Benders Decomposition for Network Design Problems with Underlying Tree Structure*

Abdelhamid Benchakroun[1] *Jacques A. Ferland*[2]
Viviane Gascon[3]

[1]Département de mathématiques et d'informatique
Université de Sherbrooke

[2]Département d'Informatique et de Recherche Opérationnelle
Université de Montréal

[3]Département de Sciences de la Gestion et de l'Économie
Université du Québec à Trois-Rivières

Abstract
We present a class of network design problems with underlying tree structure. The problem is formulated as a mixed nonlinear programming model including investment integer variables associated with the equipments to be installed and continuous variables associated with the use of the network. The generalized Benders decomposition method is used to solve it. In this article, we introduce simple procedures to determine the optimal multipliers used to specify the cuts in the master problem. The procedures do not require to solve any optimization problem as they take advantage of the underlying tree structure of the problem.

1 Introduction

Several practical problems in the field of computer communications [8], telecommunications [16], and electricity distribution [1, 2, 3, 6, 7] can be formulated as network design problems where both the topology of the network and the size of the equipments to use have to be determined.

The model is formulated using a directed graph $G = (N, A)$. The set N of nodes includes a unique source node s; all the other nodes $j \in N, j \neq s$, are demand nodes

*The first author was supported in part by NSERC grant OGP036512. The second author was supported in part by NSERC grant OGP008312 and FCAR grant 93-ER-1654. The third author was supported in part by NSERC grant OGP0153137.

requiring d_j units. (Note that any problem with several source nodes can be transformed into a model with a unique super source node using additional arcs from the super source to the source nodes of the original problem.) The arcs in A constitute the set of potential locations for feeder lines. Whenever several equipments of different capacity are available for a line, then different arcs are associated with the different equipments. Moreover, if a feeder line can be used in both directions, then the arcs associated with it are duplicated in each direction.

We assume that the topology of the solution network has to be a spanning directed tree rooted at the source node s. This is a realistic constraint for a large number of practical applications. In the problem formulation, with each arc $a \in A$ is associated a real valued flow variable x_a

$$x_a = \text{ flow in arc } a \in A \ ,$$

and a decision variable y_a to specify the topology

$$y_a = \begin{cases} 1 & \text{if arc } a \text{ is constructed} \\ 0 & \text{otherwise .} \end{cases}$$

Denote by x and y the vector of flow variables x_a and the vector of decision variables y_a, respectively. The model is summarized as follows:

$$\begin{array}{lllr}
\text{Min} & f_1(x) + f_2(y) + f_3(x,y) & & \\
\text{Subject to:} & \displaystyle\sum_{a\in\Gamma_j^-} x_a \ - \sum_{a\in\Gamma_j^+} x_a = d_j & j \in N, j \neq s & (1.1) \\
& x_a \ - \ g_a(y_a) \leq 0 & a \in A & (1.2) \\
& x_a \geq 0 & a \in A & \\
& y \in Y & &
\end{array}$$

where the constraint $y \in Y$ restricts the network structure to be a spanning directed tree rooted a the source node s, and Γ_j^- and Γ_j^+ denote the set of arcs entering node j and leaving node j, respectively. Furthermore

$$g_a(y_a) = \begin{cases} 0 & \text{if } y_a = 0 \\ U_a & \text{if } y_a = 1 \end{cases}$$

where $U_a > 0$ is the capacity of the equipment associated with arc $a, f_1(x)$ is the flow cost, $f_2(y)$ is the investment cost for the equipment, and $f_3(x,y)$ is a function specified according to the practical application. Constraints (1.1) are the usual flow conservation constraints, and constraints (1.2) are the arc capacity constraints.

Note that it is quite common practice to specify an objective function where investement and management costs are added. The general formulation in model (P) allows for any relative weights of these costs to be included. In particular, such an objective function is used in the electricity distribution planning model [3] where f_2 denotes the investment cost for the network infrastructure (variables y associated with the substations and the lines), and f_3 is the cost associated with power loss in the lines depending of the infrastructure (y) and the power flow in the lines (x). In this application, f_1 is equal to 0.

It is also worthy to note that in most applications, the mathematical models are formulated such that the convexity and the differentiability hypothesis are verified.

In [8] Gavish analyzes network design problems having underlying spanning tree structure and linear economic function. He also describes solution methods for these problems. Magnanti and Wong [11] study numerous transportation applications with underlying network design structure, and they introduce a unifying framework allowing to derive network design algorithms. Minoux in [13] is also presenting a synthesis of the models of network design problems and their solution methods.

