Operations Research Letters 9 (1990) 127–132
North-Holland

March 1990

CLASSIFICATION OF TRAVELLING SALESMAN PROBLEM FORMULATIONS

André LANGEVIN and François SOUMIS
École Polytechnique, Montréal, Canada

Jacques DESROSIERS
École des Hautes Études Commerciales, Montréal, Canada

Received November 1988
Revised August 1989

Several single-commodity, two-commodity and multi-commodity flow formulations have recently been introduced for the travelling salesman problem. The purpose of this paper is to clarify the relations between these formulations and with other classical formulations. Some results are probably known by researchers in the area. However they have not yet been published. This paper groups them together.

travelling salesman problem

1. Classical formulations

The classical formulation of the travelling salesman problem consists in modeling it as an assignment problem (1)–(4) with integrality constraints (5) and subtour breaking constraints:

$$\text{min.} \quad \sum_i \sum_j c_{ij} x_{ij} \tag{1}$$

$$\text{s.t.} \quad \sum_i x_{ij} = 1, \quad j = 1, \ldots, n, \tag{2}$$

$$\sum_j x_{ij} = 1, \quad i = 1, \ldots, n, \tag{3}$$

$$0 \leqslant x_{ij} \leqslant 1 \quad \text{for all } i,\ j, \tag{4}$$

$$x_{ij} \quad \text{integer for all } i,\ j, \tag{5}$$

$$+\ \text{subtour breaking constraints.}$$

The coefficient c_{ij} represents the cost (or the distance) to travel from node i to node j. We state $c_{ii} = \infty$ for all i. The variable x_{ij} takes the value 1 if the route goes along arc (i, j) and 0 otherwise. The equations (2) and (3) form the assignment constraints and ensure that each node is visited only once. Dantzig, Fulkerson and Johnson [2] use these $O(n^2)$ binary variables and $O(n^2)$ constraints with the following subtour breaking constraints:

$$\sum_{i \in Q} \sum_{j \in Q} x_{ij} \leqslant |Q| - 1 \quad \text{for all } Q \subseteq \{1,2,\ldots,n\} \quad \text{and} \quad 2 \leqslant |Q| \leqslant n-1. \tag{6}$$

Miller, Tucker and Zemlin [11] present another formulation based on the assignment model but using a polynomial number of subtour breaking constraints by means of $O(n^2)$ supplementary continuous variables. These constraints are:

$$u_i - u_j + n x_{ij} \leqslant n - 1 \quad \text{for all } j \neq 1 \quad \text{and} \quad i \neq j. \tag{7}$$

Finally, Gavish and Graves [5] propose a formulation using $O(n^2)$ binary variables x_{ij}, $O(n^2)$ continuous variables z_{ij} and n^2+3n constraints. The variables z_{ij} describe the flow of a single commodity to node one from every other nodes. The subtour breaking constraints are given by

$$\sum_j z_{ij} - \sum_{j \neq 1} z_{ji} = 1, \quad i = 2,\ldots,n, \tag{8}$$

$$z_{ij} \leqslant (n-1)x_{ij}, \quad i=2,\ldots,n, \quad j=1,\ldots,n, \tag{9}$$

$$z_{ij} \geqslant 0 \quad \text{for all } i,\ j. \tag{10}$$

Let $v_{\mathrm{LP}}(\mathrm{DFJ})$, $v_{\mathrm{LP}}(\mathrm{MTZ})$ and $v_{\mathrm{LP}}(\mathrm{GG})$ denote the optimal value of the linear programming relaxations for the Dantzig, Fulkerson and Johnson [2], Miller, Tucker and Zemlin [11], and Gavish and Graves [5] formulations respectively. Wong [12] demonstrates that $v_{\mathrm{LP}}(\mathrm{MTZ}) \leqslant v_{\mathrm{LP}}(\mathrm{GG}) \leqslant v_{\mathrm{LP}}(\mathrm{DFJ})$ and notes that the difference between $v_{\mathrm{LP}}(\mathrm{GG})$ and $v_{\mathrm{LP}}(\mathrm{DFJ})$ can be very large. Held and Karp [7] for their part experimentally observe that, for many problems, $v_{\mathrm{LP}}(\mathrm{DFJ})$ is very close (within 0.5%) to the optimal solution of the travelling salesman problem.

