Journal of Global Optimization **19:** 1–26, 2001.

Effective Relaxations and Partitioning Schemes for Solving Water Distribution Network Design Problems to Global Optimality

HANIF D. SHERALI[1], SHIVARAM SUBRAMANIAN[2] and G.V. LOGANATHAN[3]
[1]*Virginia Polytechnic Institute and State University, Department of Industrial and Systems Engineering (0118), 250 New Engineering Building, Blacksburg, VA 24061, USA (e-mail: hanifs@vt.edu);* [2]*United Airlines World Headquarters, Research and Development, Elk Grove Village, USA;* [3]*Virginia Polytechnic Institute and State University, Department of Civil Engineering, Blacksburg, USA*

(Received 13 August 1999; accepted in revised form 3 July 2000)

Abstract. In this paper, we address the development of a global optimization procedure for the problem of designing a water distribution network, including the case of expanding an already existing system, that satisfies specified flow demands at stated pressure head requirements. The proposed approach significantly improves upon a previous method of Sherali et al. (1998) by way of adopting tighter polyhedral relaxations, and more effective partitioning strategies in concert with a maximal spanning tree-based branching variable selection procedure. Computational experience on three standard test problems from the literature is provided to evaluate the proposed procedure. For all these problems, proven global optimal solutions within a tolerance of 10^{-4}% and/or within 1$ of optimality are obtained. In particular, the two larger instances of the Hanoi and the New York test networks are solved to global optimality for the very first time in the literature. A new real network design test problem based on the Town of Blacksburg Water Distribution System is also offered to be included in the available library of test cases, and related computational results are presented.

Key words: Global Optimization, Water Distribution Systems, Reformulation-Linearization Technique, Branch-and-Bound, Maximal Spanning Tree

1. Introduction

The design of water distribution systems has consistently received a great deal of attention because of its importance to society. Even so, many of the existing pipe networks in older urbanized areas function at severely reduced levels of efficiency, and in some cases, are inadequate with respect to meeting the required pressure and flow demands. The investments associated with the installation, expansion and maintenance of water distribution systems are very high, and account for the largest proportion in municipal maintenance budgets. An important component in this process of designing a cost effective water distribution system or extending a pre-existing network is to design the sizes of the various pipes that are capable of satisfying the flow demand, in addition to satisfying the minimum pressure head and hydraulic redundancy requirements. However, this least cost pipe design problem is a hard nonconvex optimization problem having a number of local optima, and has hence proven difficult to solve. A number of research efforts over the last two decades have focused on solving this problem, most of them generating improved suboptimal solutions for several standard test problems from the literature, with no adequate lower bounds to evaluate the prescribed solutions. Three notable exceptions discussed below that are capable of providing solutions within a proven tolerance of a global optimum are the methods due to Eiger et al. (1994), Sherali and Smith (1997), and Sherali et al. (1998). Our development in the present paper is a further enhancement of these three procedures. For other related expositions, we point the reader to the survey by Lansey and Mays (1985), the holistic integrated pipe-reliability-and-cost network optimization approach and implementation discussion in Sherali and Smith (1993), and the recent application of genetic algorithms as described in Dandy et al. (1996).

A first global optimization approach to the least cost pipe sizing decision was proposed by Eiger et al. (1994). A branch-and-bound algorithm is developed in this paper, based on partitioning the hyperrectangle restricting the flows into several subrectangles. At each node of the branch-and-bound tree, a subgradient-based heuristic is applied to determine an upper bound via the nonsmooth, nonconvex, projection of the problem onto the space of the flow variables. An independent relaxed, duality-based linear programming formulation is used to compute lower bounds.

Sherali and Smith (1997) present another global optimization approach for an arc-based formulation of the problem, in contrast with the loop-and-path based formulation employed by Eiger et al. (1994). They employ a Reformulation-Linearization Technique (RLT) to construct tight linear programming relaxations for the given problem in order to compute lower bounds. The procedure is embedded in a branch-and-bound scheme. Convergence to an optimal solution is induced by coordinating this process with an appropriate partitioning scheme. Several test problems from the literature are solved to exact global optimality for the first time using this approach. In particular, these results indicate that some of the solutions reported by Eiger et al. (1994) are in error due to a degree of infeasibility in the flow conservation constraints.

Sherali et al. (1998) provide an alternative global optimization approach that enhances the method of Eiger et al. (1994), thereby enabling them to solve the Hanoi network test problem to within 0.486% of optimality. This significantly improved upon the previously best solution reported in the literature. In this procedure, they derive a linear lower bounding problem by relaxing the nonlinear constraints in the transformed space via polyhedral outer approximations. Upper bounds are computed by solving a projected linear program which uses the flow conserving solution generated by the lower bounding problem. These bounding strategies are embedded within a branch-and-bound algorithm. A partitioning scheme is employed that induces a convergent process toward a feasible solution that lies within any prescribed accuracy tolerance of global optimality.

The global optimization method addressed in the present paper provides a further enhancement of Sherali et al.'s (1998) polyhedral outer approximation scheme by way of (a) employing tighter relaxations; (b) using a maximal spanning tree based projected space partitioning scheme that dramatically reduces the computational effort, and (c) designing a more effective branching variable selection strategy. An alternative RLT-based relaxation that is more effective than the one proposed by Sherali and Smith (1997) is also developed and tested for comparison purposes. The foregoing enhancements are shown to substantially reduce computational effort while determining proven global optimal solutions that lie at least within 10^{-4}% of optimality to standard test problems available in the literature. The results provide improved incumbent solutions over those previously reported in the literature for all these problems, and particularly so, for the two larger Hanoi and New York City problems. In fact, for the latter problem, not even a reasonable lower bound had previously been reported. We derive a tight global lower bound for this problem for the first time, solving this test case to within 10^{-6}% (or \$0.4) of optimality.

The remainder of this paper is organized as follows. Section 2 presents the network optimization model, and Section 3 derives the proposed linear programming lower bounding problem and the upper bounding heuristic. The branch-and-bound algorithm is described in Section 4. Section 5 presents extensive computational results on test problems from the literature, as well as for a new Blacksburg network that is offered as another test case to researchers. Finally, Section 6 concludes the paper with a discussion on possible algorithmic variants and further enhancements.

2. Model Formulation

Consider a distribution network $G(N, A)$ comprised of a set of reservoirs or supply nodes and a set of consumption or demand nodes. Let these nodes be collectively identified by the index set $N = \{1, 2, \ldots, n\}$, where the set of source nodes is denoted by $S \subset N$ and the set of demand nodes is denoted by $D \subset N$ such that $N = S \cup D$. Associate with each node a quantity b_i that represents the net water supply rate or demand rate corresponding to node i in the index set N. We will assume that $b_i > 0$ for $i \in S$ and $b_i \leqslant 0$ for $i \in D$. To ensure feasibility, we assume that the total supply rate is at least equal to the total demand rate.

For each pipe (new or existing) that connects a certain designated node pair i and j, where $i, j \in N, i < j$, we create a (notationally) directed arc $(i, j) \in A$. For each $(i, j) \in A$, let L_{ij} denote the pipe length corresponding to a connection in the network between the nodes i and j. If we are working on the more general problem of expanding an existing network, the problem becomes one of designing new connections as well as constructing parallel pipe links that need to be installed between certain specified node pairs i and j, similar to the consideration of Loganathan et al. (1995). (When the existing pipe is being *replaced*, we simply treat this case similar to that of designing a new connection between the corresponding nodes.) Subramanian (1999) describes the transformations necessary to avoid parallel multi-arcs in the resulting model formulation. Hence, we assume that each $(i, j) \in A$ represents a unique arc. Let us denote the set of arcs in A that are to be *newly designed* as P, and let $A - P$ represent the *existing links* in the network. We assume that each link that needs to be designed is constructed from segments of lengths having standard available diameters, chosen from the set $\{d_k, k = 1, \dots, K\}$. Also, let us denote by c_k the cost per unit length for a pipe of diameter d_k.