Benders decomposition seems to be very appropriate to deal with this problem including investment and flow variables since the problem reduces to a flow problem once the investment variables are fixed. In Section 2, the solution approach is summarized. Then, in Section 3, we introduce simple procedures to determine the optimal multipliers used to specify the cuts in the master problem. The procedures do not require to solve any optimization problem as they take advantage of the underlying spanning directed tree structure of the problem. In Section 4, we give closed forms of the cuts and a solution method based on Lagrangean relaxation to deal with the master problem. Finally, some applications are given in section 5.

2 Solution Approach using Benders Decomposition Method

The model described in Section 1 can be written as follows:

$$\begin{array}{rl} \text{Min} & f_1(x) \ + \ f_2(y) \ + \ f_3(x,y) \\ \text{Subject to:} & G(x,y) \leq 0 \quad (P) \\ & x \in X \\ & y \in Y \end{array}$$

where $G(x,y)$ is the set of capacity constraints (1.2) and

$$X = \left\{ x : \sum_{a \in \Gamma_j^-} x_a \ - \sum_{a \in \Gamma_j^+} x_a = d_j, \ j \in N, \ j \neq s; \ x_a \geq 0, a \in A \right\} .$$

Assume that f_1 is convex and differentiable on X and that f_3 is also convex and differentiable in x on X for fixed values of $y \in Y$. Furthermore, $G(x, y)$ is linear in x for fixed values of $y \in Y$. To solve (P) with the generalized Benders decomposition [9], we first complete the projection of (P) onto the space of the complicating investment variables y as follows:

$$\begin{array}{rl} \text{Min} & f_2(y) \ + \ v(y) \\ \text{Subject to:} & y \in Y \cap V \end{array} \qquad (P_p)$$

where $v(y) = \inf\limits_{x \in X} \{f_1(x) + f_3(x, y) : G(x, y) \leq 0\}$ and $V = \{y : G(x, y) \leq 0$ for some $x \in X\}$.

Assuming some classical hypothesis and referring to Lagrangean duality theory [9], it is easy to see that

$$v(y) = \sup_{u \geq 0} \inf_{x \in X} \left\{f_1(x) + f_2(x, y) + u^T G(x, y)\right\}$$

and a point $y \in Y$ is also in the set V if and only if y satisfies the system

$$\inf_{x \in X} \left\{\lambda^T G(x, y)\right\} \leq 0 \qquad \lambda \in \Lambda$$

where $\Lambda = \left\{\lambda : \lambda \geq 0 \sum\limits_{a \in A} \lambda_a = 1\right\}$.

Now, by using the definition of supremum as the smallest upper bound, problem (P_p) is equivalent to:

$$\begin{array}{rll} \text{Min} & f_2(y) \ + \ y_0 & \\ \text{Subject to:} & \inf\limits_{x \in X} \left\{f_1(x) + f_3(x, y) + u^T G(x, y)\right\} \leq y_0 & u \geq 0 \\ & \inf\limits_{x \in X} \left\{\lambda^T G(x, y)\right\} \leq 0 & \lambda \in \wedge = \left\{\lambda : \lambda \geq 0, \ \sum\limits_{a \in A} \lambda_a = 1\right\} \\ & y \in Y & \end{array} \qquad (MP)$$

(MP) is denoted the master problem.

A relaxation strategy is used to deal with (MP) by solving a sequence of smaller problems denoted relaxed master problem (MPR) and specified as follows:

$$\begin{array}{rll} \text{Min} & f_2(y) \ + \ y_0 & \\ \text{Subject to:} & \inf\limits_{x \in X} \left\{f_1(x) + f_3(x, y) + u^{k^T} G(x, y)\right\} \leq y_0 & 1 \leq k \leq K \\ & \inf\limits_{x \in X} \left\{\lambda^{r^T} G(x, y)\right\} \leq 0 & 1 \leq r \leq R \\ & y \in Y & \end{array} \qquad (MPR)$$

where $u^k \geq 0, 1 \leq k \leq K$, and $\lambda^r \in \wedge, 1 \leq r \leq R$. Given an optimal solution $(\bar{y}, \bar{y}_0)$ of (MPR), the subproblem $(SP \cdot \bar{y})$ is solved:

$$\begin{array}{rl} \text{Min} & f_1(x) \ + \ f_3(x, \bar{y}) \quad (SP \cdot \bar{y}) \\ \text{Subject to:} & G(x, \bar{y}) \leq 0 \\ & x \in X \end{array}$$

If $(SP \cdot \bar{y})$ is feasible and $v(\bar{y}) \leq \bar{y}_0$, then it follows from Lagrangean duality that $(\bar{y}, \bar{y}_0)$ is an optimal solution of (MP). Otherwise, if $(SP \cdot \bar{y})$ is feasible and $v(\bar{y}) > \bar{y}_0$, then using the vector of optimal multipliers $\bar{u}$ associated with constraints $G(x, \bar{y}) \leq 0$ an additional constraint (cut of type I)

$$\inf_{x \in X} \left\{ f_1(x) + f_3(x, y) + \bar{u}^T G(x, y) \right\} \leq y_0$$

is introduced to specify a new relaxation (MPR) of (MP).