2. Multi-commodity flow formulations

Wong [12] presents a formulation for the TSP using the flow of $2(n-1)$ commodities, $Y^k = (y_{ij}^k)$, $k=2,\ldots,n$, and $Z^k = (z_{ij}^k)$, $k=2,\ldots,n$:

$$\text{(Wong)} \quad \min. \quad \sum_i \sum_j c_{ij} x_{ij}$$

$$\text{s.t.} \quad (2),\ (3),\ (4),\ (5),$$

$$\sum_j \left(y_{ij}^k - y_{ji}^k \right) = \begin{cases} 1 & \text{if } i=1, \\ -1 & \text{if } i=k, \\ 0 & \text{if } i \neq 1 \text{ and } k, \end{cases} \quad k=2,\ldots,n, \tag{11}$$

$$\sum_j \left(z_{ij}^k - z_{ji}^k \right) = \begin{cases} -1 & \text{if } i=1, \\ 1 & \text{if } i=k, \\ 0 & \text{if } i \neq 1 \text{ and } k, \end{cases} \quad k=2,\ldots,n, \tag{12}$$

$$y_{ij}^k \leqslant x_{ij}, \quad z_{ij}^k \leqslant x_{ij} \qquad \text{for all } i,\ j,\ k, \tag{13}$$

$$y_{ij}^k \geqslant 0, \quad z_{ij}^k \geqslant 0 \qquad \text{for all } i,\ j,\ k. \tag{14}$$

Constraints (11) and (12) ensure that a unit of commodity Y^k travels from node 1 (source of commodity Y^k) to node k (the sink of commodity Y^k) while one unit of commodity Z^k travels from node k to node 1. This implies that the subgraph supporting the flow is strongly connected, i.e., that there is on this subgraph a path going from any node to any other node (passing through the source). This formulation requires $O(n^2)$ binary variables, $O(n^3)$ continuous variables and $2(n^3+n^2+n)$ constraints. Only the x_{ij} variables are subject to integrality since the y_{ij}^k and z_{ij}^k variables are necessarily integer whenever the x_{ij}'s are, because the flow matrix for commodities Y^k and Z^k is totally unimodular (the x_{ij}'s correspond to the capacities on the arcs). Wong [12] demonstrates that

$$v_{\mathrm{LP}}(\mathrm{Wong}) = v_{\mathrm{LP}}(\mathrm{DFJ}). \tag{15}$$

The proof is based on two ideas. First, in presence of constraints (2) and (3), subtour breaking constraints (6) can be rewritten as

$$\sum_{i \in Q} \sum_{j \in \overline{Q}} x_{ij} \geqslant 1 \quad \text{for all } Q \subseteq \{1, 2, \ldots, n\}. \tag{6'}$$

Second, considering x_{ij} as an arc capacity, constraints (6′) on cut capacities are equivalent (by max flow–min cut theorem) to the existence of flows Y^k and Z^k carrying respectively one unit from node 1 to node k and from node k to node 1 conversely.

Since the number of subtour breaking constraints grows exponentially in the Dantzig, Fulkerson, and Johnson formulation but only polynomially in the Wong formulation, the potential for problem simplification is very significant. The Wong formulation, however, uses an additional number of variables.

Claus [1] proposes a $(n-1)$-commodity formulation that relaxes Wong's formulation by eliminating the z_{ij}^k variables and the constraints in which they appear. Langevin [8] introduces a formulation that replaces, for each arc (i, j), the two inequalities (13) in Wong's formulation by only one more restrictive constraint $y_{ij}^k + z_{ij}^k \leqslant x_{ij}$ for all i, j and k. Loulou [9] considers a formulation with even more restrictive constraints: $y_{ij}^k + z_{ij}^k = x_{ij}$ for all i, j and k. Obviously we have

$$v_{\mathrm{LP}}(\text{Claus}) \leqslant v_{\mathrm{LP}}(\text{Wong}) \leqslant v_{\mathrm{LP}}(\text{Langevin}) \leqslant v_{\mathrm{LP}}(\text{Loulou}). \tag{16}$$

We demonstrate by the following proposition that the LP relaxations of the $(n-1)$- and $2(n-1)$-commodity formulations considered by Claus, Langevin, and Loulou give the same bound as the LP relaxation of the Dantzig, Fulkerson and Johnson formulation. However, these LP formulation cannot be used to directly solve TSPs of substantial size because of the large number of constraints $[O(n^3)]$.