Associated with each link connecting node pairs $(i, j) \in A$ is the decision variable q_{ij} that represents the flow rate (m^3/hr). Note that this variable may be nonnegative or negative, thus permitting flow in either direction. A positive flow value means that flow is along the specified conventional direction of the arc. The value q_{ij} associated with each link is assumed to lie between some analytically determined minimum and maximum bounds $q\text{min}_{ij}$ and $q\text{max}_{ij}$, that may appropriately be of either sign. We define the hyperrectangle restricting the flows $\boldsymbol{q}$ as $\Omega = \{\mathbf{q} : \mathbf{q}min \leqslant \mathbf{q} \leqslant \mathbf{q}\text{max}\}$, where the (boldface) notation $\mathbf{q}\,\text{min}$ and $\mathbf{q}\,\text{max}$ with the subscripts dropped denotes the corresponding vectors of lower and upper bounds. Sherali et al. (1998) discuss procedures for determining these bounds on the flows from the network configuration using logical arguments (without making any *a priori* assumption on the nature of the optimal flow distribution).

Our next set of decision variables relates to the lengths of pipe segments having different standard diameters, that comprise each link of the network. Let x_{ijk} denote the length of segment of diameter k in the link $(i, j) \in A$, and let $\mathbf{x}_{ij}$ be the vector having components $(x_{ijk}, k = 1, \dots, K)$. We assume that the variable x_{ijk} is fixed at a value $\tilde{x}_{ijk}$ for all arcs $(i, j) \in A - P$, $\forall k = 1, \dots K$.

Let us now consider the energy heads at the various nodes in the network. For each node $i \in N$, let E_i denote its ground elevation, and let H_i (a variable) denote the established head above E_i. Additionally, for the source nodes $i \in S$, let F_i denote the fixed maximum available energy head, and suppose that there is an opportunity to further raise this head by an amount H_{si} at an annualized cost $c_{si} > 0$ per unit energy head, as suggested by Rowell and Barnes (1982). Correspondingly, for each demand node $i \in D$, suppose that there is the requirement that at a flow equilibrium, the established head $(H_i + E_i)$ at this node lies in the interval $[H_{iL}, H_{iU}]$ where $H_{iL} < H_{iU}$.

The pressure loss (or head-loss) in a pipe due to friction, given by $[(H_i + E_i) - (H_j + E_j)]$ for a link (i, j), can be described by the empirical Hazen-William equation as follows (see Walski, 1984), where the sign depends on the direction of flow.

$$\Phi(q, C_{HW}, d, x) = (1.52)\ 10^4\ \text{sign}(q)|q/C_{HW}|^{1.852} d^{-4.87} x \qquad (2.1)$$

where

Φ = pressure head-loss (in meters) assuming smooth-flow conditions in a given pipe segment,
q = water flow rate in the pipe (m^3/hr),
C_{HW} = Hazen-Williams coefficient based on roughness and diameter,
d = pipe diameter (in centimeters),
x = pipe length (in meters).

For our model, the head loss in a pipe that has several potential segments of varying diameter and roughness is computed as follows:

$$\Phi_{ij}(q_{ij}, \mathbf{x}_{ij}) = \sum_{k=1}^{K} \Phi(q_{ij}, C_{HW(ijk)}, d_k, x_{ijk}), \quad \text{where } \mathbf{x}_{ij} \equiv (x_{ijk}, k = 1, \ldots, K). \tag{2.2}$$

The network optimization problem NOP, restricted on Ω, can now be formulated as follows.

NOP(Ω):

$$\text{Minimize} \sum_{(i,j)\in P} \sum_{k=1}^{K} c_k x_{ijk} + \sum_{i\in S} c_{si} H_{si} \tag{2.3a}$$

subject to

$$\Phi_{ij}(q_{ij}, \mathbf{x}_{ij}) = (H_i + E_i) - (H_j + E_j) \quad \forall (i, j) \in A \tag{2.3b}$$

$$\sum_{k=1}^{K} x_{ijk} = L_{ij} \quad \forall (i, j) \in P \tag{2.3c}$$

$$\sum_{j\in FS(i)} q_{ij} - \sum_{j\in RS(i)} q_{ji} = b_i \quad \forall i \in D \tag{2.3d}$$

$$\sum_{j\in FS(i)} q_{ij} - \sum_{j\in RS(i)} q_{ji} \leqslant b_i \quad \forall i \in S \tag{2.3e}$$

$$q\ \text{min}_{ij} \leqslant q_{ij} \leqslant q\ \text{max}_{ij} \quad \forall (i, j) \in A \tag{2.3f}$$

$$H_i + E_i \leqslant F_i + H_{si} \quad \forall i \in S \tag{2.3g}$$

$$H_{iL} \leqslant H_i + E_i \leqslant H_{iU} \quad \forall i \in D \tag{2.3h}$$

$$H_{si} \geqslant 0 \quad \forall i \in S \tag{2.3i}$$

$$x_{ijk} \geqslant 0 \quad \forall (i, j) \in P, k = 1, \ldots, K$$

$$x_{ijk} = \tilde{x}_{ijk} \quad \forall (i, j) \in A - P, k = 1, \ldots, K. \tag{2.3j}$$

The objective function, Equation (2.3a), denotes the total cost of the pipes and the cost of the additional head generated at each source node. Constraints (2.3b) are the conservation of energy equations, and along with Constraints (2.3h), ensure that the hydraulic energy loss over each chain in the network is such that the minimum head requirements (H_{iL}) are met for each demand node. It may be noted that

Constraints (2.3b) implicitly enforce that the hydraulic energy loss in each loop in the network is zero. The link length constraints are represented by Equation (2.3c). Equations (2.3d) and (2.3e) enforce conservation of flow at all nodes, and Equation (2.3f) bounds the flow value in each link to lie in a specified valid or implied interval. These bounds that define Ω will be suitably modified during the course of the algorithm for solving Problem NOP. Constraints (2.3g) and (2.3h) represent restrictions on the maximum variable head at each source node, and the head requirements at each demand node, respectively. Finally, Constraints (2.3i) and (2.3j) enforce logical nonnegativity restrictions, and require that the variables x_{ijk} are fixed at the corresponding pre-specified values $\tilde{x}_{ijk}$ for the existing arcs $(i, j) \in A - P$ in the network.

Our principal set of decision variables are the lengths x_{ijk} of the different segments comprising each link $(i, j) \in P$, and the additional head H_{si} to be developed at each source node $i \in S$. The resulting heads H_i at each node $i \in N$ (above the elevation E_i of the node) and the flows q_{ij} in the links $(i, j) \in A$ are also problem variables that happen to be governed by the foregoing design variables.

3. Lower and Upper Bounding Problems

The frictional head loss expression in the Constraints (2.3b) cause NOP(Ω) to become nonlinear and nonconvex. Following Eiger et al. (1994) and Sherali et al. (1998), we take advantage of the monotone nature of these constraints to transform the problem NOP(Ω) into a set of newly defined variables, and accordingly, develop suitable relaxations for the flow conservation constraints that turn out to be nonlinear in the projected space of these new decision variables.

Since the relations derived subsequently hold true for each link, the subscripts defining the links will be dropped for convenience. Equation (2.2) which appears in Constraints (2.3b) can be written as follows using Equation (2.1), for any link having a flow q and a length segment vector $\mathbf{x} = (x_k, k = 1, \ldots, K)$.

$$\Phi(q, \mathbf{x}) = \sum_{k=1}^{K} \operatorname{sign}(q)|q|^{1.852}(1.52)10^4(C_{HW(k)})^{-1.852} d_k^{-4.87} x_k. \tag{3.1}$$

Denoting

$$v(q) \equiv \operatorname{sign}(q)|q|^{1.852}, \tag{3.2}$$

Equation (3.1) can be rewritten as follows,

$$\Phi(q, \mathbf{x}) = \sum_{k=1}^{K} v(q)\alpha_k x_k, \tag{3.3}$$

$$\text{where } \alpha_k \equiv (1.52)10^4(C_{HW(k)})^{-1.852} d_k^{-4.87}. \tag{3.4}$$

By the monotonicity of $v(q)$, we can represent its value for any q as some convex combination of its minimum and maximum values $v(q\text{min})$ and $v(q\text{max})$, henceforth denoted as $v$min and $v$max, respectively.

$$v(q) = \lambda v\text{min} + (1 - \lambda)v\text{max}, \quad \text{for some } 0 \leqslant \lambda \leqslant 1. \tag{3.5}$$