If $(SP \cdot \bar{y})$ is not feasible, then a $\bar{\lambda} \in \wedge$ such that

$$\inf_{x \in X} \left\{ \bar{\lambda}^T G(x, \bar{y}) \right\} > 0$$

is identified, and an additional constraint (cut of type II)

$$\inf_{x \in X} \left\{ \bar{\lambda}^T G(x, y) \right\} \leq 0$$

is introduced to specify a new relaxation (MPR).

Finally, referring to [9, Theorem 2.4], the finite convergence follows from the fact that Y is a finite discrete set.

3 Solving the Subproblem and Generating the Cuts

The solution approach is an iterative procedure implementing Benders decomposition. At each iteration, the current relaxed master problem (MPR) is solved first. Denote $(\bar{y}, \bar{y}_0)$ the optimal solution obtained. Next, the subproblem $(SP \cdot \bar{y})$ associated with $\bar{y}$ is solved. If $(SP \cdot \bar{y})$ is feasible and its optimal value $v(\bar{y})$ verifies $v(\bar{y}) \leq \bar{y}_0$, then the procedure stops because $(\bar{y}, \bar{y}_0)$ is optimal for (MP) . Otherwise, a new constraint (either a cut of type I or a cut of type II) is added to (MPR) to generate a new current relaxed master problem, and a new iteration is initiated.

In this section we analyse several properties of the subproblem $(SP \cdot \bar{y})$, and in the following section we introduce solution approaches to deal with (MPR) .

3.1 Solution of the Subproblem

First, note that the spanning directed tree structure in constraint $\bar{y} \in Y$ implies that if $(SP \cdot \bar{y})$ is feasible, then there is a unique feasible solution. The following proposition shows how easy it is to obtain this solution illustrated in figure 3.1. Denote $T(\bar{y}) = \{a \in A : \bar{y}_a = 1\}$, the spanning directed tree associated with $\bar{y}$.

Proposition 3.1: Let $\bar{y} \in Y$. Then the subproblem $(SP \cdot \bar{y})$ has at most one feasible solution.

Proof.: It is sufficient to note that since $\bar{y} \in Y$, the network structure associated with $\bar{y}$ is a spanning directed tree rooted at s. Hence, there exists a unique solution $\bar{x} \in X$ such that $\bar{x}_a = 0$ for all arc $a \notin T(\bar{y})$. Moreover, if $\bar{y}_a = 1$, then the flow $\bar{x}_a$ in arc a is equal to the sum of the demands of all nodes reached through arc a. This is illustrated in Figure 3.1 where $\bar{x}_{a_4} = d_3$, $\bar{x}_{a_3} = d_4$, $\bar{x}_{a_2} = d_2 + d_4$, and $\bar{x}_{a_1} = d_1 + d_2 + d_4$.

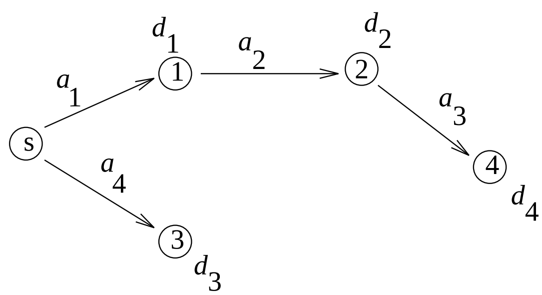

Figure 3.1

Now, if $\bar{x}_a \leq g_a(1)$ for all $a \in T(\bar{y})$, then $\bar{x}$ is the unique feasible (and hence optimal) solution of $(SP \cdot \bar{y})$. Otherwise, if $\bar{x}_a > g_a(1)$ for some $a \in T(\bar{y})$, then problem $(SP \cdot \bar{y})$ is not feasible. □

3.2 Cut of Type I

When $\bar{x}$ is feasible for $(SP \cdot \bar{y})$, then a cut of type I has to be generated whenever $v(\bar{y}) > \bar{y}_0$. The fact that $G(x, y)$ is linear in x for fixed values of $y \in Y$ guarantees the existence of optimal multipliers for problem $(SP \cdot \bar{y})$. Let $\bar{u}$ and $\bar{\pi}$ denote the vectors of optimal multipliers associated with the capacity constraints $G(x, y) \leq 0$ and the flow conservation constraints specifying X, respectively. These multipliers are obtained by solving the Kuhn-Tucker optimality conditions of problem $(SP \cdot \bar{y})$: for all $a \in A$ (assuming that a is leading from node i to node j).

