The Complexity of the Network Design Problem

D. S. Johnson
Bell Laboratories
Murray Hill, New Jersey

J. K. Lenstra
Mathematisch Centrum
Amsterdam, The Netherlands

A. H. G. Rinnooy Kan
Erasmus University
Rotterdam, The Netherlands

ABSTRACT

In the network design problem we are given a weighted undirected graph. We wish to find a subgraph which connects all the original vertices and minimizes the sum of the shortest path weights between all vertex pairs, subject to a budget constraint on the sum of its edge weights. In this note we establish NP-completeness for the network design problem, even for the simple case where all edge weights are equal and the budget restricts the choice to spanning trees. This result justifies the development of enumerative optimization methods and of approximation algorithms, such as those described in a recent paper by R. Dionne and M. Florian.

INTRODUCTION

In the network design problem we are given a weighted undirected graph. We wish to find a subgraph which connects all the original vertices and minimizes the sum of the shortest path weights between all vertex pairs, subject to a budget constraint on the sum of its edge weights. In this note we establish NP-completeness [7,8] for the network design problem. Briefly, this result implies that a polynomial-bounded method for its solution could be used to construct similar algorithms for a large number of combinatorial problems which are notorious for their computational intractability, such as the travelling salesman problem and the multicommodity network flow problem. Since none of these problems are known to be solvable in polynomial time, NP-completeness of the network design problem justifies the development of enumerative optimization methods and of approximation algorithms, such as those described by R. Dionne and M. Florian [2].

Networks, Vol. 8 (1978) 279-285 0028-3045/78/0008-0279$01.00

For our purposes, we formulate the problem in the following way.

NETWORK DESIGN PROBLEM (NDP): Given an undirected graph $G = (V,E)$, a weight function $L: E \to \mathbb{N}$, a budget B and a criterion threshold C $(B,C \in \mathbb{N})$, does there exist a subgraph $G' = (V,E')$ of G with weight $\sum_{\{i,j\}\in E'} L(\{i,j\}) \leq B$ and criterion value $F(G') \leq C$, where $F(G')$ denotes the sum of the weights of the shortest paths in G' between all vertex pairs?

By way of introduction to a quite involved NP-completeness proof for a simplified version of NDP, we shall first present a simple proof establishing NP-completeness for the general NDP.

Theorem 1: NDP is NP-complete.

Proof: Consider the following problem.

KNAPSACK: Given positive integers $t,a_1,\ldots,a_t,b$, does there exist a subset $S \subset T = \{1,\ldots,t\}$ such that $\sum_{i\in S} a_i = b$?

We will show that KNAPSACK is reducible to NDP, *i.e.*, that for any instance of KNAPSACK an instance of NDP can be constructed in polynomial-bounded time such that solving the instance of NDP solves the instance of KNAPSACK as well. The theorem then follows from the NP-completeness of KNAPSACK [7] and the fact that NDP belongs to NP, since any feasible subgraph can be recognized as such in polynomial time.

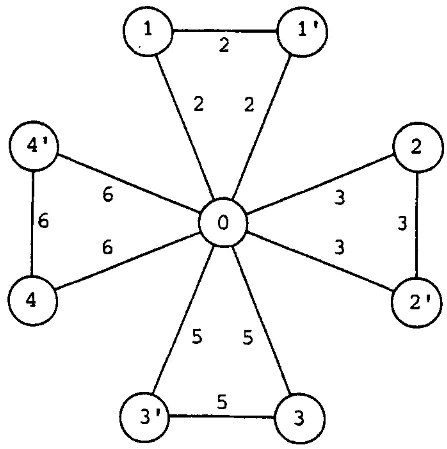
Given any instance of KNAPSACK, we write $A = \sum_{i\in T} a_i$ and define an instance of NDP as follows:

$$V = \{0\} \cup \{i,i': i \in T\},$$
$$E = \{\{0,i\},\{0,i'\},\{i,i'\}: i \in T\},$$
$$L(\{0,i\}) = L(\{0,i'\}) = L(\{i,i'\}) = a_i \quad (i \in T),$$
$$B = 2A+b,$$
$$C = 4tA-b.$$

Figure 1 illustrates this reduction. We claim that KNAPSACK has a solution if and only if $G = (V,E)$ contains a subgraph with weight at most B and criterion value at most C.

KNAPSACK: $t = 4$, $a_1 = 2$, $a_2 = 3$, $a_3 = 5$, $a_4 = 6$, $b = 7$.