Using the representation (3.5) in Equation (3.3) and rearranging terms, we get,

$$\Phi(q, \mathbf{x}) = \sum_{k=1}^{K}(v\text{min})\alpha_k(\lambda x_k) + \sum_{k=1}^{K}(v\text{max})\alpha_k(1 - \lambda)x_k. \tag{3.6}$$

We now define our new decision variables as

$$x_k^1 = \lambda x_k \text{ and } x_k^2 = (1 - \lambda)x_k, \tag{3.7}$$

so that

$$x_k = x_k^1 + x_k^2. \tag{3.8}$$

Note that for the existing pipes, x_k is fixed at a value $\tilde{x}_k$ (some possibly zero) for all k. Equation (3.6) can now be rewritten in terms of the new decision variables as

$$\Phi(q, \mathbf{x}) = \sum_{k=1}^{K}(v\text{min})\alpha_k x_k^1 + \sum_{k=1}^{K}(v\text{max})\alpha_k x_k^2. \tag{3.9}$$

Note that in the space of the new decision variables, x^1, x^2 and λ we have linearized the energy conservation constraints by substituting (3.9) on the left-hand side of (2.3b), but at the expense of introducing nonlinearity elsewhere in the problem. Specifically, this nonlinearity arises in two sets of relationships. First it occurs in the nonlinear representation (3.7) that accompanies the (linear) relationship (3.8). Second, the flow q (for each generic link) is now given via (3.2) and (3.5) by the following function $q(\lambda)$,

$$q(\lambda) = \text{sign}[\lambda\ v\text{min} + (1 - \lambda)\ v\text{max}]|\lambda\ v\text{min} + (1 - \lambda)\ v\text{max}\,|^{1/1.852}. \tag{3.10}$$

Using the foregoing transformations, we therefore obtain an alternative equivalent representation for Problem NOP (Ω) as follows.

NOP(Ω):

$$\text{Minimize} \sum_{(i,j)\in P} \sum_{k=1}^{K} c_k(x_{ijk}^1 + x_{ijk}^2) + \sum_{i\in S} c_{si} H_{si} \tag{3.11a}$$

subject to:

$$\sum_{k=1}^{K} (v\text{min}_{ij})\alpha_{ijk} x_{ijk}^1 + \sum_{k=1}^{K} (v\text{max}_{ij})\alpha_{ijk} x_{ijk}^2 = (H_i + E_i) - (H_j + E_j) \quad \forall (i, j) \in A \tag{3.11b}$$

$$\sum_{k=1}^{K} x_{ijk}^1 + \sum_{k=1}^{K} x_{ijk}^2 = L_{ij} \quad \forall (i, j) \in P \tag{3.11c}$$

$$\sum_{j\in FS(i)} q_{ij} - \sum_{j\in RS(i)} q_{ji} = b_i \quad \forall i \in D \tag{3.11d}$$

$$\sum_{j\in FS(i)} q_{ij} - \sum_{j\in RS(i)} q_{ji} \le b_i \quad \forall i \in S \tag{3.11e}$$

$$q_{ij} = \text{sign}[\lambda_{ij} v\text{min}_{ij} + (1 - \lambda_{ij}) v\text{max}_{ij}] |\lambda_{ij} v\text{min}_{ij} + (1 - \lambda_{ij}) v\text{max}_{ij}|^{1/1.852} \quad \forall (i, j) \in A \tag{3.11f}$$

$$q\text{min}_{ij} \le q_{ij} \le q\text{max}_{ij} \quad \forall (i, j) \in A \tag{3.11g}$$

$$H_i + E_i \le F_i + H_{si} \quad \forall i \in S \tag{3.11h}$$

$$H_{iL} \le H_i + E_i \le H_{iU} \quad \forall i \in D \tag{3.11i}$$

$$H_{si} \ge 0 \quad \forall i \in S \tag{3.11j}$$

$$x_{ijk}^1, x_{ijk}^2 \ge 0 \quad \forall (i, j) \in P, k = 1, \ldots, K \tag{3.11k}$$

$$0 \le \lambda_{ij} \le 1 \quad \forall (i, j) \in A \tag{3.11l}$$

$$x_{ijk}^1 = \lambda_{ij} \tilde{x}_{ijk}, \quad \text{and} \quad x_{ijk}^2 = (1 - \lambda_{ij}) \tilde{x}_{ijk} \quad \forall (i, j) \in A - P, k = 1, \ldots, K \tag{3.11m}$$

$$x_{ijk}^1 = \lambda_{ij} x_{ijk}, \quad \text{and } x_{ijk}^2 = (1 - \lambda_{ij}) x_{ijk}, \quad \text{where } x_{ijk} \ge 0, \quad \forall (i, j) \in P, k = 1, \ldots, K. \tag{3.11n}$$

We will now construct relaxations for the nonlinear relationships (3.11f) and (3.11n) in order to derive lower bounding linear programs.

3.1. LOWER BOUNDING RELAXATION LB (Ω)

To derive this relaxation for Problem NOP(Ω) as given by (3.11), we first omit (3.11n), but replace it with its following aggregate relationship based on (2.3c):

$$\sum_{k=1}^{K} x_{ijk}^{1} = \lambda_{ij} L_{ij} \quad \forall (i, j) \in P. \tag{3.12}$$

Note that the symmetric relationship

$$\sum_{k=1}^{K} x_{ijk}^{2} = (1 - \lambda_{ij}) L_{ij} \quad \forall (i, j) \in P$$

is implied by (3.12) and (3.11c). Next, we relax the nonlinear relationship (3.11f) by constructing a polyhedral outer-approximation to this function which relates q_{ij} to $\lambda_{ij} \forall (i, j) \in A$ as in Sherali et al. (1998). In the most general case, this function is concave-convex as depicted in Figure 1 for the generic representation $q(\lambda)$ stated as a function of λ as in (3.10). In this figure, $\bar{\lambda}$ is such that, if it exists, the tangential support to $q(\lambda)$ at $\lambda = (\bar{\lambda})$ passes through the coordinate $(1, q\text{min})$ in the (λ, q) space. Similarly, $\tilde{\lambda}$ is such that, if it exists, the tangential support to $q(\lambda)$ at $\lambda = \tilde{\lambda}$ passes through the coordinate (0, $q$max) in the (λ, q) space. These two values are respectively computed via Equations (3.13) and (3.14) given below using a bisection search, and each has a solution if and only if the corresponding entity $\bar{\lambda}$ or $\tilde{\lambda}$ exists.

$$q\text{min} - q(\bar{\lambda}) - (1 - \bar{\lambda}) q'(\bar{\lambda}) = 0 \quad \text{where } 0 < \bar{\lambda} < 1, \tag{3.13}$$

$$q\text{max} - q(\tilde{\lambda}) + \tilde{\lambda} q'(\bar{\lambda}) = 0 \qquad \text{where } 0 < \tilde{\lambda} < 1, \tag{3.14}$$

and where, the derivative (or slope) of $q(\lambda)$, denoted by $q'(\lambda)$, is given by

$$q'(\lambda) = \frac{(v\text{min} - v\text{max})}{1.852} |\lambda \ v\text{min} + (1 - \lambda) v\text{max}|^{\frac{-0.852}{1.852}},$$
$$\text{for } \lambda \in [0, 1], \lambda \ v\text{min} + (1 - \lambda) v\text{max} \neq 0.$$

To enhance the relaxation used in Sherali et al. (1998), we construct the following more refined support structure. For each of the following cases, the corresponding nonlinear constraint (3.11f) is replaced by a set of affine inequalities that this relationship must satisfy in the (λ, q)-space. For the general concave-convex case of Figure 1, if $\bar{\lambda}$ and $\tilde{\lambda}$ exist, we construct six supporting facets for the polyhedral approximation via tangents at $\{0, \bar{\lambda}/2, \bar{\lambda}\}$ and at $\{\tilde{\lambda}, (1 + \tilde{\lambda})/2, 1\}$ as shown in Figure 1. In case $\bar{\lambda}$ does not exist, but $\tilde{\lambda}$ does, or vice versa, we replace the corresponding undefined supports with the affine concave or convex envelope, respectively, that pass through (0, $q$max) and (1, $q$min). This case yields only four supporting facets for the polyhedral approximation.