Proposition 1.

$$v_{\mathrm{LP}}(\text{DFJ}) = v_{\mathrm{LP}}(\text{Wong}) = v_{\mathrm{LP}}(\text{Claus}) = v_{\mathrm{LP}}(\text{Langevin}) = v_{\mathrm{LP}}(\text{Loulou}).$$

Proof. Let $X = (x_{ij})$ and $Y^k = (y_{ij}^k)$, $k = 2,\dots,n$, be a solution satisfying Claus' formulation. Since $y_{ij}^k \leqslant x_{ij}$, define $z_{ij}^k = x_{ij} - y_{ij}^k \geqslant 0$. Since $\Sigma_j x_{ij} = \Sigma_j x_{ji} = 1$, a solution of Loulou's formulation is easily obtained by substituting in (11) the variables y_{ij}^k:

$$\begin{aligned}
\sum_j y_{ij}^k - \sum_j y_{ji}^k &= \sum_j \left(x_{ij} - z_{ij}^k\right) - \sum_j \left(x_{ji} - z_{ji}^k\right) \\
&= \sum_j x_{ij} - \sum_j x_{ji} - \sum_j \left(z_{ij}^k - z_{ji}^k\right) \\
&= -\sum_j \left(z_{ij}^k - z_{ji}^k\right).
\end{aligned}$$

Thus, $Z^k, k = 2,\dots, n$, satisfy (12), and hence v_{LP} (Loulou) $\leqslant v_{\mathrm{LP}}$(Claus). By relations (15) and (16), the proof is complete. □

3. Two-commodity flow formulations

Finke, Claus and Gunn [3] propose a two-commodity flow formulation for the TSP. Consider that the travelling salesman leaves node 1 with $n-1$ units of commodity Y and none of commodity Z. He distributes one unit of Y and picks up one unit of Z at each node so that the salesman travels at all times with a total of $n-1$ units. The model is written as follows:

$$\text{min.} \quad \sum_i \sum_j c_{ij}\left(y_{ij} + z_{ij}\right)/(n-1) \tag{17}$$

$$\text{s.t.} \quad \sum_j \left(y_{ij} - y_{ji}\right) = \begin{cases} n-1 & \text{for } i = 1, \\ 1 & \text{for } i \neq 1, \end{cases} \tag{18}$$

$$\sum_j \left(z_{ij} - z_{ji}\right) = \begin{cases} -(n-1) & \text{for } i = 1, \\ 1 & \text{for } i \neq 1, \end{cases} \tag{19}$$

$$\sum_j (y_{ij} + z_{ij}) = n - 1 \quad \text{for all } i, \tag{20}$$

$$y_{ij} + z_{ij} \in \{0, n-1\} \quad \text{for all, } i, j, \tag{21}$$

$$y_{ij} \geqslant 0, \quad z_{ij} \geqslant 0 \quad \text{for all } i, j. \tag{22}$$

Constraints (18) and (19) define the flow conservation equations for each commodity. Constraints (20) and (21) ensure that there is exactly one arc supporting a combined flow of $n-1$ units out of each node. The supply and demand structure implies that there is a path from source 1 to each node j and another one from each node j to 1. Therefore, the supporting graph is strongly connected as in the case of the multi-commodity flow formulations. The LP relaxation of the preceding formulation is obtained by replacing constraints (21) by $y_{ij} + z_{ij} \leqslant n-1$ for all i, j. These constraints are redundant because of (20) and can be eliminated from the LP relaxation. The corresponding LP, noted FCG, contains then $3n$ constraints whereas the integer formulation contains $O(n^2)$ additional constraints.

Lucena's formulation [10] generalizes Finke, Claus and Gunn formulation. The travelling salesman starts at node 1 and distributes d_i units of commodity Y and picks up d_i units of commodity Z at node i, $i = 2,\ldots,n$. In relations (17)–(22), d_i replaces value 1 for $i \neq 1$ while $d_2 + \cdots + d_n$ replaces value $n-1$ elsewhere. This two-commodity flow formulation has the same numbers of variables and constraints as the Finke, Claus and Gunn formulation.