NDP: G & L:



$B = 39$, $C = 249$.

Fig. 1 Equivalent instances of KNAPSACK and NDP.

It is easily seen that any feasible NDP solution can be assumed to contain a star graph $G^* = (V, \{\{0,i\},\{0,i'\}: i \in T\})$; G^* has weight $2A = B-b$ and criterion value $4tA = C+b$. Adding an edge $\{i,i'\}$ to G^* increases the weight by a_i and decreases the criterion value by a_i, since $\{i,i'\}$ will appear only in the shortest path between i and i'. The equivalence now follows in a straightforward way. □

However, since KNAPSACK can be solved in $O(tb)$ time [1], Theorem 1 does not exclude the existence of a similar *pseudo-polynomial* algorithm [4] for NDP; the above construction crucially depends on allowing arbitrary positive integers as edge weights and budget. As a stronger result, we shall now prove that NDP is NP-complete even in the simple case where all edge weights are equal and the budget restricts the choice to spanning trees.

SIMPLE NETWORK DESIGN PROBLEM (SNDP): NDP with $L(\{i,j\}) = 1$ for all $\{i,j\} \in E$ and $B = |V|-1$.

Theorem 2: SNDP is NP-complete.

Proof: As a starting point we take the following NP-complete problem [3,6].

EXACT 3-COVER: Given a family $S = \{\sigma_1,\ldots,\sigma_s\}$ of 3-element subsets of a set $T = \{\tau_1,\ldots,\tau_{3t}\}$, does there exist a subfamily $S' \subset S$ of pairwise disjoint sets such that $\bigcup_{\sigma\in S'} \sigma = T$?

We will show that EXACT 3-COVER is reducible to SNDP.

Given any instance of EXACT 3-COVER, we define an instance of SNDP as follows:

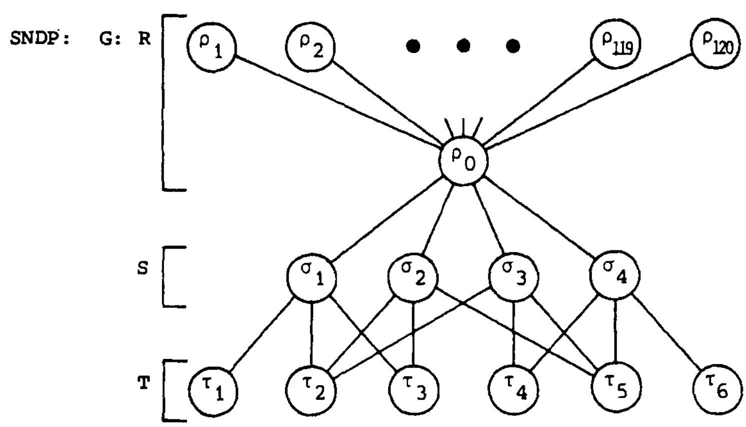
$$V = R \cup S \cup T,$$
$$R = \{\rho_0,\rho_1,\ldots,\rho_r\},$$
$$r = C_{SS} + C_{ST} + C_{TT},$$
$$E = \{\{\rho_i,\rho_0\}\colon i=1,\ldots,r\} \cup \{\{\rho_0,\sigma\}\colon \sigma\in S\} \cup \{\{\sigma,\tau\}\colon \tau\in\sigma\in S\},$$
$$C = C_{RR} + C_{RS} + C_{RT} + C_{SS} + C_{ST} + C_{TT},$$

where $C_{RR} = r^2$, $C_{RS} = 2rs+s$, $C_{RT} = 9rt+6t$, $C_{SS} = s^2-s$, $C_{ST} = 9st-6t$, $C_{TT} = 18t^2-12t$.

Figure 2 illustrates this reduction. We will prove that EXACT 3-COVER has a solution if and only if $G = (V,E)$ contains a spanning tree with criterion value at most C. We assume that G is connected, *i.e.*, $\bigcup_{\sigma\in S} \sigma = T$.

EXACT 3-COVER: $t = 2$, $s = 4$,

$S = \{\{\tau_1,\tau_2,\tau_3\},\{\tau_2,\tau_3,\tau_5\},\{\tau_2,\tau_4,\tau_5\},\{\tau_4,\tau_5,\tau_6\}\}$.



$C = 17656$.

Fig. 2 Equivalent instances of EXACT 3-COVER and SNDP.

