A Heuristic Method for the Optimal Design of

Water Distribution Systems

by

Mahesh Shah

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Civil Engineering

APPROVED:

Dr. G. V. Loganathan, Chairman

Dr. C. Y. Kuo

Dr. J. M. Wiggert

September, 1987

Blacksburg, Virginia

A Heuristic Method for the Optimal Design of

Water Distribution Systems

by

Mahesh Shah

Dr. G. V. Loganathan, Chairman

Civil Engineering

(ABSTRACT)

The water distribution system design problem consists of finding a minimum cost combination of network layout and sizes of system components so as to satisfy flow demands, minimum and maximum head requirements and a reliability criterion. A two step procedure is proposed to find a near optimal design. The first step considers obtaining a near optimal tree layout using a heuristic tree search algorithm. Two different methods are followed for the tree search - one for single source networks and the other for multiple source networks. The second step adds loop forming redundant links to the tree layout in such a way that every demand node has two paths from source node(s). The methodology is applied to a single source network and a multiple source network. In both the cases better results are achieved than those obtained previously by other researchers.

Acknowledgements

This work would not have been possible without the guidance and support of my principal advisor, Dr. G. V. Loganathan. His sharp thinking, keen insights and encouragement kept me going when everything seemed bleak. The training he has provided will stand me in good stead in my future pursuits.

I would also like to express my sincerest gratitude to Dr. C. Y. Kuo and Dr. J. M. Wiggert for reviewing the manuscript. Their constructive criticisms and helpful suggestions have gone a long way in improving the quality of this thesis.

My thanks are also due to all my friends at Virginia Tech who have helped me from time to time. In particular, I would like to thank Hemang for spending his valuable time in helping me with the computer programs. The help extended by Asif and Nathan in proof reading is also appreciated.

Last but not the least, I would like to thank my parents whose moral support has been a continuing source of encouragement in pursuing my goals.

Table of Contents

List of Illustrations

List of Tables

List of Symbols

Symbol	*Chapters*	*Definition*
A	2,3	Adjacency matrix
A	1,2	Parameter in pump head equation
adj(i)	2	Set of nodes adjacent to node i
a_{ij}	2,3	Element of adjacency matrix
B	1,2	Parameter in pump head equation
C	2,4	$\alpha C_2/\beta$
CT	2,3	Set of cotree links
C_1	2,4	Cost parameter in link cost function
C_2	2,4	Exponent in link cost function
C_3	2,4	Cost parameter in pump cost function
C_{hw}	1	Hazen-Williams coefficient
C_{Lk}	2	Cost of kth diameter pipe per unit length
C_P	2	Cost of pumping unit head
C_S	2	Cost of elevating reservoir by unit head
C_T	2	Cost of tree network T
C_Y	2	Set of nodes on a loop
$C_{(i,j)}$	3	Cardinality of reconnecting set $RC_{(i,j)}$
D	1	Diameter of a pipe

Symbol	*Chapters*	*Definition*
D_I	2	Set of available diameters $\{d_1, \; d_2, \ldots, \; d_{N_D}\}$
$D_{(i,j)}$	1,2	Diameter of link (i,j)
deg(i)	2	Degree of node i
F	1	Vector of constants for continuity and energy equations
F_i	2	Fixed head at node i
F_L	2	Link cost function
F_P	2	Pump cost function
F_S	2	Storage cost function
f	1	Darcy-Weischbach friction factor
f	1	Number of fixed grade (head) nodes
H_0	1,2	Parameter in pump head equation
H_i	1,2	Head at node i
$\underline{H}_i$	2	Minimum head at node i
$\overline{H}_i$	2	Maximum head at node i
H_{Si}	2	Additional storage height
$H_{f(i,j)}$	1,2	Friction headloss in link (i,j)
$H_{L(i,j)}$	1,2	Total headloss in link (i,j)
$H_{m(i,j)}$	1,2	Minor headloss in link (i,j)
$H_{P(i,j)}$	1,2	Pump head in link (i,j)
h_i	1	Pressure head at node i
J	1	Jacobian of continuity and energy equations
J	2,4	Hydraulic gradient in a link
$K_{f(i,j)}$	1,2	Coefficient in friction headloss equation
$\overline{K}_{f(i,j)}$	2,4	$C_1[K_{f(i,j)}/J]^{C_2/\beta}$
k_i	1,2	Set of nodes adjacent to node i

