in R} y_R \leq \sum_{R: a \in R} f_R K \quad \forall \text{ edge } a \quad (2')$$

$$\sum_{l: a \in l} f_l \leq \Gamma_a \quad \forall \text{ edge } a \quad (3')$$

$$f \geq 0 \quad (4')$$

$$y \geq 0 \quad (5')$$

The variables are the vector of frequencies  $f$  and the vector of passenger flows  $y$ . The total cost only considers travel times and a fixed cost per bus. In the original model there is a fixed cost per line, but to be consistent we put them equal to zero. Waiting times and transfers play no role.

The first equation imposes that passengers on all routes  $R$  serving a given O-D pair add up to the corresponding O-D demand. Note that this means that passenger assignment to routes is endogenous aiming at minimizing total cost, which may not coincide with the individual preference of each user. As  $K$  is the size of each bus, the second equation imposes bus capacity constraints. The third equation relates frequency to streets capacity,  $\Gamma_a$ , but in our model this is not considered so, in practice, we used  $\Gamma_a = +\infty \quad \forall a$ .

### **Cenek (2010)**

First assume that all passengers will take the shortest route to go from their origin to their destination; this permits the construction of arc-specific weights as the number of passengers that use that arc under this assumption. Then, the following procedure is applied:

(1) Select the arc with the largest weight from those that finish in a center (if that weight is 0, finish). Define the node  $u$  as the other extreme of that arc.

(2) Extend the line. Select the arc with the highest weight from those that are incident in  $u$ . Update  $u$  as the other extreme of that arc. If  $u$  is a center or has no other incident arc with positive weight, go to (3). Otherwise, repeat (2).

(3) The line has been built and added to the set of lines. Calculate  $w$  as the minimum weight of the arcs present in this line. Subtract  $w$  from the weight of all the arcs in the line. Go to (1).

To analyze the application to our scheme, some route notation is needed. A route will be denoted by its initial node (whose zone will be always denoted by  $i$ ) and the final node (denoted by  $j$  if different from  $i$ ). Note that if the destination is the own subcenter or the CBD, there is only one possible route. If the destination is a foreign subcenter, when the route goes through the CBD it is marked with an  $H$ ; if it goes across the subcenters ring is marked with an  $F$ . For example  $LP_iCBD$  is a line that starts in a periphery, stops in the subcenter from the same zone and finish its tour in the CBD. In this case there is no need to specify  $H$  or  $F$  because there is only one route.  $LSC_iSC_jH$  is a route that starts in a subcenter and goes to another, but stopping previously in the CBD.

To apply this heuristic to our scheme, let us note that the weight of each arc depends only on the types of nodes that it links, e.g.  $w$  of  $P_i - SC_i$  is  $a$  or  $w$  of  $SC_i - CBD$  is  $a\left(\alpha + \frac{3}{7}\gamma\right) + b\left(\tilde{\alpha} + \frac{3}{7}\tilde{\gamma}\right)$ . Depending on the values of the parameters  $a, \alpha$  and  $\beta$ , the largest weight of an arc will be either  $P_i - SC_i$  or  $SC_i - CBD$ . As all the arcs are incident in a center, it would be impossible to have lines that tour more than one arc, so we do not stop when arriving at another center.

First case:  $a > a\left(\alpha + \frac{3}{7}\gamma\right) + b\left(\tilde{\alpha} + \frac{3}{7}\tilde{\gamma}\right)$

The first line added is  $LP_i - SC_i$ ; then, we start with an arc  $SC_i - CBD$  and is then extended to  $CBD - SC_j$ , so the line is  $LSC_iSC_jH$  (we assume that the final subcenter is the opposite to the initial subcenter); finally, the last type starts again with arc  $SC_i - CBD$  but is then extended to  $SC_i - SC_{i+1}$ : this line is denoted  $LSC_iCBDb$ .

As a result, when finding optimal frequencies the line  $LSC_iCBDb$  will always present a null frequency. This means that the structure is a mixture between the feeder-trunk structure (because all the passengers from the periphery take the feeder bus to their subcenter) and the hub and spoke structure (because the CBD will be a hub where almost all the passengers that go to a foreign subcenter will transfer).

Second case:  $a < a\left(\alpha + \frac{3}{7}\gamma\right) + b\left(\tilde{\alpha} + \frac{3}{7}\tilde{\gamma}\right)$

In this case the first line starts with an arc  $SC_i - CBD$ , and it extends to  $P_i - SC_i$ , so the line is  $LP_iCBD$ ; the second type starts again with arc  $SC_i - CBD$  and it extends to

$CBD - SC_j$ , so the line is  $LSC_iSC_jH$  (as in the first case, we assume that the final subcenter is the opposite to the initial subcenter); finally, the last type of line starts with arc  $SC_i - SC_{i+1}$ , it extends to  $SC_{i+1} - SC_{i+2}$  and so on, so finally the line is the circular line presented in some previous structures.