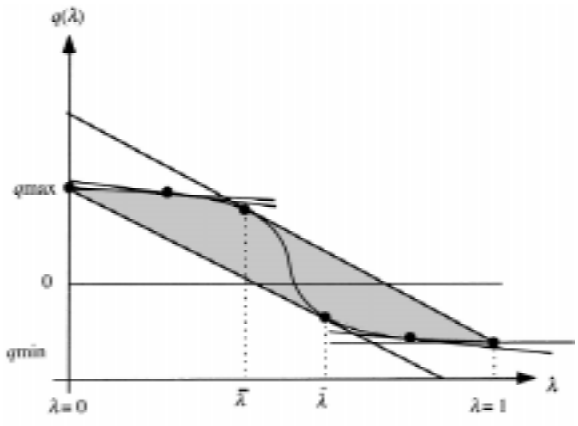

Figure 1. Polyhedral outer-approximation for relating q(λ) to λ.

In case $q(\lambda)$ is a concave function of λ (i.e., $q\max > 0$ and $q\min \geq 0$), we construct the affine convex envelope along with tangential supports at the points λ = 0, 0.25, 0.5, 0.8, and 0.9. Similarly, if $q\max \leq 0$ and $q\min < 0$, so that $q(\lambda)$ is a convex function of λ, we construct its affine concave envelope along with tangential supports at the points λ = 0.1, 0.2, 0.5, 0.75, and 1. Each of these cases produce six facets for the polyhedral approximation that replaces (3.11f) in the relaxation LB(Ω). We also experimented with using four facets for the polyhedral approximation, instead of the six used here, similar to the consideration of Sherali et al. (1998). We provide some comparative computational experience for this in Section 5.

3.2. LOWER BOUNDING RELAXATION RLT(Ω)

As an alternative to LB(Ω), we derive a further enhanced lower bounding procedure that is motivated by the Reformulation-Linearization Technique (RLT) of Sherali and Tuncbilek (1992) for solving polynomial programming problems. Our purpose here is to study the tradeoff between a quicker versus a more involved, but stronger, lower bounding procedure with respect to the overall effort for solving the problem. To construct such a lower bounding problem RLT(Ω), we augment Problem LB(Ω) by incorporating certain additional constraints that are generated using the RLT concept as follows.

3.2.1. *Reformulation Step*

The following quadratic valid constraints are generated based on the products of the stated pairs of inequalities (written in the form $\{\cdot\} \geq 0$), or based on products of equations with variables.

(a) Using the pipe length constraints in (3.11c), generate the equality product constraints

$$\left(\sum_{k=1}^{K} x_{ijk}^{1} + \sum_{k=1}^{K} x_{ijk}^{2}\right) q_{ij} = L_{ij} q_{ij} \quad \forall (i, j) \in P.$$

(b) Multiply each constraint in LB(Ω) that represents a linear inequality in q_{ij} and λ_{ij} for the corresponding polyhedral approximation to (3.11f), with each corresponding variable $x_{ijk}, \forall (i, j) \in P, k = 1, \ldots, K$.

3.2.2. *Linearization Step*

Linearize the resulting product constraints generated above by substituting $y_{ijk} = q_{ij} x_{ijk}, \forall (i, j) \in P, k = 1, \ldots, K$, and by using Equations (3.7) and (3.8). This produces a linear programming lower bounding problem RLT(Ω) that incorporates certain additional valid inequalities that must be satisfied by any feasible solution to the nonlinear problem NOP(Ω) given by (3.11).

3.2.3. *Upper Bounds*

For computing upper bounds on the least cost pipe sizing problem NOP(Ω), we fix the conserving flow solution as obtained via the lower bounding problem LB(Ω) or RLT(Ω) within the Problem NOP(Ω), and solve the resulting linear programming problem. If this problem is feasible, it yields an upper bounding completion to this fixed flow. As an alternative, a more refined, but computationally non-intensive local search heuristic could be employed to derive improved upper bounds.

4. A Branch-and-Bound Algorithm

We embed the lower and upper bounding schemes described in the foregoing section in a branch-and-bound procedure to solve NOP(Ω) globally to any specified percentage tolerance (100ϵ%) of optimality. Each branch-and-bound node principally differs in the specification of the hyperrectangle Ω. The hyperrectangle associated with node *t* of the branch-and-bound tree at the main iteration or stage *S* of the procedure is denoted by $\Omega^{S,t} = \{\mathbf{q} : \mathbf{qmin}^{S,t} \leq \mathbf{q} \leq \mathbf{qmax}^{S,t}\}$. In our implementation of the branch-and-bound procedure, we successively partition the hyperrectangle defined by the initial bounds $\Omega^{1,1} \equiv \Omega$ on the flow variables into smaller and smaller hyperrectangles. At any stage *S* of the branch-and-bound algorithm, we have a set of active or nonfathomed nodes denoted as T_S. We select an active node t^* in T_S that has the least lower bound (this is termed as the *global lower bound* at stage *S*), breaking ties arbitrarily, and partition this node using a variable selection strategy that is described below. The selection of a branching variable according to this strategy ensures convergence of the overall procedure to a global optimum for NOP(Ω) using the general theory discussed in Sherali et al. (1998). This process continues by solving the bounding problems for the resulting two node subproblems, and then fathoming the nodes for which the lower bound is greater than or equal to $UB(1-\epsilon)$, where UB is the value of the current incumbent solution, and $0 < \epsilon < 1$ is a suitable tolerance. (We used $\epsilon = 10^{-6}$ in our computations, thereby obtaining solutions to within 10^{-4}% of optimality.) Whenever the set of active nodes is empty, the process terminates. A formal statement of this algorithm is given in Sherali et al. (1998). Here, we focus on the principal difference based on (a) a reduction in the potential set of branching variables, and (b) a strategy for selecting a branching variable from this set. These two features are described in succession below.

4.1. MAXIMAL SPANNING TREE BASED APPROACH FOR REDUCING THE CANDIDATE SET OF BRANCHING VARIABLES (MSTR)

We can reduce the number of possible candidates for selecting a branching variable by restricting these to be a set of independent arcs in A. To see this, suppose that at the beginning of the branch-and-bound procedure, we construct a maximal spanning tree for the distribution network via Kruskal's (1956) procedure, using arc weights $(q\max_{ij} - q\min_{ij}) \forall (i, j) \in A$. The supply nodes are connected to a dummy sink via slack arcs having a large weight for this purpose, in order to balance supply and demand, and thereby obtain an equality flow conservation system. Hence, all these slack arcs are a part of the maximal spanning. Let B denote the set of arcs in this spanning tree. The remaining arcs $\{A - B\}$ in the network are designated as non-tree arcs and form the set from which the branching variables are selected. Note that such a spanning tree yields a valid basis for the underlying network flow problem, and given the flows on the independent nonbasic arcs, the corresponding flows on the dependent basic arcs are uniquely determined. Hence, only the nonbasic arcs corresponding to this (fixed) basis are selected for partitioning flow intervals. Upon fixing the flow bounds for the set of nonbasic arcs, the flow bounds for the basic arcs are updated using the representation of the dependent basic variables in terms of the independent nonbasic variables (see Bazaraa et al., 1990). Since most water distribution networks are almost 'tree-like', this reduction in the potential set of branching variables is substantial.

Given the flow bounds on the nonbasic arcs, the basic flow bounds are updated following the reverse thread (post-order) recursive tree traversal procedure that is typically employed for computing flows in network simplex implementations (see Bazaraa et al. (1990), for example). At any step in this process, while examining a node j with the basic arc $(i, j) \in B$ leading into node j (node i being the predecessor of node j), we perform the following update operation (a similar operation is performed if the corresponding arc is (*j*, *i*), leading out of node *j*):

$$q\max_{ij} = \min\left\{ q\max_{ij}, -b_j - \sum_{\substack{k\in RS(j)\\ k\neq i}} q\min_{kj} + \sum_{k\in FS(j)} q\max_{jk} \right\} \tag{4.1}$$

$$q\min_{ij} = \max\left\{ q\min_{ij}, -b_j - \sum_{\substack{k\in RS(j)\\ k\neq i}} q\max_{kj} + \sum_{k\in FS(j)} q\min_{jk} \right\} \tag{4.2}$$

Hence, as the flow interval lengths for nonbasic arcs shrink to zero, so do the corresponding interval lengths for the basic arcs since the right-hand sides of (4.11) and (4.12) coincide in this case at each step.