Symbol	*Chapters*	*Definition*
$k_{i,CT}$	2	Set of nodes adjacent to node i in CT
L_p	2	Set of links on a loop
$L_{(i,j)}$	1,2,3	Length of link (i,j)
$\mathscr{L}$	1	Set of existing links
$\mathscr{L}$	2,3,4	Set of potential links
l	1	Number of links
N	1,2,4	Set of nodes
N_1, N_2	3	Two connected sets of nodes created by removal of a link from a tree
N_D	2	Number of available diameters
N_t	2	Number of demand patterns
$N_{(p,q)}$	3	Number of reconnecting sets in which link (p,q) occurs
n	1,2	Number of nodes
P	2	Set of links in which pumps can be located
P	3	Parallel link indicator
p	1	Number of independent cycles
Q	1	Vector of flows
$Q_{(i,j)}$	1,2	Flow in link (i,j)
q_i	1,2,4	External flow at node i (demand or supply)
R	3	Set of all reconnecting links (union of all reconnecting sets for a tree)
RL	3,4	Set of redundant links
R_{min}	2	Minimum level of reliability
$RC_{(i,j)}$	3,4	Reconnecting set for link (i,j)
R_s	2	Measure of system reliability
r	1	Resistance coefficient $= K_f L/D^{\beta}$

Symbol	*Chapters*	*Definition*
S	2,3	Set of source nodes
T	2,3	Set of tree links
T^*	2	Optimal tree
t	2	Index indicating demand pattern
v_i	1	Velocity of fluid at node i
$X_{(i,j)k}$	2	Length of diameter d_k in link (i,j)
Z_i	1	Elevation head at node i
α	1,2,4	Exponent of flow in friction head loss formula
β	1,2,4	Exponent of diameter in friction head loss formula
η	2	Pump efficiency
γ	2	Parameter in pump cost function
Φ	2,3	Empty set
$\in$	1,2,3,4	Member of (or belongs to)
$\forall$	1,2,3,4	For all
Σ	1,2,3,4	Summation
$\cup$	2,3	Union (of sets)
$\mid$	2,3	Such that
$\lvert X \rvert$	1,2,3	Absolute value Cardinality of X (number of elements in X)

1.0 Water Distribution Systems : An Overview

1.1 Introduction

A water distribution system can be defined as a utility service which taps water from nature and distributes it in a form suitable to the consumers through a network of interconnected facilities. Because of their size and complexity, the construction, operation and maintenance of water distribution systems require large capital outlays. Their impact on social and industrial growth is immense. A few facts about water distribution systems are in order :

- the demand for water is generally increasing because of population and industrial growth (Goodman, 1984).
- the timing and scale of investment decisions amounting to $15 billion per year depend on forecasts of future water use (Boland, 1980).
- excessive amount of water is being lost through leaks in various cities : 17% in Boston, 15% in St. Louis, 14% in Pittsburgh, 14% in Tulsa (Choate and Walter , 1981).
- corrosion caused by aggressive water not only carries the threat of distribution system deterioration but also poses a potential health hazard (Millette, et al., 1980).
- deterioration of aging water supply systems in old urbanized areas and expansion of existing systems pose special problems and rehabilitation costs are in the order of $75 to $110 billion in 1972 dollars (Mays and Cullinane, 1986).

The classical approach to steady state water distribution system problem involves the following subproblems : (i) planning, (ii) design and (iii) analysis. The *planning* aspect concerns itself with the determination of requirements of the system. At the planning stage, the water demand for the period under consideration is estimated. It also determines the level of service and the reliability that the system should provide. The traditional *design* involves the selection of topological layout and sizes of distribution system components. The *analysis* problem determines flows in each pipe (or heads at different nodes) for a network configuration, specified pipe diameters and the flow demands.

It is the design problem which is of particular concern in this study. Mays and Cullinane (1986) have criticized the conventional trial and error design approaches by pointing out the lack of optimality of the design obtained. Such procedures also fail to quantify the reliability of the system. There is a definite need for improving the existing methodologies to ensure better service at low cost. In the present study, the layout and reliability aspects are incorporated in the design. Specifically, the present study is an attempt in answering the following question :

Given Water demands at various locations, minimum and maximum allowable pressures and restrictions on locations of reservoirs, pumping stations and pipes

To Find Locations and sizes of pipes, pumping stations, and reservoirs such that the cost of the system is minimized subject to the constraints on flows and pressure heads; at the same time every demand node in the system has two paths from source node(s).

In this chapter, various aspects of planning, analysis and design of water distribution systems are discussed. A review of the past work is provided. The basic elements of the proposed methodology are summarized at the end of this chapter. Subsequent chapters describe the complete design methodology.

1.2 Considerations in Planning a Water Distribution System

A water distribution system should be able to meet the projected water demand for a specified planning period with a specified level of reliability. Water demand is estimated based on the population projection. Next, a reliable source of water is located and then a distribution system is designed to transport water from the source to the demand nodes. Often the quality of water at the source may not meet the drinking water standards and suitable treatment facilities should be provided. A detailed account of the above aspects is given in Linsley and Franzini (1979). In this study, the design of water distribution system is of main concern which requires estimates of spatially disaggregated demands and pressure requirements. In the sequel, some general guidelines with regard to the input parameters of a water distribution system are outlined :

1. A map of the water district is obtained showing contours and other characteristics like streets, buildings, etc., and the fire risks to be protected.

