Thinking about this as an integer program, let x_{ij} be a 0,1 variable, equal to 1 if a link exists between nodes i and j. (For $1 \leq i \leq n$ and $1 \leq j \leq i-1$).

Then the set of all possible trees is given by the constraints:

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1} x_{ij} = n-1$$

$$\sum_{i=1}^{n} x_{ij} + x_{ji} = 1 \forall j \in leafnodes$$

$$\sum_{i=1}^{n} x_{ij} + x_{ji} \geq 0 \forall j \in hubnodes$$

Note that there are no more than

$$\frac{\frac{(n-1)n}{2}!}{(n-1)!^2}$$

possible trees for n nodes (this is not a tight upper bound: for 4 nodes, there are 16 possible trees, and the above formula gives an upper bound of 20).

I have a sort of formula for calculating the cost function to minimise within this set of 0,1 IP constraints. However, not only is this margin too small to contain it, it's pretty crapulent to use or calculate. I need to work on this some more, to reformulate that function so that it can be used in an IP/LP solution.

– hugo