4.2. BRANCHING VARIABLE SELECTION STRATEGY

The recommended rule combines a pair of strategies based primarily on the fact that the polyhedral approximations in the (q, λ) space are less exact when the flow is distant from both of its bounds. Hence, when partitioning any node of the branch-and-bound tree, the base-strategy is to select the branching variable (*r*, *s*) according to

$$(r, s) \in \text{argmax}\{\min\{q\max_{ij} - \hat{q}_{ij}, \hat{q}_{ij} - q\min_{ij}\} : (i, j) \in A - B\}, \tag{4.3}$$

where *B* is the set of tree arcs in the MSTR procedure described above, and $\hat{q}$ is the flow solution produced by solving the lower bounding relaxation. The bounding intervals for q_{rs} in the two children node subproblems are then taken as $[q\min_{rs}, \hat{q}_{rs}]$ and $[\hat{q}_{rs}, q\max_{rs}]$, respectively.

However, even if the nonbasic flow values are driven close to either their lower or upper bounding interval end-points as the algorithm progresses, if the implied interval bounds on the basic arcs as determined by (4.11) and (4.12) are not sufficiently tight, the same situation may not be the case for these basic arcs, thereby making the process stall using this strategy. Hence, when a substantial improvement in the global lower bound is not obtained for any pair of successive stages, i.e., if at any stage $S + 1$, it turns out that $0.9GLB_{S+1} \leq GLB_S$, then at this stage, we adopt an alternative branching variable selection strategy. As explained below, this also induces convergence to a global optimum.

The proposed switchover strategy is based on simply partitioning the longest flow interval among the nonbasic arcs, i.e., the branching variable index (*r*, *s*) is selected as

$$(r, s) \in \underset{(i,j)}{\text{argmax}}\{q\max_{ij} - q\min_{ij} : (i, j) \in A - B\}. \tag{4.4}$$

Having selected (*r*, *s*) according to (4.14) in this case, we bi-partition the interval $q\min_{rs}, q\max_{rs}]$ by cutting it at the value 0 if $q\min_{rs} < 0 < q\max_{rs}$, or by bisecting this interval, otherwise. We found this combination strategy to be more effective than using either scheme on its own, or using the strategy prescribed in Sherali et al. (1998) (see Subramanian (1999) for detailed comparative results).

To see how this induces convergence, let us define δ_{ij}, for each $(i, j) \in A$, as the discrepancy in the actual (Equation (2.3b)) versus the approximate (Equation (3.11b)) head-loss computation relative to the solution produced by the relaxed solution, and denote $\Delta = \max\{\delta_{ij} : (i, j) \in A\}$. Note that if $\Delta = 0$ (empirically measured by $\Delta = 10^{-6}$ in our implementation), then we will have achieved an optimal solution to the current node subproblem, since then, the solution ($\boldsymbol{q}, \boldsymbol{H_s}, \boldsymbol{H}, \boldsymbol{x}$) where $\mathbf{x} \equiv \mathbf{x}^1 + \mathbf{x}^2$, that is produced by the relaxation LB(Ω) (or RLT(Ω)) will be feasible to NOP(Ω) as defined in (2.3), yielding the same objective value. Hence, this node can be fathomed after updating the incumbent solution. Moreover, when $q\text{min}_{ij} = q\text{max}_{ij}\ \forall (i, j) \in A - B$, we will also have $q\text{min}_{ij} = q\text{max}_{ij}\ \forall (i, j) \in B$ by (4.1) and (4.2), and then as shown in Sherali et al. (1998), this will imply that the condition $\Delta = 0$ holds true, thereby indicating optimality for the current subproblem. Hence, if we examine any path of the branch-and-bound tree associated with selecting the node having the least lower bound at each stage, since the condition $0.9\ GLB_{S+1} > GLB_S$ can hold only finitely often along this path, this induces a process whereby $\{q\text{max}_{ij} - q\text{min}_{ij}\} \to 0\ \forall (i, j) \in A - B$, leading as above to $\Delta \to 0$ along such a path. Therefore, a global optimum is recovered in the limit.

5. Computational Experience

In this section, we apply the proposed branch-and-bound algorithm to three standard test problems from the literature, and a newly generated Blacksburg network. The algorithms were implemented on a SUN SPARC 10 UNIX workstation, using the CPLEX 6.0 callable library to solve the linear programming problems. The computer code was written in C++. The algorithm was implemented using a termination criterion of $\epsilon = 10^{-6}$. In addition, we also experimented with using a lesser number of supporting hyperplanes (four) in lieu of six as discussed in Section 3 for constructing the lower bounding linear programs, with and without RLT enhancements. The computational results and the best design configuration for each of these four networks are presented sequentially below.

Test Problem 1: Two-Loop Network

This is a single source test problem originally presented by Alperovits and Shamir (1977). (See Sherali et al. (1998) for a description of the network configuration and for the set of initial bounds on the flows. The latter were logically determined as previously stated in Section 3.) The set of commercially available pipes were taken as having diameters *d* (inches) given by {1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24}, with the corresponding costs per unit length ($/meter) being {2, 5, 8, 11, 16, 23, 32, 50, 60, 90, 130, 170, 300, 550}. Note that the original Alperovits and Shamir

Table Ia. Computational results for the Two-Loop Network

n_s	Global Lower Bound	Global Upper Bound	# LP solved	CPU time (sec)
4	403385.10	403385.40	141	11.64
6	403385.09	403385.41	147	18.05

Table Ib. Optimum design for the Two-Loop Network ($\epsilon = 10^{-6}$)

Arc #	Segments having Length (m)	Flow (m^3/hr)	Diameter (inches)	Head-Loss (m)
1	1000	1120.0	18	6.749242
2	795.408540	368.331588	10	11.982436
2	204.591460	368.331588	12	1.268322
3	1000	651.668412	16	4.393325
4	999.998776	0.975686	1	18.857409
4	0.001224	0.975686	2	0.000001
5	310.354145	530.692726	14	1.786092
5	689.645855	530.692726	16	2.071341
6	11.139149	200.692726	8	0.161574
6	988.860851	200.692726	10	4.838426
7	98.493861	268.331588	8	2.446478
7	901.506139	268.331588	10	7.553516
8	1000	−0.692726	1	−10.000004

(1977) test problem excludes certain pipe diameters, whereas several authors have later solved this problem by including all the possible aforementioned diameters. To enable a comparison, as well as from a practical viewpoint, we permit the selection of all commercially available pipe diameters. Table 1a summarizes the results obtained using different numbers of supports (n_s) per arc, and Table 1b provides the key information regarding the best design obtained.

Among the more recent results on this problem, the heuristic of Loganathan et al. (1995) found a solution having a total cost of \$403,657. Sherali et al. (1998) obtained a somewhat improved solution with an objective value of \$403,390. Earlier, Sherali and Smith (1997) had recently obtained a global lower bound of \$403,385 on this problem, along with a feasible solution of \$403,386, which is \$1 within global optimality. Their algorithm, when implemented on the same computer and using CPLEX 2.0 to solve the LP relaxations, enumerated only 49 nodes, but consumed 342 CPU seconds due to the size of their lower bounding problem. The best

Table IIa. Computational results for the Hanoi network

n_s	Global Lower Bound	Global Upper Bound	# LP solved	CPU time
4	6055536.43	6055542.48	1533	245.44 sec
6	6055536.31	6055542.37	1075	552.82 sec

solution presented in Table 1b is obtained using our branch-and-bound procedure with four supporting hyperplanes per arc. This solution is the most accurate one reported in the literature, and has an objective value that lies within $0.3 of global optimality, and was derived while consuming only 12 CPU seconds. We comment that the use of the procedure MSTR resulted in a reduction in computational effort by a factor of 5, hence underscoring the usefulness of this scheme.

Test Problem 2: Hanoi Network

The Hanoi network is a single source network consisting of three basic loops, thirty-two nodes and thirty-four links. The network configuration, arc definitions, and flow bounds are given in Sherali et al. (1998), and the other data appears in Fujiwara and Khang (1990). Table 2a presents the results obtained using our algorithm, and Table 2b provides the key design parameters for the best solution reported.