2. Key locations (or nodes) are identified depending on the demand. Demarcation of districts is generally done on the basis of natural bonds between them. The probable flow demands are determined on the basis of population of the service area and the land use type, e.g., commercial or residential (Goodman, 1984).

3. In determining service pressures, some rules of thumb are followed. Typically, service pressures are maintained between 30 and 100 psi, 40 to 50 psi provide good flow rates to 3 to 4 storeyed buildings. Municipalities typically require owners of tall buildings to install booster pumps in order to avoid the need for high pressures in the mains which can result in increased leaks and higher pumping costs.

4. Fire protection requirements in terms of pressures and flows are estimated at the nodes. The rate of flow for fire fighting is given as $Q = 1020\sqrt{P}(1 - 0.1\sqrt{P})$ in which Q is in (gallons per minute) GPM and P is the population in thousands (Davis and Sorensen, 1969). According to the recommendations of the National Board Of Fire Underwriters, a fire pressure (pressure required at the nozzle of a fire hose for protection against fire) of 75 psi is necessary where more than 10 buildings exceed 3 stories in height, 60 psi in less risky areas and 50 psi in thinly populated areas. Fire hydrants should be located every 300 ft. in the mercantile and industrial areas, and every 600 ft. in the residential areas, so that every building in the city limits will be within 500 ft. of a hydrant (Texas Water Utility Association, 1979).

5. Next, depending on the physical conditions like topography, street right of ways etc., all potential links (a link is defined as a segment connecting two nodes and has no branches) of the network are determined. The link lengths depend on how pipes are laid . Main pipes are laid along the streets at some standardized positions along the curbs. For streets more than 40 to 50 ft. wide, it is more economical to install mains behind the curb on each side of the street as this reduces lengths of service pipes. Some of the rules for laying of water mains are reproduced here from the Washington Suburban Sanitary Commission manual (WSSC, 1981) :

   - water mains shall be laid 7 ft. from the centerline of the street right-of-way on the side of the street on high ground.

- the curvature of the centerline of the mains shall not exceed an allowable maximum . The allowable values depend on diameter and material of the pipe. These values are given in WSSC (1981).
- normal cover over water mains shall be a minimum of 4 ft. from the finished grade or the existing grade, whichever is lower.
- minimum clearance of 1 ft. above the sewer shall be kept.

Another important aspect in the water distribution system design is the level of service provided when failures occur which is assessed in terms of the reliability of the system. Reliability is generally defined as the probability that a system performs its mission within specified limits for a given period of time in a specified environment. For a large system like the water distribution system, the calculation of overall reliability depends on the reliability of its subsystems like pipes, pumps, reservoirs etc.. In order to predict water distribution system reliability and help make maintenance decisions, theories are required to estimate the component failures quantitatively. Damelin, et al., (1972) have developed a simulation model to evaluate the reliability of supply systems in which shortfalls are caused by the failure of pumping equipment. Shamir and Howard (1981) define a reliability factor in terms of the relative capacity lost due to failures and derived the probability distribution function (pdf) of the reliability factor based on the pdf of lost capacity. Such a pdf involves desired system capacity as a parameter and can be determined for a chosen reliability value.

With regard to pipe line failures, the reasons for failures are the soil movement, corrosion, temperature, stresses due to overburden and improper laying. Clark, et al. (1982) have conducted a repair frequency analysis of leaks and breaks for two systems and concluded that the maintenance rate of a pipe increases exponentially with time. Two regression equations were developed to relate the leak events to specific parameters like pipe type, diameter, soil type and pressure. They conclude that metallic pipes are more reliable than reinforced concrete pipes and that larger diameter pipes have a longer period before the first break. Kettler and Goulter (1985) point out that pipes with increased wall thickness perform better against corrosion. They also conclude that predominant mode of failure for asbestos cement pipes is circumferential cracking while cast-iron pipes mainly break at joints.

The reliability analysis under pipe failures is of great interest recently. Morgan and Goulter (1985) define a reliable network as the one which has two independent paths from the sources to the demand nodes and the problem becomes deterministic. Andreou, et al. (1987a, 1987b) have presented a methodology for analyzing random pipe failures. They use a Proportional Hazard model to represent the early stage of few breaks and Poisson model to describe later stages of multiple breaks. A detailed review of the reliability concepts can be found in Mays and Cullinane (1986).

1.3 Water Distribution System Analysis

The problem of analysis of water distribution systems concerns itself with the evaluation of the flows and pressures in the system, given the layout and sizes of its components. Such an analysis is necessary to evaluate the performance of the system under changing demands. The analysis problem can be stated as the one of solving a set of simultaneous nonlinear equations based on the energy equation and linear equations based on the continuity equation.