$$\bar{u}_a + \bar{\pi}_j - \bar{\pi}_i + \nabla_a f_3(\bar{x},\bar{y}) + \nabla_a f_1(\bar{x}) \geq 0 \quad (3.1)$$
$$(\bar{u}_a + \bar{\pi}_j - \bar{\pi}_i + \nabla_a f_3(\bar{x},\bar{y}) + \nabla_a f_1(\bar{x}))\,\bar{x}_a = 0 \quad (3.2)$$
$$\bar{u}_a(\bar{x}_a - g_a(\bar{y}_a)) = 0 \quad (3.3)$$
$$\bar{u}_a \geq 0 \quad (3.4)$$

where $\nabla_a f_3(\bar{x},\bar{y})$ and $\nabla_a f_1(\bar{x})$ are the partial derivatives with respect to x_a of $f_3(x,y)$ and $f_1(x)$, respectively.

To derive optimal multipliers (which are not unique), two separate cases are analyzed:

i) $a \in T(\bar{y})$

Hence the capacity constraint $x_a \leq g_a(\bar{y}_a) = g_a(1)$ allows x_a to be positive since
$g_a(1) = U_a > 0$. Furthermore, if

$$\bar{u}_a + \bar{\pi}_j - \bar{\pi}_i + \nabla_a f_3(\bar{x},\bar{y}) + \nabla_a f_1(\bar{x}) = 0$$

and
$$\bar{u}_a = 0 \ , \quad (3.5)$$

then conditions (3.1) to (3.4) are satisfied. It follows that for these arcs $a \in T(\bar{y})$

$$\bar{\pi}_j - \bar{\pi}_i + \nabla_a f_3(\bar{x},\bar{y}) + \nabla_a f_1(\bar{x}) = 0 \ . \quad (3.6)$$

Now, since $\bar{y}$ induces a spanning directed tree rooted at source node s, it follows from (3.6)
that a vector $\bar{\pi}$ of optimal multipliers is determined as follows:

$$\bar{\pi}_s = 0 \quad (3.7)$$
$$\bar{\pi}_j = \bar{\pi}_i - \nabla_a f_3(\bar{x},\bar{y}) - \nabla_a f_1(\bar{x}) \quad , \quad j \in N, \ j \neq s \quad (3.8)$$

where a is the unique arc leading from i to j such that $\bar{y}_a = 1$.

ii) $a \notin T(\bar{y})$

Hence $\bar{x}_a = 0$, and it follows that conditions (3.1) to (3.4) are satisfied if

$$\bar{u}_a + \bar{\pi}_j - \bar{\pi}_i + \nabla_a f_3(\bar{x}, \bar{y}) + \nabla_a f_1(\bar{x}) \geq 0$$

and

$$u_a \geq 0 \ .$$

Using the values of π_i and π_j determined in (3.7) and (3.8), it follows that

$$\bar{u}_a = [\pi_i - \pi_j - \nabla_a f_3(\bar{x}, \bar{y}) - \nabla_a f_1(\bar{x})]^+ = \max\{0, \bar{\pi}_i - \bar{\pi}_j - \nabla_a f_3(\bar{x}, \bar{y}) - \nabla_a f_1(\bar{x})\} \quad (3.9)$$

satisfies conditions (3.1) to (3.4).

Finally, it is interesting to note that since Lagrangean duality applies, it follows that

$$\inf_{x \in X} \left\{ f_1(x) + f_3(x, \bar{y}) + \bar{u}^T G(x, \bar{y}) \right\} = v(\bar{y}) > \bar{y}_0,$$

and then the cut of type I specified with $\bar{u}$ is efficient to eliminate solution $\bar{y}$.

These developments can be summarized in the following theorem.

Theorem 3.1: Assume that $\bar{x}$ is the optimal solution of problem $(SP \cdot \bar{y})$ and that $v(\bar{y}) > \bar{y}_0$. Then $\bar{u}$ and $\bar{\pi}$ specified in (3.5), and (3.7) to (3.9) are such that

$$\inf_{x \in X} \left\{ f_1(x) + f_3(x, \bar{y}) + \bar{u}^T G(x, \bar{y}) \right\} > \bar{y}_0 \ .$$

3.3 Cut of Type II

Recall that a cut of type II is introduced whenever the subproblem $(SP \cdot \bar{y})$ is not feasible. Hence, if $\bar{x}$ is not feasible for $(SP \cdot \bar{y})$, it follows that there exists an arc $a \in A$ such that

$$\bar{x}_a > g_a(\bar{y}_a) = g_a(1) \ = U_a > 0 \ .$$

A cut of type II associated with such a $\bar{y} \in Y$ is formulated as

$$\inf_{x \in X} \left\{ \bar{\lambda}^T G(x, y) \right\} \leq 0 \ .$$

We introduce a procedure to determine a vector $\bar{\lambda} \in \wedge$ such that $\inf_{x \in X} \{\bar{\lambda}^T G(x, \bar{y})\} > 0$. Hence this cut is efficient in the sense that it eliminates $\bar{y}$ and, therefore, it reduces the feasible domain of the relaxations of (MP) obtained by adding this constraint.