Let $G' = (V,E')$ be some spanning tree of G and let $F_{PQ}(G')$ denote the sum of the weights of all shortest paths in G' between vertex sets P and Q ($P,Q \subset V$). We clearly have $\{\rho_i,\rho_0\} \in E'$ for all $i = 1,\ldots,r$. If $\{\rho_0,\sigma\} \notin E'$ for some $\sigma \in S$, then

$$\begin{aligned} F(G') &> F_{RR}(G') + F_{RS}(G') + F_{RT}(G') \\ &\geq C_{RR} + C_{RS} + 2(r+1) + C_{RT} \\ &> C; \end{aligned}$$

therefore, we may assume that $\{\rho_0,\sigma\} \in E'$ for all $\sigma \in S$. It follows that in G' each vertex in T is adjacent to exactly one vertex in S. Straightforward calculations show that we now have

$$F_{PQ}(G') = C_{PQ} \quad \text{for } P = R,S \text{ and } Q = R,S,T.$$

Denoting the number of vertices in S being adjacent in G' to exactly h vertices in T by s_h $(h = 0,1,2,3)$, we have

$$\begin{aligned} F_{TT}(G') &= 4(3t(3t-1)/2) \\ &\quad - 2\left|\{\{\tau,\tau'\}\colon \tau \neq \tau',\ \{\{\sigma,\tau\},\{\sigma,\tau'\}\} \subset E' \text{ for some } \sigma \in S\}\right| \\ &= (18t^2-6t) - (2s_2+6s_3) \\ &= C_{TT} + 6(t-s_3) - 2s_2. \end{aligned}$$

It is easily seen that $F_{TT}(G') = C_{TT}$ if and only if $s_3 = t$, $s_2 = s_1 = 0$, $s_0 = s-t$. The first condition is now equivalent to $F(G') \leq C$, the second one to the existence of an EXACT 3-COVER solution. This completes the proof. □

Various related types of network design problems have been discussed in the literature; an excellent survey has been given by R. T. Wong [10]. For instance, the problems dealt with by A. J. Scott [9] are generalizations of NDP and hence NP-complete.

Another variation has been introduced by T. C. Hu [5]. Given a complete graph with a distance and a requirement for each vertex pair, we wish to find a spanning tree which minimizes the total cost of communication, where the cost of communication for a pair of vertices is the distance of the path between them multiplied by their requirement. The case where all distances are equal can be solved by polynomial time [5]; for the case where all requirements are equal, NP-completeness follows easily as a corollary to Theorem 2.

Finally, we mention some results with respect to the complexity of a network design problem due to F. Maffioli. Given a weighted graph with a specific vertex ρ and an integer k, we wish to find a spanning tree of minimum total weight subject to the constraint that each subtree incident with ρ contains at most k other vertices. The case $k = 2$ can be formulated as a matching problem; the case $k = 3$ can be proved NP-complete by a reduction of 3-DIMENSIONAL MATCHING, even if all edge weights are equal.

REFERENCES

1. Bellman, R. E. and S. E. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, Princeton, New Jersey, 1962.

2. Dionne, R. and M. Florian, "Exact and Approximate Algorithms for Optimal Network Design," to appear in *Networks*, 9, 1, 1979.

3. Garey, M. R., R. L. Graham and D. S. Johnson, "Some NP-Complete Geometric Problems," *Proc. 8th Annual ACM Symp. Theorey Comput.*, 1976, pp. 10-22.

4. Garey, M. R. and D. S. Johnson, "Strong NP-Completeness Results: Motivation, Examples and Implications," *J. Assoc. Comput. Mach.*, 25, 1978, pp. 499-508.

5. Hu, T. C., "Optimum Communication Spanning Trees," *SIAM J. Comput.*, 3, 1974, pp. 188-195.

6. Hyafil, L. and R. L. Rivest, "Constructing Optimal Binary Decision Trees is NP-Complete," *Information Processing Letters*, 5, 1976, pp. 15-17.

7. Karp, R. M., "Reducibility Among Combinatorial Problems," in *Complexity of Computer Computations*, R. E. Miller and J. W. Thatcher, eds., Plenum Press, New York, 1972, pp. 85-103.

8. Karp, R. M., "On the Computational Complexity of Combinatorial Problems," *Networks*, 5, 1975, pp. 45-68.

9. Scott, A. J., "The Optimal Network Problem: Some Computational Procedures," *Trans. Res.*, 3, 1969, pp. 201-210.

10. Wong, R. T., "A Survey of Network Design Problems," Working Paper OR 053-76, Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1976.

Received February 1977; Revised July 1977.