1.3.1 Terminology

A water distribution network will be assumed to consist of n nodes numbered 1 through n and a set of l links connecting specified pairs of nodes. The set of nodes will be denoted by N and the set of links by $\mathscr{L}$. A link connecting nodes i and j will be denoted by (i,j). For example, pipe network shown in Fig. 1(a) consists of four nodes N = { 1,2,3,4 } and five links $\mathscr{L}$ = {(1,2), (2,3), (3,4), (4,1), (4,2)}. Removal of link (1,2) results in Fig.1(b). Removal of node 1 results in Fig. 1(c).

The flow $Q_{(i,j)}$ in link (i,j) is considered positive if it is from node i to node j and negative otherwise. The energy head at node i, H_i, is expressed as :

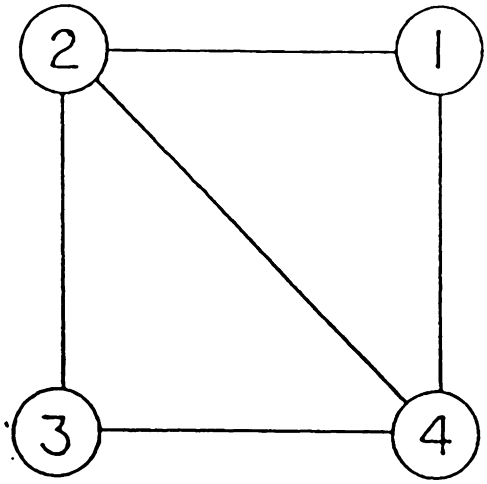
$$H_i = Z_i + h_i + \frac{v_i^2}{2g} \tag{1.1}$$

Where elevation head Z_i is the vertical height of node i above an arbitrary datum, h_i is the pressure head (fluid pressure divided by specific weight of the fluid), $\frac{v_i^2}{2g}$ is the velocity head, and g is the acceleration due to gravity.

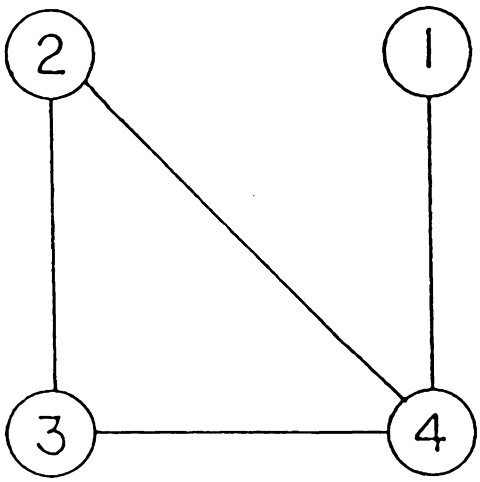
1.3.2 Energy Equation

The algebraic sum of the head losses along a path defined by the set of nodes and links $\{i_1 \; ; \; (i_1, i_2) \; ; \; i_2 \; ; \; (i_2, i_3) \; ; \; i_3 \; ; \; (i_3, i_4) \; ; \; \ldots i_{m-1} \; ; \; (i_{m-1}, i_m) \; ; \; i_m\}$ is equal to the difference between the head at node i_1 and that at node i_m

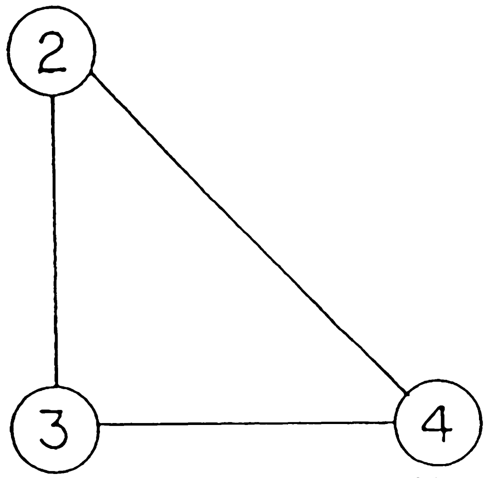
$$H_{i_1} - H_{i_m} = \Delta H = \sum_{k=2}^{m} H_{L(i_{k-1}, i_k)} \tag{1.2}$$



(a) Given Network



(b) Removal of link (1,2)



(c) Removal of node 1

Figure 1. Example Network

In which $H_{L_{(i_{k-1}, i_k)}}$ = head loss in link (i_{k-1}, i_k) which is positive if the flow is from node i_{k-1} to node i_k. For a loop, $i_1 = i_m$, so that $\Delta H = 0$.