Eiger et al. (1994) reported a solution having an objective function value of $6,026,660 for this problem using an optimality tolerance of 0.5%. However, their solution contains some violations in the flow conservation constraints, as shown by Sherali et al. (1998), who obtained a solution having an objective value of $6,058,976, which is the best solution previously reported in the literature. The best solution found by our algorithm has an objective value of $6,055,542 which is significantly better than the values reported in the literature for this test problem. This solution was obtained within 4 minutes of CPU time, using four supporting hyperplanes per arc in LB(Ω), and is within 10^{-4}% of optimality (or $6 of global optimality) as verified by our global lower bound.

The original data for the Hanoi test network presented in Fujiwara and Khang (1990) used a C_{HW} value of 162.5. For the sake of comparison, the above run was repeated using this value for the Hazen-William coefficient. A global lower bound of 4,954,941.69 and a corresponding feasible solution having an objective cost of 4,954,945.29 was obtained after solving 541 linear programs and expending 2 minutes of CPU time. This solution is within $4 of global optimality and significantly improves upon the best objective cost of 5,562,000 reported in Fujiwara and Khang (1990).

Table IIb. Optimum design for the Hanoi network ($\epsilon = 10^{-6}$)

Arc #	Dia	Length (m)	Flow (m^3/hr)	Head-Loss (m)	Arc #	Dia	Length (m)	Flow (m^3/hr)	Head-Loss (m)
1	40	100	19940	2.859823	19	24	400	2403.036233	2.734511
2	40	1350	19050	35.477043	20	40	2200	7831.762902	11.145477
3	40	900	7965.200865	4.704430	21	16	491.360368	1415	9.074382
4	40	1150	7835.200865	5.830782	21	20	1008.639632	1415	6.283494
5	40	1450	7110.200865	6.141881	22	12	500	485	5.159782
6	40	450	6105.200865	1.437398	23	40	2650	5141.762902	6.158480
7	40	850	4755.200865	1.709164	24	30	1230	3501.070442	5.694622
8	40	850	4205.200865	1.361194	25	30	1300	2681.070442	3.671724
9	30	74.233811	3680.200865	0.376960	26	20	850	1186.762902	3.823008
9	40	725.766189	3680.200865	0.907902	27	12	299.999150	286.762902	1.169823
10	30	950	2000	1.559305	27	16	0.000850	286.762902	0.000001
11	24	1200	1500	3.427313	28	12	750	83.237098	0.295907
12	24	3500	940	4.206805	29	16	1500	595.692460	5.580180
13	16	253.562337	1155.200865	3.216195	30	12	1999.999879	305.692460	8.778997
13	20	546.437663	1155.200865	2.338011	30	16	0.000121	305.692460	0
14	16	500	540.200865	1.551952	31	12	1600	−54.307540	−0.286252
15	12	550	260.200865	1.791357	32	16	150	−414.307540	−0.284832
16	12	2730	−133.036233	−2.566981	33	16	748.166501	−519.307540	−2.158642
17	16	1750	−998.036233	−16.930625	33	20	111.833499	−519.307540	−0.108844
18	20	419.803213	−2343.036233	−6.654887	34	24	950	1324.307540	2.154262
18	24	380.196787	−2343.036233	−2.480223					

Table IIIa. Computational results for the New York network ($\epsilon = 10^{-6}$)

n_s	Global Lower Bound	Global Upper Bound	# LP solved	CPU time
4	37878580.89	37878581.28	12953	93 min
6	37878580.89	37878581.28	10957	58 min

Test Problem 3: New York Network

The New York test network configuration and data are given in Loganathan et al. (1985) and Fujiwara and Khang (1990). The original network has 20 nodes and 21 arcs, while the expanded network (with parallel arcs) has 26 nodes and 33 arcs (see Subramanian (1999) for details). Since there is only a single source node, the initial flow bounds for the arcs were calculated using the procedure described in Sherali and Smith (1997). The computational results and the best design obtained are presented in Tables 3a and 3b, respectively. The coefficients used in the head-loss equation are the same as that used in Fujiwara and Khang (1990) and Loganathan et al. (1985). The flow rate exponent value was set equal to 1.85, while the head-loss coefficient was set equal to 851500 to conform with the flow rates being measured in cubic-feet per second, the pipe diameters in inches, and the head-losses in feet.

The New York test network problem was first analyzed using parallel links by Lai and Schaake (1969) and they obtained a solution having an objective value of $77.61($10^6$). Fujiwara and Khang (1990) applied their two-phase approach to this problem, but the solution presented by them was infeasible. Quindry et al. (1981) obtained a solution having a total cost of $63.581($10^6$), while Gessler (1985), Bhave (1985), and Morgan and Goulter (1985) obtained solutions having costs of $41.2($10^6$), $40.18($10^6$), and $39.018($10^6$), respectively. Loganathan et al. (1995) used a simulated annealing based heuristic procedure to further improve the objective value to $38.04($10^6$). All these approaches are heuristic in nature and simply seek to determine (at best) local optimal solutions, providing no indication of a competitive global lower bound on the optimum value. Using our procedure, we were able to obtain such a global lower bound of value 37878580.89 and a feasible solution having a cost of 37878581.28. This is the best solution reported thus far in the literature and lies within 10^{-6}% (or within $0.4) of optimality. Note that this result is obtained by considering 4-inch pipe increments for the set of available pipe diameters. All the results presented above since 1985 use 12-inch pipe diameter increments for the sake of computational ease. Hence, for the sake of comparison, our procedure (using six supports per arc) was run using 12-inch pipe increments. The results produced a global lower bound value of 38,067,895 along with a corresponding feasible solution of 38,067,935, after solving 2467 linear programs. It was observed that while the optimal pipe diameters coincided with that obtained by

Loganathan et al. (1995), the corresponding pipe segments lengths were different. This is due to the fact that the head-loss values obtained by Loganathan et al. (1995) have a feasibility tolerance of 0.1, while the results presented in this paper are obtained using a more precise feasibility tolerance value of 10^{-6}. Consequently, the objective cost corresponding to the optimal solution is higher than that obtained by Loganathan et al. (1995), and represents a relatively more accurate estimate of the actual solution.

It can be seen from Table 3a that the computational times are significantly higher for this test case, as compared with the computational efforts for the previous two test problems. One important reason for this difference is that no logical test based schemes were used to generate tight initial flow interval bounds. In fact, the initial feasible solution was itself near-optimal, but was polished to the final solution only toward the tail-end of the branching procedure. The results for this test problem also differ from the previous two in that the introduction of additional hyperplanes results in an improvement in the computational effort, both in terms of the number of nodes enumerated and the CPU time expense.

Test Problem 4: Blacksburg Test Network

Figure 2 depicts a network representation of a newly expanded subdivision of the water distribution system in the town of Blacksburg, Virginia. The network data for this problem was acquired from the public works department of the town, along with other problem parameters such as pressure requirements, locations of fire hydrants, cost factors, pipe quality that is reflected via the associated C_{HW} value, and demand requirements. The unit pipe costs used in this problem represent real-life values and include installation costs as well. The link and node data for the network are presented in Tables 4a and 4b. The set of pipes whose diameters are fixed is listed in Table 4a. A Hazen-Williams coefficient value of 120 was used for all the links. Expression (2.1) was used to compute the head-losses. The flow rates were converted using double precision from gallons per minute (gpm) into units of m^3/hr, the pipe diameters were specified in centimeters, and the head-losses in meters. The computational results and the best design obtained are presented in Tables 4c and 4d, respectively. The optimal flow values can be computed using the optimal head-loss and pipe diameter values given in Table 4d via Equation (2.1) (using appropriate coefficients for the measurement units specified above).

The computational results for the Blacksburg network also exhibit that the introduction of additional supporting hyperplanes (six instead of four per arc), results in a decrease in the number of nodes enumerated, as well as in the CPU time expended. The best solution obtained for this test problem has an objective value of 577066, along with a best global lower bound of 577066.