4. Relations between two-commodity and multi-commodity flow formulations

It is commonly supposed that the two-commodity formulation from Finke, Claus and Gunn is weaker than the multi-commodity formulations from Wong or Claus. We establish that this is the case by showing that Finke, Claus and Gunn formulation is obtained by aggregating Loulou's formulation. Constraints (18) and (19) are respectively obtained by summing the groups of equations (11) and (12) over k and by aggregating the commodities Y^k and Z^k such that

$$y_{ij} = \sum_{k=2}^{n} y_{ij}^k \quad \text{and} \quad z_{ij} = \sum_{k=2}^{n} z_{ij}^k. \tag{23}$$

Furthermore, by summing the equations $y_{ij}^k + z_{ij}^k = x_{ij}$ over k, we have $y_{ij} + z_{ij} = (n-1)x_{ij}$, from which we obtain (21) since x_{ij} is binary. Constraints (20) are obtained by summing over j the preceding equation and by using the assignment equations.

Lucena's constraints are also obtained from Loulou's formulation by aggregating the constraints (11) and (12) with the weights d_k and by defining the flows by

$$y_{ij} = \sum_{k=2}^{n} d_k y_{ij}^k \quad \text{and} \quad z_{ij} = \sum_{k=2}^{n} d_k z_{ij}^k. \tag{24}$$

Lucena's linear bound is clearly weaker than the multi-commodity bound. However it can be weaker or stronger than the Finke, Claus and Gunn bound, depending on the problem instances. Finally, Finke, Claus and Gunn [3] demonstrate that the bound obtained from their formulation is greater than or equal to the solution of the assignment problem and they give an example where their bound is strictly better than that of the corresponding assignment problem. Furthermore, Loulou [9] finds an example for which Finke, Claus and Gunn bound is smaller than that of Wong.

5. Relations between single-commodity and two-commodity flow formulations

We demonstrate by the following proposition the equivalence between the LP relaxations of the formulations of Gavish and Graves, and of Finke, Claus and Gunn. The GG formulation contains $O(n^2)$ constraints while the FCG formulation constraints $O(n)$ constraints only.

Proposition 2.

$$v_{\mathrm{LP}}(\mathrm{GG}) = v_{\mathrm{LP}}(\mathrm{FCG}).$$

Proof. Gavish and Graves solutions can be constructed from Finke, Claus and Gunn solutions by using relation $x_{ij} = (y_{ij} + z_{ij})/(n-1)$. The reciprocal result can be obtained by using the inverse relation. □

Finke, Claus and Gunn formulation is also very interesting since it is well suited to modelling TSP's with additional technological constraints. For example, Finke and Kusiak [4] use this model in the area of flexible manufacturing systems and Lucena [10] in the area of vehicle routing.

6. Conclusion

The paper presents relations between various LP relaxations of TSP formulations (Figure 1). We have shown the equivalence of the LP relaxation of the $(n-1)$- and $2(n-1)$-commodity flow formulations containing $O(n^2)$ constraints with that of the formulation of Dantzig, Fulkerson and Johnson [2] that contains $O(2^n)$ constraints. This allows using the formulation which is easiest to solve according to the type of approach since the corresponding bounds are all equivalent.

We have also shown the equivalence of the 1- and 2-commodity formulations as LPs, that is the formulation by Gavish and Graves [5] containing $O(n^2)$ constraints and that of Finke, Claus and Gunn [3] containing $3n$ constraints. Even though $(n-1)$- and $2(n-1)$-commodity formulations produce tighter