The procedure is quite similar to the one introduced by Gavish [8] to deal with the capacitated minimal spanning tree problem. The basic idea is to determine the multipliers by giving a bigger weight to the arcs where the constraint is violated. By doing that, we hope to force new arcs into the network. Furthermore, we want to determine the multipliers without having to solve an optimization problem.

Denote

- $T^+(\bar{y}) = \{a \in T(\bar{y}) : \bar{x}_a > g_a(\bar{y}_a)\}$;
- for all $a \in T(\bar{y})$,

  $N(a) = \{j \in N : j$ is reached from s through $a\}$;
- $N_1 = \bigcup_{a \in T^+(\bar{y})} N(a)$;
- $N_2 = N - N_1$.

To illustrate, consider Figure 3.2

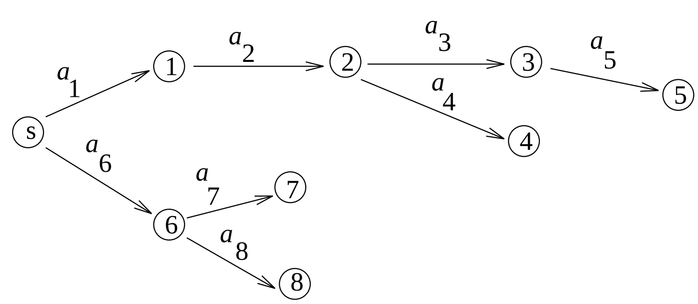

Figure 3.2

Suppose that $T^+(\bar{y}) = \{a_2, a_7\}$. Then

$$N(a_2) = \{2, 3, 4, 5\} \quad , \quad N(a_7) = \{7\}$$
$$N_1 = \{2, 3, 4, 5, 7\} \quad , \quad N_2\{s, 1, 6, 8\} \ .$$

Now, let

$$\rho_j = \begin{cases} 1 & \text{if } j \in N_1 \\ 0 & \text{if } j \in N_2 \ . \end{cases}$$

Furthermore, for all arcs $a \in A$ leading from i to j, let

$$\tilde{\lambda}_a = [\rho_j - \rho_i]^+ = \max\{0, \rho_j - \rho_i\} \ . \tag{3.10}$$

The following result indicates that the vector obtained by normalizing $\tilde{\lambda}$ is appropriate to generate a cut of type II.

Theorem 3.2: Assume that the vector $\tilde{\lambda}$ is specified by (3.10). Then

a) $\bar{x}$ is an optimal solution of $\inf_{x \in X} \left\{ \tilde{\lambda}^T G(x, \bar{y}) \right\}$, and

b) $\tilde{\lambda}^T G(\bar{x}, \bar{y}) > 0$.

Proof:

a) The problem $\inf_{x \in X} \left\{ \tilde{\lambda}^T G(x, \bar{y}) \right\}$ can be written explicitly as

$$\begin{array}{ll} \text{Min} & -\sum_{a \in A} \tilde{\lambda}_a g_a(\bar{y}_a) + \sum_{a \in A} \tilde{\lambda}_a x_a \\ \text{Subject to:} & \\ & \sum_{a \in \Gamma_j^-} x_a - \sum_{a \in \Gamma_j^+} x_a = d_j \quad , \quad j \in N, j \neq s \qquad (3.11) \\ & \qquad x_a \geq 0 \quad , \quad a \in A \ . \end{array}$$

Discarding the constant term $-\sum_{a \in A} \tilde{\lambda}_a g_a(\bar{y}_a)$ from the objective function, the dual problem of the resulting linear programming problem is as follows:

$$\begin{array}{ll} \text{Max} & \sum_{\substack{j \in N \\ j \neq s}} d_j \ \pi_j \\ \text{Subject to:} & \\ & -\pi_i + \pi_j \leq \tilde{\lambda}_a \quad , \quad a \in A \text{ leading from } i \neq s \text{ to } j \\ & \pi_j \leq \tilde{\lambda}_a \quad , \quad a \in A \text{ leading from } s \text{ to } j \ . \end{array}$$