The headloss, $H_{L(i,j)}$ between two nodes i and j of a link (i,j) is equal to the sum of frictional headloss $H_{f(i,j)}$ and minor headlosses $H_{m(i,j)}$ minus the head gained due to pumping, $H_{p(i,j)}$. This may be written as

$$H_{L(i,j)} = H_i - H_j = H_{f(i,j)} + H_{m(i,j)} - H_{p(i,j)} \qquad \text{for all} \quad (i,j) \in \mathscr{L} \tag{1.3}$$

These losses (or gains) are expressed as functions of flow $Q_{(i,j)}$, length $L_{(i,j)}$ and diameter $D_{(i,j)}$ of link (i,j) :

$$\begin{aligned} H_{f(i,j)} &= K_{f(i,j)} L_{(i,j)} Q_{(i,j)} | Q_{(i,j)} |^{\alpha - 1} / D^{\beta}_{(i,j)} \\ H_{m(i,j)} &= K_{m(i,j)} | Q_{(i,j)} | Q_{(i,j)} / D^4_{(i,j)} \\ H_{p(i,j)} &= H_0 + A | Q_{(i,j)} | + B | Q_{(i,j)} | Q_{(i,j)} \end{aligned} \tag{1.4}$$

The values of K_f, α and β depend on the type of frictional head loss formula and the system of units used. If the empirical Hazen-Williams formula and SI system of units are used then $\alpha = 1.852$, $\beta = 4.87$ and $K_f = 10.7/C_{hw}^{1.852}$, where C_{hw} is the Hazen-Williams coefficient which depends on the age and material of the pipe. Hazen-Williams formula is widely used in practice because of its simplicity. However, it is valid only for water. Analytical equations like Darcy-Weishbach equation are also available which hold good for any fluid. For Darcy-Weischbach equation $\alpha = 2, \beta = 5$ and $K_f = 8f/\pi^2 g$, where f is the Darcy-Weischbach friction factor which is a function of flow $Q_{(i,j)}$, diameter $D_{(i,j)}$, fluid viscosity and pipe roughness. Because of this dependence on flow the friction factor, f, has to be updated as $Q_{(i,j)}$ changes whereas in the Hazen-Williams formula C_{hw} is treated as a constant. The coefficient K_m depends on the type of bend, valve or fitting. Minor losses are generally negligible except when the valves are present. The equation for pumping head is valid only for centrifugal pumps which are widely used in water distribution systems. The parameters H_0, A and B depend on the pump used (Walski, 1984).

1.3.3 Continuity Equation

The Continuity equation is given as

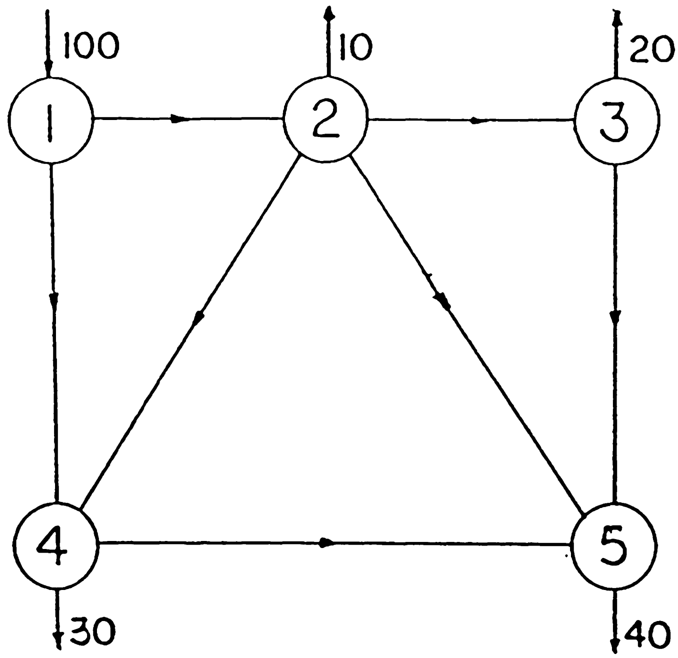
$$\sum_{k \varepsilon k_i} Q_{(k,i)} + q_i = 0 \qquad \text{for all } \; i \in N \tag{1.5}$$

Where $Q_{(k,i)}$ is the flow in the link (k,i), considered to be positive if the flow is from node k to node i ; k_i = set of nodes adjacent to nodes i (a node is said be adjacent to another one if a link exists between them) ; q_i is the external flow at node i, considered positive if the flow is entering node i (supply) and negative otherwise (demand).

1.3.4 Solution of Network Equations

For a balanced network (i.e. supply equals demand) if there are no reservoirs then one of the continuity equations becomes redundant because the continuity equation for the last node can be generated by simply taking the difference of the sum of the incoming and outgoing flows for the remaining nodes (see Fig. 2). From graph theory (see Appendix B) for a network (simple graph) with n nodes and l links, the number of primary loops (independent cycles) denoted as p can be written as (Harary, 1972 ; Berge, 1985)

$$p = l + 1 - n \tag{1.6}$$



Q(1,2) + Q(1,4)	=	100	node 1
Q(2,3) + Q(2,5) + Q(2,4) - Q(1,2)	=	-10	2
Q(3,5) - Q(2,3)	=	-20	3
Q(4,5) - Q(2,4) - Q(1,4)	=	-30	4
Q(2,5) + Q(4,5) + Q(3,5)	=	40	5

Figure 2. Redundancy of One Continuity Equation.