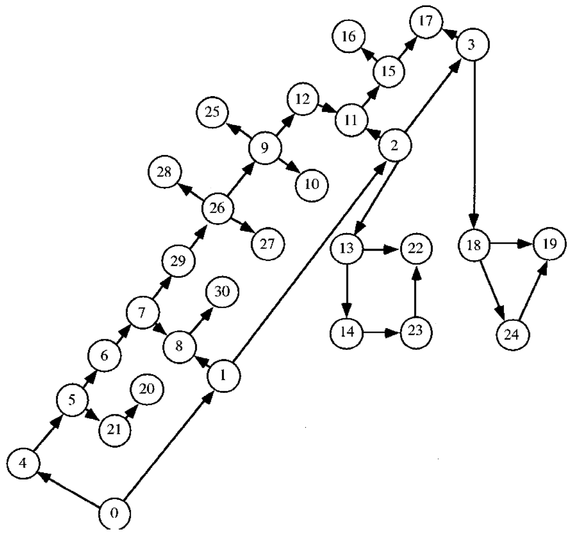

Figure 2. Blacksburg test network configuration.

Discussion on Algorithmic Strategies and the Use of RLT(Ω)

The computational experience on the foregoing test problems clearly indicated that the maximal spanning tree reduction procedure (MSTR) is an indispensable strategy. In fact, for the two larger New York and Blacksburg test problems, when we suppressed the procedure MSTR, we were unable to obtain good quality feasible solutions within the time limit of 10 CPU hours. Furthermore, the use of six versus four hyperplanes per arc in the polyhedral approximation of the flow relationships improved the relative performance for these two larger test problems, while for the other two test problems, it resulted in only a marginally greater CPU time expense.

However, in the cases where the introduction of additional supports was beneficial, a significant reduction in computational time was observed. Hence, we recommend the use of the six prescribed supporting hyperplanes per arc in the lower bounding problem, but suggest experimenting with additional supports.

As far as the use of the lower bounding scheme RLT(Ω) is concerned, it was generally observed that this yielded much tighter lower bounds, and resulted in fewer branch-and-bound nodes being enumerated as compared with using LB(Ω). On the other hand, the compromise with respect to computational effort was not favorable for these test cases, although as problem size increased, the relative benefit of using the tighter lower bounding formulation RLT(Ω) became more pronounced. Table 5a illustrates this phenomenon. However, when we switched off the strategy MSTR, as seen from Table 5b for the Hanoi network, the RLT scheme enumerated significantly fewer nodes and also consumed lesser CPU time, in comparison with using the relaxation LB(Ω). In general, it is apparent that the relative computational

Table IIIb. Optimum design for the New York network ($\epsilon = 10^{-6}$)

Link index	**Dia (inches)**	**Segment length (ft)**	**Flow (gpm)**	**New Cost ($)**	**Head-Loss (ft)**
1	180	11600	879.302755	7987090.05	5.729565
2	180	19800	786.902755	13633136.46	7.963917
3	180	7300	694.502755	5026358.39	2.330383
4	180	8300	606.302755	5714900.64	2.060921
5	180	8600	518.102755	5921463.31	1.596524
6	180	19100	429.902755	13151156.89	2.510590
71	132	4800	201.677399	2249808.73	0.704444
72	132	4800	201.677399	2249808.73	0.704444
7	112	2408.899652	140.025356	920962.07	0.400677
7	116	7191.100348	140.025356	2871547.08	1.008212
8	132	12500	253.502755	5858876.90	2.800706
9	180	9600	58.500000	6610005.56	0.031515
10	204	11200	131.124553	9006418.33	0.088965
11	204	14500	484.997245	11660095.17	1.295005
12	204	12200	836.297245	9810562.83	2.985471
13	204	24100	953.397245	19379882.31	7.515528
14	204	21100	1045.797245	16967448.83	7.808112
15	204	15500	1138.197245	12464239.66	6.708414
161	72	13200	17.039827	2917822.79	0.383495
162	72	13200	17.039827	2917822.79	0.383495
16	100	26399.999870	40.460173	8769939.75	0.766990
16	104	0.000130	40.460173	0.05	0.000000
171	72	15600	69.403955	3448336.02	6.090610
172	72	15600	69.403955	3448336.02	6.090610
17	96	0.000063	164.796045	0.02	0.000000
17	100	31199.999937	164.796045	10364474.28	12.181220
181	60	12000	38.897986	2115833.46	3.900627
182	60	12000	38.897986	2115833.46	3.900627
18	76	9745.580952	78.202014	2303611.69	3.646560
18	80	14254.419048	78.202014	3590655.93	4.154695
191	60	7200	64.220140	1269500.08	5.917147
192	60	7200	64.220140	1269500.08	5.917147
19	72	0.195798	119.652552	43.28	0.000209
19	76	14399.804202	119.652552	3403753.70	11.834085
20	60	38400	13.872692	6770667.09	1.853176
211	72	13200	80.619266	2917822.79	6.799253
212	72	13200	80.619266	2917822.79	6.799253
21	68	10595.684537	75.508042	2181889.52	6.386791
21	72	15804.315463	75.508042	3493499.38	7.211714

Total cost	Existing cost	New cost
217,700,927	179,822,346	37,878,581

Table IVa. Arc data for the Blacksburg network

Arc	**Length (ft)**	**Fix dia. (inches)**	**Arc**	**Length (ft)**	**Fix dia. (inches)**	**Arc**	**Length (ft)**	**Fix dia. (inches)**
(0, 1)	1363	–	(6, 7)	95	–	(15, 17)	1009	–
(0, 4)	194	–	(7, 8)	419	–	(18, 19)	408	–
(1, 2)	1832	–	(7, 29)	208	–	(18, 24)	1181	–
(1, 8)	151	–	(8, 30)	110	–	(21, 20)	113	–
(2, 3)	888	–	(9, 10)	451	6	(22, 13)	701	6
(2, 11)	155	–	(9, 12)	59	–	(23, 22)	351	6
(2, 13)	309	–	(9, 25)	416	6	(24, 19)	967	6
(3, 17)	699	–	(11, 15)	303	12	(26, 9)	271	–
(3, 18)	1151	–	(12, 11)	823	–	(26, 27)	317	6
(4, 5)	1098	–	(13, 14)	766	8	(26, 28)	424	6
(5, 6)	578	–	(14, 23)	382	–	(29, 26)	730	6
(5, 21)	611	6	(15, 16)	758	10			–

Table IVb. Node data for the Blacksburg network

Node index	**Supply or demand (gpm)**	**Elevation (ft)**	**Node index**	**Supply or Demand (gpm)**	**Elevation (ft)**
0	1548.63	2163	16	−52.11	2149
1	−52.11	2141	17	−10.38	2109
2	−50.58	2132	18	−103.65	2144
3	−25.77	2121	19	−52.11	2149.5
4	−13.84	2153.5	20	−10.96	2140
5	−53.65	2141.5	21	−51.35	2141.5
6	−51.73	2129	22	−11.73	2144
7	−200.58	2127	23	−51.54	2156.5
8	−11.35	2127	24	−102.50	2178
9	−10.77	2109.5	25	−51.54	2118
10	−52.11	2121	26	−52.50	2099.5
11	−100.77	2139	27	−50.96	2102
12	−27.11	2110	28	−25.58	2098.5
13	−100.77	2136.5	29	−41.92	2120
14	−25.77	2143	30	−51.35	2123
15	−51.54	2144.5			

Table IVc. Computational results for the Blacksburg network

n_s	**Global Lower** Bound	**Global Upper** Bound	**# LP** solved	**CPU time**
4	577066	577067	2089	27 min
6	577066	577067	1975	24 min

efficiency of the RLT-enhanced scheme improves as the size of the problem ($|A|$, $|N|$, K) and the number of possible choices for selecting the branching variable increase. This suggests that for problems having several more independent variables (networks having more looped structures for reliability purposes) than those analyzed in this paper, such an RLT-enhanced procedure might be beneficial. Furthermore the relative advantage of the RLT approach might become favorable with further advances in linear programming technology. This is open to future investigation.