Obviously, $\bar{x}$ is a feasible solution of the (primal) linear programming problem (3.11). To show that $\bar{x}$ is an optimal solution of (3.11), it is sufficient to identify a dual feasible solution $\bar{\pi}$ such that

$$\sum_{a \in A} \tilde{\lambda}_a \bar{x}_a = \sum_{\substack{j \in N \\ j \neq s}} d_j \bar{\pi}_j \ . \tag{3.12}$$

Now, since $\tilde{\lambda}_a = [\rho_j - \rho_i]^+$, it follows that $\tilde{\lambda}_a \geq \rho_j - \rho_i$. Hence ρ is a dual feasible solution. Furthermore, substitute

$$\begin{aligned} \bar{\pi}_j &= \rho_j \\ \tilde{\lambda}_a &= [\rho_j - \rho_i]^+ \\ d_j &= \sum_{a \in \Gamma_j^-} \bar{x}_a - \sum_{a \in \Gamma_j^+} \bar{x}_a \end{aligned}$$

in relation (3.12):

$$\sum_{a \in A} [\rho_j - \rho_i]^+ \bar{x}_a = \sum_{\substack{j \in N \\ j \neq s}} \left(\sum_{a \in \Gamma_j^-} \bar{x}_a \ - \ \sum_{a \in \Gamma_j^+} \bar{x}_a \right) \rho_j \ . \tag{3.13}$$

If arc $a \notin T(\bar{y})$, then $\bar{x}_a = 0$, and hence, $[\rho_j - \rho_i]^+ \bar{x}_a = 0$. Moreover, referring to the definition of N_1, we have $\rho_j \geq \rho_i$ for all $a \in T(\bar{y})$, where a is leading from i to j. Therefore, $[\rho_j - \rho_i]^+ = \rho_j - \rho_i$. Then, the left hand side of (3.13) reduces to $\sum_{a \in T(\bar{y})} (\rho_j - \rho_i) \bar{x}_a$,

but

$$\begin{aligned} \sum_{a \in T(\bar{y})} (\rho_j - \rho_i) \bar{x}_a &= \sum_{a \in T(\bar{y})} \rho_j \bar{x}_a \ - \ \sum_{a \in T(\bar{y})} \rho_i \bar{x}_a \\ &= \sum_{\substack{j \in N \\ j \neq s}} \rho_j \sum_{a \in \Gamma_j^-} \bar{x}_a \ - \ \sum_{\substack{j \in N \\ j \neq s}} \rho_j \sum_{a \in \Gamma_j^+} \bar{x}_a \\ &= \sum_{\substack{j \in N \\ j \neq s}} \rho_j \left(\sum_{a \in \Gamma_j^-} \bar{x}_a - \sum_{a \in \Gamma_j^+} \bar{x}_a \right) \end{aligned}$$

Hence, (3.13) is verified, and $\bar{x}$ is an optimal solution of $\inf_{x \in X} \left\{ \tilde{\lambda}^T G(x, \bar{y}) \right\}$.

b) Since $(SP \cdot \bar{y})$ is not feasible, it follows that there exists at least one arc $a \in T(\bar{y})$ such that $\bar{x}_a > g_a(\bar{y}_a), \rho_j = 1$, and $\rho_i = 0$. Furthermore, for all arcs $a \notin T(\bar{y}), \bar{x}_a = g_a(\bar{y}) = 0$. Also, for all arcs $a \in T(\bar{y})$ such that $\bar{x}_a \leq g_a(\bar{y}_a)$, it follows from the definitions of N_1 and N_2, that $\tilde{\lambda}_a = 0$. Hence it follows that

$$\tilde{\lambda}^T G(\bar{x}, \bar{y}) = \sum_{a \in A} \tilde{\lambda}_a \left(\bar{x}_a - g_a(\bar{y}_a)\right) > 0 \ .$$

□

The vector of multipliers $\bar{\lambda} \in \wedge$ used to specify the cut is obtained by normalizing $\bar{\lambda}$; i.e.,

$$\bar{\lambda} = \frac{\tilde{\lambda}}{\sum_{a \in A} \tilde{\lambda}_a} \ .$$

4 Closed Forms of the Cuts and Lagrangian Relaxation

In Section 3.2, we show that closed forms exist for cuts of type II since $\inf_{x \in X} \left\{ {\lambda^r}^T G(x, y) \right\} \leq 0$ can be replaced by ${\lambda^r}^T G(x^r, y) \leq 0$ where $x^r \in X$ is used to identify λ^r. To see how closed forms for cuts of type I can be obtained, two different cases are considered. In the first case, assume that f_3 and G are separable in x and y : i.e., $f_3(x, y) = f_4(x) + f_5(y)$ and $G(x, y) = G_1(x) + G_2(y)$. Then reffering to Geoffrion [9], it follows that the k^{th} cut of type I can be formulated as

$$v(y^k) - f_5(y^k) - {u^k}^T G_2(y^k) + f_5(y) + {u^k}^T G_2(y) \leq y_0 \tag{4.1}$$

where u^k and $v(y^k)$ are the vector of multipliers and the optimal value of $(SP \cdot y^k)$, respectively.