For a system with j junction nodes and f fixed grade nodes, $n = j + f$ and therefore Eqn. (1.6) can be written as

$$l = j + f + p - 1 \tag{1.7}$$

There are l number of Q's to be computed. Based on Equations (1.2) and (1.5) one can write j continuity equations and (f+p) energy equations. Of course, one of these equations will be redundant. Various methods are available for solving these equations, notable among them are the Hardy-Cross method, the Newton-Raphson method and the linear theory method (Wood and Charles, 1972). The Hardy-Cross method selects initial flows based on Eqn. (1.5). Because there are more unknowns than the number of equations of the type Eqn. (1.5) there are infinite number of such initial solutions. Let Q_1^0, $Q_2^0, \ldots$, Q_p^0 be the initial flows in p pipes in path P. The essence of the method is to adjust Q_k 's so that they will also satisfy the energy equations, Eqn. (1.2). It is achieved by perturbing each Q_k^0 by some ΔQ. Let the perturbed flow be given as

$$Q_k = Q_k^0 \pm \Delta Q \tag{1.8}$$

where the ($\pm$) sign depends on the direction of Q_k^0.

For simplicity, taking into account only the friction losses, from Eqs.(1.2) and (1.4) one gets

$$\Delta H = \sum_{k=1}^{p} r_k Q_k \,|\, Q_k \,|^{\alpha - 1} \tag{1.9}$$

in which $r_k = K_f L/D^\beta$ is the resistance coefficient. By using Eqn.(1.8) in Eqn.(1.9) and expanding by Taylor series about Q_k^0 for the first two terms of the series Eqn.(1.9) becomes

$$\Delta H = \sum_{k=1}^{p} r_k Q_k^0 |Q_k^0|^{\alpha-1} + \sum_{k=1}^{p} r_k \alpha |Q_k^0|^{\alpha-1} \Delta Q \quad \text{for path P} \tag{1.10}$$

which can be solved for ΔQ. It should be noted that in Eqn.(1.10) the *same step size*, ΔQ is used for all Q_k^0 for the perturbation, for a fixed path. This ΔQ will in general vary for different paths and new estimates of Q^0 are obtained from Eqn.(1.8).

For the Newton-Raphson method Eqn.(1.10) is written as

$$\Delta H = \sum_{k=1}^{p} r_k Q_k^0 |Q_k^0|^{\alpha-1} + \sum_{k=1}^{p} r_k \alpha |Q_k^0|^{\alpha-1} (Q_k - Q_k^0) \quad \text{for path P} \tag{1.11}$$

From Eqn.(1.11) it is seen that the Newton-Raphson method uses a *variable step size* in terms of $(Q_k - Q_k^0)$ which in general improves the convergence rate. However the Newton's method may not converge for all starting points. It is observed that in Eqn.(1.11) the second sum is simply the directional derivative in the direction $(Q - Q^0)$. If there are p energy and continuity equations, then Eqn.(1.11) can be written separately for the energy equations and together with the continuity equations can be written in the matrix form as

$$[J]_{p \times p} \{Q\}_{p \times 1} = \{F\}_{p \times 1} \tag{1.12}$$

In which J is the Jacobian (matrix of first derivatives for the p continuity and energy equations) and F is the vector of constants from the continuity and energy equations. The solution to Eqn.(1.12) is used as new Q^0 and depends on the nature of [J]. It must be nonsingular.

It is noted that in the Hardy-Cross method Eqn.(1.10) can be solved for the only unknown ΔQ from just one energy equation whereas in the Newton's method Eqn.(1.11) involves p unknown Q_k's and therefore requires all p continuity and energy equations and the matrix inversion in Eqn.(1.12). The implication is that the Hardy-Cross method can be applied by hand whereas the Newton's method is not suitable for hand calculations.

The linear theory method uses a linearization of the form

$$rQ^{\alpha} = rQ_0^{\alpha-1}Q \tag{1.13}$$

in which Q_0 is a chosen point of linearization and r is a known resistance coefficient. Once Q_0 is chosen, equations of the form Eqn.(1.12) can be constructed and solved for the new estimates of Q_0. However it is to be noted that the linearization in Eq.(1.13) lacks theoretical support.

Dillingham and Cleasby (1965) have reported that the Hardy Cross method may lead to divergence in flow computations and no solution can be obtained. Collins (1980) has presented a simple example wherein the linear theory method fails to converge while Newton-Raphson technique converges faster. Moreover Collins, et al. (1979) point out that the solution of the analysis problem will be nonunique if the pumps in the system have nonmonotonic head-discharge characteristic curves. They claim that such pumps are not uncommon. In general it is a good practice to use several different algorithms to verify the solution and with different starting points. The above remarks are meant to emphasize the complexity of the pipe network problems in general. However all these methods have worked very well for most practical problems. For further details on these three methods the reader is referred to Jeppson (1976).