6. Summary and Conclusions

In this paper, we have proposed enhanced lower bounding procedures along with significantly more effective branching and partitioning strategies for determining global optimal solutions to water distribution network design problems. A new test problem dealing with the water distribution system in Blacksburg, Virginia, is also introduced to the literature in this paper. Results obtained on this as well as three other standard lems from the literature demonstrate the efficacy of the proposed methodology. Improved solutions are reported for each of the latter problems, significantly so for the two larger cases of the Hanoi and the New York test networks for which solutions proven to lie within 10^{-4}% of optimality are derived for the first time in the literature. Further enhancements in algorithmic efficiency can be achieved by including a more effective preprocessor to deduce valid, tighter initial bounds on the flow variables. The algorithm can also benefit by computing sharper upper bounds by using some local optimization scheme, rather than simply evaluating the flow solution produced by the lower bounding problem. The application of efficient schemes such as those described in Sherali and Smith (1997) to obtain tight flow bounds or upper bounds for each node subproblem is more critical in the case of the network design problems having several independent variables. Such problems can also benefit via the construction of tighter lower bounding problems through the use of an additional, suitable number of supporting hyperplanes in the approximation of the flow relationships, once the direction of flow in any link is determined, as well as through the proposed RLT constructs. Such investigations and further computational tests are proposed for future research.

Table IVd. Optimum design for the Blacksburg network ($\epsilon = 10^{-6}$)

Arc	**Dia (inches)**	**Length (ft)**	**Arc**	**Dia (inches)**	**Length (ft)**
(13, 14)	8.0	766	(18, 24)	10.0	0.001887
(23, 22)	6.0	351	(18, 24)	12.0	1180.998113
(26, 28)	6.0	424	(7, 29)	10.0	208
(14, 23)	4.0	63.645446	(9, 25)	6.0	416
(14, 23)	6.0	318.354554	(9, 10)	6.0	451
(9, 12)	10.0	59	(26, 9)	6.0	92.889470
(2, 3)	16.0	888	(26, 9)	8.0	178.110530
(3, 18)	16.0	1151	(22, 13)	6.0	701
(12, 11)	12.0	823	(0, 4)	16.0	194
(11, 15)	12.0	303	(21, 20)	2.0	12.848676
(15, 16)	10.0	758	(21, 20)	3.0	100.151324
(0, 1)	24.0	1363	(1, 8)	3.0	43.747654
(2, 11)	12.0	69.689856	(1, 8)	4.0	107.252346
(2, 11)	16.0	85.310144	(5, 6)	12.0	577.999064
(2, 13)	10.0	309	(5, 6)	16.0	0.000936
(24, 19)	6.0	967	(26, 27)	6.0	317
(8, 30)	4.0	0.000108	(29, 26)	6.0	730
(8, 30)	6.0	109.999892	(4, 5)	16.0	1098
(1, 2)	20.0	13.727890	(5, 21)	6.0	611
(1, 2)	24.0	1818.272110	(3, 17)	1.0	698.988391
(15, 17)	3.0	714.434858	(3, 17)	2.0	0.011609
(15, 17)	4.0	294.565142	(6, 7)	12.0	95
(18, 19)	10.0	408	(7, 8)	4.0	419

Node #	**Head (ft)**	**Node #**	**Head (ft)**
0	184.660000	16	93.308922
1	168.722421	17	46.160000
2	134.430362	18	90.189925
3	129.949917	19	80.263327
4	185.105409	20	46.160000
5	148.536939	21	72.661969
6	99.580936	22	66.147108
7	94.038161	23	46.160000
8	51.706685	24	46.160000
9	97.880803	25	53.029402
10	46.160000	26	75.785923
11	112.829253	27	46.160000
12	102.771162	28	66.662345
13	104.838176	29	97.379870
14	80.826940	30	46.160000
15	103.426382		

Table Va. Comparative results between using LB(Ω) versus RLT(Ω) for $\epsilon = 10^{-3}$

Network	**Nodes (RLT(Ω))/ Nodes (LB(Ω))**	**Time (RLT(Ω))/ Time (LB(Ω))**
Two-Loop	97/99 = 0.98	33/8 = 12.5
Hanoi	115/167 = 0.69	167/51 = 3.27
New York	6871/13673 = 0.50	255/71 = 3.59
Blacksburg	2734/3445 = 0.81	51/12 = 4.25

Table Vb. Comparative results between using LB(Ω) versus RLT(Ω) without the procedure MSTR for $\epsilon = 10^{-3}$

Network	**Nodes (RLT(Ω))/ Nodes (LB(Ω))**	**Time (RLT(Ω))/ Time (LB(Ω))**
Two-Loop	521/887 = 0.59	109/61 = 1.79
Hanoi	2905/15681 = 0.19	49/65 = 0.75

Acknowledgments

This material is based upon research supported by the National Science Foundation under Grant Number DMI-9812047, the United States Geological Survey under Grant Number 1434-HQ-96-6R-02703, and the Virginia Water Resources Research Center.

References

Alperovits, E. and Shamir, U. Design of Optimal Water Distribution Systems, *Water Resources Research* 13: (1977), 885–900.

Bazaraa, M.S., Jarvis, J. J. and Sherali, H.D. (1990), *Linear Programming and Network Flows*, second edition, John Wiley and Sons, New York, NY.

Bhave, P.R. (1983), Optimization of Gravity-Fed Water Distribution Systems: Theory, *Journal of Environmental Engineering*, ASCE, 109(1), 189-205.

Bhave, P.R. (1985), Optimal Expansion of Water Distribution Systems, *Journal of the Environmental Engineering Division*, ASCE 111, No. EE2: 177–197.

Dandy, G.C., Simpson, A.R. and Murphy, L.J. (1996), An Improved Genetic Algorithm for Pipe Network Optimization, *Water Resources Research*, 32(2): 449–458.

Eiger, G., Shamir, U. and Ben-Tal, A. (1994), Optimal Design of Water Distribution Networks, *Water Resources Research* 30(9): 2637–2646.

Fujiwara, O. and Khang, D.B. (1990), A Two-Phase Decomposition Method for Optimal Design of Looped Water Distribution Networks, *Water Resources Research* 26(4): 539–549.

Gessler, J.M. (1985), Pipe Network Optimization by Enumeration, in H.C. Tonro (ed.), *Computer Applications in Water Resources*, pp. 527–581.

Kruskal, J, B. (1956), On the Shortest Spanning Tree of a Graph and the Traveling Salesman Problem, *Proceedings of the American Mathematical Society* 7: 48–50.

Lai, D. and J. Schaake, (1969), Linear Programming and Dynamic Programming Applications to Water Distribution Network Design. Report 116, Department of Civil Engineering, MIT, Cambridge.

Lansey, K. and Mays, L. (1985), A Methodology for Optimal Network Design, *Computer Applications in Water Resources*, 732–738.

Loganathan, G.V., Greene, J.J. and Ahn, T.J. (1995), Design Heuristic for Globally Minimum Cost Water-Distribution Systems, *Journal of Water Resources Planning and Management* 121(2): 182–192.

Morgan, D.R. and Goulter, I.C. (1985), Optimal Urban Water Distribution Design, *Water Resources Research*, 21(5): 642–652.

Quindry, G., Brill, E.D. and Liebman, J.C. (1981), Optimization of Looped Water Distribution Systems, *Journal of Environmental Engineering Division*, ASCE, 107, EE4: 665–679.

Rowell, W.F. and Barnes, J.W. (1982), Obtaining Layouts of Water Distribution Systems, *Journal of the Hydraulics Division*, ASCE, 108, HY1: 137–148.

Sherali, H.D. and Smith, E.P. (1993), An Optimal Replacement-Design Model for a Reliable Water Distribution Network System, *Integrated Computer Applications in Water Supply* 1: 61–75. Research Studies Press Ltd., Somerset, England.

Sherali, H.D. and Smith, E.P. (1997), A Global Optimization Approach to a Water Distribution Network Design Problem, *Journal of Global Optimization*, 11: 107–132.

Sherali, H.D., Totlani, R. and Loganathan, G.V. (1998), Enhanced Lower Bounds for the Global Optimization of Water Distribution Networks, *Water Resources Research* 34(7): 1831–1841.

Sherali, H.D. and Tuncbilek, C.H. (1992) A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique, *Journal of Global Optimization* 2: 101–112.

Subramanian, S. (1999), Optimization Models and Analysis of Routing, Location, Distribution, and Design Problems on Networks. Doctoral Dissertation, Virginia Polytechnic Institute and State University.

Walski, T.M. (1984), *Analysis of Water Distribution Systems*, Van Nostrand Reinhold Company, New York, NY.