In the second case when the separability assumption does not hold for f_3 or G, we suggest to approximate the k^{th} cut of type I by the following closed form :

$$f_1(x^k) + f_3(x^k, y) + {u^k}^T G(x^k, y) \leq y_0 \tag{4.2}$$

where x^k is the (unique) optimal solution of subproblem $(SP \cdot y^k)$ at the iteration where u^k is determined.

Obviously, in this case the resulting problem $(\overline{MPR})$ approximating (MPR) has a smaller feasible domain and may have an optimal value larger than or equal to that of (MPR). Hence, the optimal value of $(\overline{MPR})$ may exceed the theoretical lower bound that would be generated with (MPR). Nevertheless, this strategy of using the approximating problem $(\overline{MPR})$ is justified to reduce the computational effort. Moreover, this approach has shown to be successful to deal with the problems analyzed in [2, 3, 6, 7].

It is interesting to note that the preceding approximation (4.2) coincides with the (exact) cut (4.1) when the separability assumption holds. Indeed,

$$\begin{aligned} f_1(x^k) + f_3(x^k, y) + u^{k^T} G(x^k, y) &= f_1(x^k) + f_4(x^k) + f_5(y) + u^{k^T} G_1(x^k) + u^{k^T} G_2(y) \\ &\quad + f_5(y^k) - f_5(y^k) + u^{k^T} G_2(y^k) - u^{k^T} G_2(y^k) \\ &= v(y^k) - f_5(y^k) - u^{k^T} G_2(y^k) + f_5(y) + u^{k^T} G_2(y) \end{aligned}$$

since

$$v(y^k) = f_1(x^k) + f_4(x^k) + f_5(y^k),$$

and by complementarity,

$$u^{k^T} [G_1(x^k) + G_2(y^k)] = 0.$$

Hence, (4.2) is valid in both cases. To reduce the burden of notation, we use it instead of (4.1) and we refer to $(\overline{MPR})$ in general, with the implicit convention that $(\overline{MPR})$ is in fact (MPR) whenever the separability assumption is verified.

Now, we can use Lagrangean relaxation [5] to deal with $(\overline{MPR})$ in order to take advantage of the spanning directed tree constraint. Indeed, if δ_k and ν_r denote the Lagrangean multipliers associated with the k[th] cut of type I and the r[th] cut of type II, respectively, then the Lagrangean relaxation problem is as follows:

$$\begin{aligned} \min_{y \in Y} \Big\{ f_2(y) \quad &+ \quad \left(1 - \sum_{k=1}^{K} \delta_k\right) \; y_0 \; + \sum_{k=1}^{K} \delta_k \left[f_1(x^k) + f_3(x^k, y) + u^{k^T} G(x^k, y) \right] \\ &+ \quad \sum_{r=1}^{R} \nu_r \; \lambda^{r^T} G(x^r, y) \Big\} \end{aligned}$$

This problem can be solved using an algorithm to determine spanning directed tree problem. Such algorithms have been developed by Tarjan [14, 15] and Edmonds [4].

A sub-gradient optimization technique [10] can be used to solve the Lagrangean dual, but unfortunately, this solution $\bar{y}$ is not feasible for $(\overline{MPR})$ in general. Of course, if $(\overline{MPR})$ does not include any cut of type II, or if the optimal solution $\bar{y}$ of the Lagrangean dual satisfies all cuts of type II, then the value of y_0 is easily adapted to make $\bar{y}$ a feasible solution of $(\overline{MPR})$ as follows:

$$\bar{y}_0 = \max_{1 \le k \le K} \left\{ f_1(x^k) + f_3(x^k, \bar{y}) + u^{k^T} G(x^k, \bar{y}) \right\} \; .$$

If some cuts of type II are not satisfied with $y = \bar{y}$, then ad hoc procedures have to be used to reduce their violation. Such a procedure is used in [7] for the electricity distribution network design problem where a pair of nodes (i, j) is selected to reduce the violation of one of the cuts of typeII by fixing an appropriate $y_a = 1$ for some arc a leading from i to j. Then the minimum spanning directed tree is solved with this additional constraint $y_a = 1$.