1.4 Design of Water Distribution Systems

1.4.1 Preliminaries

Traditional approach in pipe network design was mainly based on engineering judgement and certain rules of thumb. Several different sets of pipe sizes are guessed and for each guess the performance of the system under various conditions of demand, pipe breakages etc., is simulated by using a Hardy-Cross type network solver. The solution giving the least cost is selected. This type of trial and error approach does not guarantee an optimal solution.

More recent work on design has focussed on the use of mathematical optimization techniques. The structure of a mathematical program can be characterized by an objective function and a set of constraints. Objective function expresses the criterion of design and the constraints express the requirements of the system. A verbal statement of the water distribution design problem can be stated as :

Minimize : Cost of the system

subject to

1. continuity equations
2. energy equations
3. the heads at the nodes should lie within minimum and maximum permissible bounds
4. the system should have "sufficient" reliability

The cost of the system includes pipe, pumping, operation and maintenance costs. These costs are functions of pipe diameters, pumping heads etc. which depend on the flow and other requirements of the system. The best combination of decision variables like diameters, pumping heads etc., which satisfy the constraints while yielding the best objective function value is called the optimal solution.

1.4.2 Tree Networks

For rural distribution systems where the reliability is not a major issue, tree-like topologies are satisfactory. The design of tree networks involves two decisions : the selection of the tree layout from a given set of potential links, and the sizes of the links. The problem of selecting the sizes for a given tree layout is a relatively easier one to solve. A tree with n nodes has (n-1) links. If the tree network has only a single source, and if the demands at the remaining (n-1) nodes are known, then (n-1) independent continuity equations can be solved for the (n-1) link flows. These flows can then be substituted into energy equations to choose the optimal diameters to satisfy the pressure head constraints. For fixed flows and diameters the head loss expression is linear in pipe length. In such cases the constraints and the objective function are linear and the resulting problem can be solved by linear programming techniques. The simultaneous selection of layout and pipe sizes is a difficult problem because of the large number of tree layouts which can be generated from a given set of nodes. It is also observed that in an optimal tree it is quite likely that each node will be connected to its nearest neighbors. By perturbing such trees in a systematic way (Rothfarb, et al., 1970) a near optimal tree can be found without too much computational burden.

1.4.3 Looped Networks

For a large towns and cities, tree networks are not reliable because many nodes can be cutoff from supply even if one link fails. Thus redundancy has to be provided by loop forming links. As pointed out in the previous section the continuity equations are sufficient to solve for the flows in *single source* tree networks. However in looped networks in addition to continuity equations energy equations are needed. The energy equations are nonlinear and the problem in general becomes one of minimizing a nonlinear objective function over a nonconvex constraint set (see Appendix A). At present algorithms are not available which can guarantee a globally optimum solution for this problem. Moreover, if the reliability constraint is not included, then optimization by its very nature, removes redundancy from the network (Templeman, 1982). In fact, it can be shown that the optimization of a looped network results in a tree layout by making the flows zero in the loop forming links (Delfino, 1973). Jacoby (1968) has presented examples with looped layouts. The loop structure is maintained by specifying certain minimum diameter for pipes. One of the example problems is solved by this author by relaxing the minimum diameter condition with the use of GRG2 code (Lasdon and Waren, 1982). As expected, the solution resulted in a tree layout with less cost. The details of the solution are given in Appendix C. The traditional methods of specifying minimum diameter or minimum flows to obtain looped layouts do not provide for a truly optimal design because in such procedures the redundancy is forced in the system rather than brought forth naturally through a reliability constraint.

The past research in this area can be grouped into two categories :

A) approaches which have focussed on optimization of *fixed looped networks*, wherein redundancy is provided by specifying minimum diameters or flows or, by considering various demand patterns.

B) more recent approaches which have addressed the combined optimization-reliability-layout problem in which the layout is *not* fixed but to be determined optimally. The reliability is incorporated explicitly by considering link failures and multiple demand patterns. A brief review of some of the models in each category is given here.

Group A : Methods for given layouts

Gradient search techniques are widely used for the solution of nonlinear optimization problems. In gradient search techniques, typically the solution procedure starts with an initial feasible solution. A move to another feasible solution is made by taking a suitable step in the direction of improvement of the objective function value. The direction is obtained from the gradient of the objective function. The procedure is continued until no improvement in the objective function value is found. Because of the nonconvexity of the looped network optimization problem, if the move is always made in the direction of improvement, the procedure may terminate at a local optimum (Jacoby, 1968). Jacoby therefore used a solution technique in which the step size and the direction at every iteration were adjusted in a random fashion. The pipe diameters were rounded off to the nearest integers and the solution was checked with a Hardy-Cross analyzer.