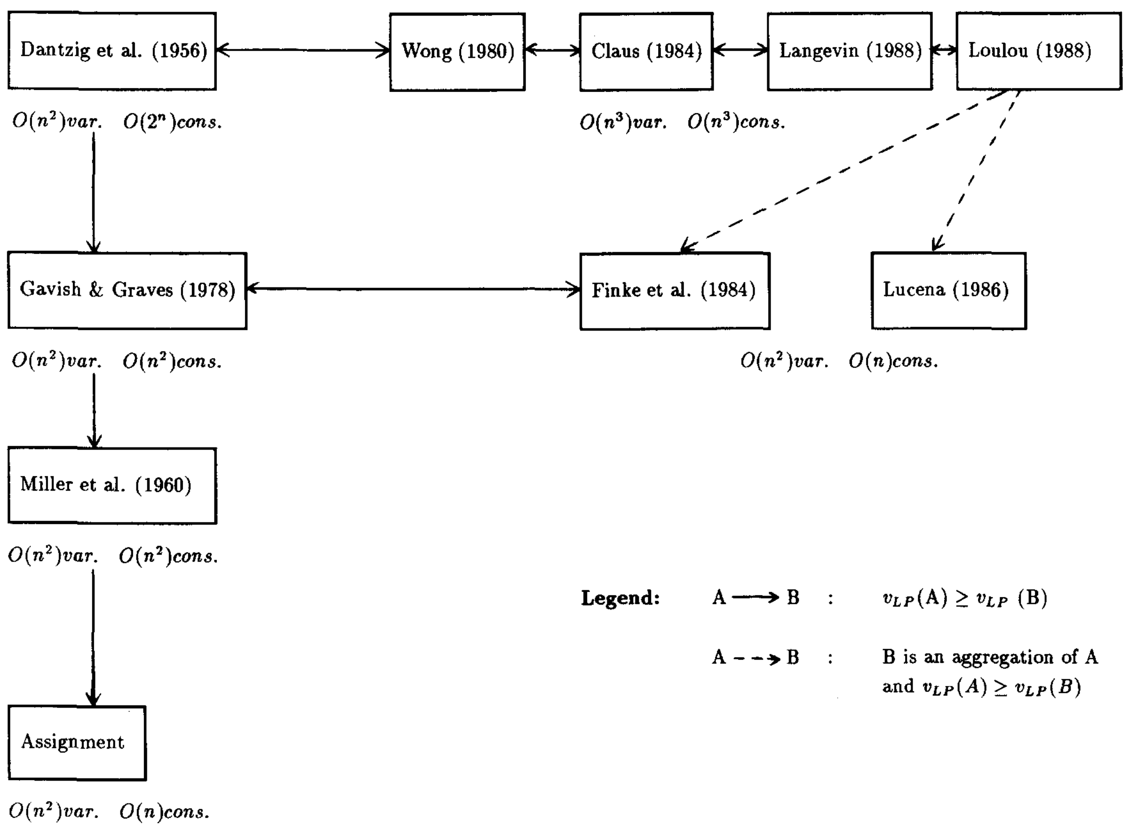

Fig. 1. Classification of TSP relaxations.

bounds than 1- and 2-commodity formulations, the experimentation of Gavish and Srikanth [6] on the Lagrangian relaxations gives us clues on what could be done with the formulation of Finke, Claus and Gunn. Hence different sizes of the formulations as LPs may lead to computational advantages.

Acknowledgements

This research was supported by the Federal Government of Canada (NSERC Program) and by the Quebec Government (FCAR Program).

References

[1] A. Claus, "A new formulation for the travelling salesman problem", *SIAM J. Algebraic Discrete Methods* **5**, 21–25 (1984).
[2] G.B. Dantzig, D.R. Fulkerson and S.M. Johnson, "Solutions of large scale travelling salesman problem", *Oper. Res.* **2**, 393–410 (1954).
[3] G. Finke, A. Claus and E. Gunn, "A two-commodity network flow approach to the travelling salesman problem", *Congr. Numer.* **41**, 167–178 (1984).
[4] G. Finke and A. Kusiak, "Network approach to modelling of flexible manufacturing modules and cells", *RAIRO* **19**, 359–370 (1985).
[5] B. Gavish and S.C. Graves, "The travelling salesman problem and related problems", Working paper GR-078-78, Operations Research Center, Massachusetts Institute of Technology, 1978.
[6] B. Gavish and K. Srikanth, "An optimal solution method for large-scale multiple travelling salesman problems", *Oper. Res.* **34**, 698–717.
[7] M. Held and R.M. Karp, "The travelling salesman problem and minimum spanning trees II", *Math. Programming* **1**, 6–25 (1971).
[8] A. Langevin, "Planification de tournées de véhicules", PhD thesis, Dept. of Applied Mathematics, École Polytechnique, Montréal, Canada, H3C 3A7, 1988.
[9] R.J. Loulou, "On multicommodity flow formulation for the TSP", Working paper, McGill University, Montreal, 1988.
[10] A. Lucena, "Exact solution approaches for the vehicle routing problem", PhD thesis, Imperial College of Science and Technology, University of London, 1986.
[11] C.E. Miller, A.W. Tucker and R.A. Zemlin, "Integer programming formulation of travelling salesman problems", *J. ACM* **7**, 326–329 (1960).
[12] R.T. Wong, "Integer programming formulations of the travelling salesman problem", *Proceedings of the IEEE International Conference of Circuits and Computers*, 149–152 (1980).