5 Applications

This approach has been used to deal with the electricity distribution problem [3, 7] that can be formulated as a network design problem over a horizon of several periods. The power originates from several substations(or sources), and it has to be distributed to load locations (or destinations) through a network. The capacities of the transformers at the substations, the distribution network topology and the capacities of the feeder lines have to be adjusted during the planning horizon according to variations in destination demands.

As indicated in the introduction, this problem can be formulated as a model of type (P). For each period of the planning horizon, a node is associated with each substation (source) and with each load location. A special node is associated with the supersource linked with each the nodes associated with the substations. The arcs in A correspond to potential combination of transformers installed in the substations over the horizon (arcs from supersource to substations) or to potential locations and types for feeder lines. Whenever several combinations of transformers or several types of feeder lines of different capacity are available, then they are associated with different arcs.

The variables y are used to specify the topology of the distribution network and the equipments used at each period of the horizon. Hence, the capacity of the arcs are specified via the function $G(x, y)$ in order to determine the power flow ${x_{ij}}^t$ in each line (i, j) at each period t.

The cost to modify the network topology from one period to the next and the cost to operate the equipments at each period are included in $f_2(y)$. Furthermore, f_3 denotes the cost associated with the power loss (in lines) that depends of the network topology, of the equipments used (variables y) and of the power flow in the lines (variables x).

An extension of this approach has been used in the capacity and flow assignment of data networks problem [12] wich is a mixed-integer non linear model to optimize jointly the assignment of capacities and flows in a communication network. Discrete capacities are considered and a total delay constraint models the grade of service of the desired network. Generalized Benders decomposition induces convex subproblems which are multicommodity flow problems on different topologies with fixed capacities.

Besides the classical cuts (cuts of type I and II), two other kinds of cuts are generated: connexity cuts and spanning tree cuts. The introduction of the connexity cuts and spanning tree cuts reduces the combinatorial choice between capacities lowering then the number of iterations.

References

[1] R.N. Adam and M.A. Laughton, "Optimal Planning of Power Networks Using Mixed Integer Programming", *Proc. IEEE* **121**, 139–146 (1974).

[2] A. Benchakroun, "*Un Modèle de Planification des Réseaux de Distribution d'Énergie Électrique*", Ph.D. Dissertation, Département d'Informatique et de Recherche Opérationnelle, Université de Montréal (1988).

[3] A. Benchakroun, J.A. Ferland and R. Cléroux, "Distribution System Planning Through a Generalized Benders Decomposition Approach", *European Journal of Oper. Res.* **62**, 149-162 (1991).

[4] J. Edmonds, "Optimum Branchings", *Journal of Research of the National Bureau of Standards-B, Mathematics and Mathematical Physics* **71**B, 233-240 (1967).

[5] M.L. Fisher, "The Lagrangean Relaxation Method for Solving Integer Programming Problems", *Management Science* **27**, 1-18 (1981).

[6] V. Gascon, "*Modèle de Distribution d'Énergie Électrique avec Structures de Réseau*", Ph.D. Dissertation, Département d'Informatique et de Recherche Opérationnelle, Université de Montréal (1990).

[7] V. Gascon, A. Benchakroun, J.A. Ferland, "Electricity Distribution Planning Model: A Network Design Approach Solving the Master Problem of the Benders Decomposition Method", *INFOR* **31**, 205-220 (1993).

[8] B. Gavish, "Topological Design of Centralized Computer Networks-Formulation and Algorithms", *Networks* **12**, 355-377 (1992).

[9] A.M. Geoffrion, "Generalized Benders Decomposition", *Journal of Optimization Theory and Applications* **10**, 237-260 (1972).

[10] M. Held, P. Wolfe and H.P. Crowder, "Validation of Subgradient Optimization", *Mathematical Programming* **6**, 62-88 (1974).

[11] T.L. Magnanti and R.T. Wong, "Network Design and Transportation Planning: Models and Algorithms", *Transportation Science* **18**, 1-55 (1984).

[12] P. Mahey, A. Benchakroun, F. Boyer, "Capacity and flow assignment of data networks by generalized Benders decomposition", **To appear**.

[13] M. Minoux, "Network Synthesis and Optimum Network Design Problems: Models, Solution Methods and Application", *Networks* **19**, 313-360 (1989).

[14] R.E. Tarjan, "Finding Optimum Branchings", *Networks* **7**, 25-35 (1977).

[15] R.E. Tarjan, "Depth-First Search and Linear Graph Algorithms", *Siam J. Comput.* **1**, 146-160 (1972).

[16] N. Zadeh, "On Building Minimum Cost Communication Network Over Time", *Networks* **4**, 19-34 (1974).