A Linear Programming Gradient (LPG) search technique was presented by Alperovitz and Shamir (1977). Their approach consists in converting the nonlinear program to a linear one by specifying the initial flows and assuming each link to be further split up into different known diameter pipes of unknown lengths. The solution from the linear program is then utilized in calculating the negative gradient direction (steepest descent) with regard to a perturbation vector of flows and a fixed step size is used for further movement. Quindry, et al. (1979) suggested a correction to the gradient expression used by Alperovitz and Shamir. Fujiwara, et al. (1987) presented a rigorous approach for the LPG technique based on the sensitivity theorem of Fiacco (1983, theorem 3.4.1). They also pointed out the deficiencies of the steepest descent direction method and suggested the use of a quasi-Newton method with variable step size.

Heuristic search techniques offer another avenue for optimization. Ormsbee and Contractor (1981) used the Complex method of Box (1965). For higher dimensional problems the complex method performs poorly. Gessler (1983) used an enumeration technique in which various guesses were examined systematically. The examination was based on a property - if a set S of pipe sizes is not feasible then all sets which have pipe sizes equal to or less than the corresponding sizes in set S cannot be feasible either. Such enumeration searches can be computationally burdensome if many possibilities cannot be discarded apriori. Bhave (1983a, 1983b) used shortest paths from source nodes to demand nodes to arrive at a "design tree". Redundant links, whose positions were assumed to be known, were added by specifying a minimum diameter or minimum flows. Optimal sizes were then obtained using a linear program. The shortest path tree being the optimal layout can be proved if one assumes the hydraulic gradient to be constant for all links and no external energy is supplied (Rowell, 1979). In a general case, it need not be an optimal tree.

Group B : Methods which include layout selection

Rowell and Barnes (1982) used a two-step procedure : the identification of the optimal tree layout from a given set of potential links and, the addition of redundant links to the tree layout. First, a tree layout was selected by solving a nonlinear programming

(NLP) problem. The NLP model was formulated by assuming constant hydraulic gradient for all links. The redundant links were then added to the tree layout by performing a failure analysis of the tree layout. The diameters of the redundant links were optimized using an integer program. The selection of redundant links was based on flow demands only and failed to take into account the hydraulics of the network as a whole. Moreover, hydraulic gradients need not remain constant when the redundant links are added (Goulter and Morgan, 1982).

Morgan and Goulter (1985) devised an LP based heuristic procedure for the least cost design of looped systems. They combined a linear program with a Hardy Cross solver. The sizes were assumed first and the flows were obtained using the network solver. Knowing the flows the sizes were then revised using a linear program. This process was repeated until no improvement was found. At every iteration, uneconomical pipes were removed from the network based on the flows in various links.

1.5 Summary

Based on the literature review, it is clear there is a need for development of new models which can address the following problems simultaneously

1. selection of the layout and sizes of the components

2. incorporation of sufficient reliability to address the link failure problem.

It is difficult to achieve a global optimal solution of the water distribution problem because of its nonconvexity. Also, such an optimization removes the redundancy from the system yielding a tree layout if suitable reliability constraints are not imposed. However, if a near optimal tree layout can be identified by using an efficient search technique then the redundant links can be added to the tree layout to insure service in case of failure of tree links. Such a two step procedure is developed in Chapters 2 and 3. The first step is the *Optimization Model* in which a tree layout is obtained through a heuristic tree search algorithm. Two avenues are followed for obtaining the tree layout : one for single source networks and the other for multiple source networks. The Optimization Model is discussed in Chapter 2. The second step is the *Reliability Model* in which the redundant links are superimposed on the tree network obtained from the Optimization Model in such a way that every demand node has two paths from source nodes. The Reliability Model is discussed in Chapter 3. In Chapter 4 applications to a single source network and a multiple source network are described. In both the cases better solutions are obtained than those obtained by previous researchers.

2.0 Optimization Model

2.1 Introduction

This introduction presents an overview of the chapter. In section 2.2 a general formulation of the water distribution system optimization problem is presented. Because the problem is nonconvex only local optima can be obtained (see Appendix A). A relaxed water distribution system optimization problem without the reliability constraint of the general formulation is presented in section 2.3 and pertinent literature is reviewed. Because optimization does not permit redundancy, the local optima should be tree layouts unless suitable reliability constraints are imposed to obtain looped layouts. Since a tree layout would qualify to be an optimal solution to the relaxed problem, an efficient tree search algorithm can be adopted without much computational burden. Section 2.4 presents such a tree search algorithm which systematically perturbs tree layouts to obtain an optimal tree layout without exhaustive enumeration. The algorithm also employs a linear program to optimize the sizes of the system components for each newly generated layout. After sufficient exploration, current best tree layout with its optimal system components is declared as near optimal.

2.2 Problem Statement

Following parameters are assumed to be known before the design is carried out :

1. A set of nodes, N, steady state flow demands and minimum/maximum pressure requirements at each node. Variations in demand can be considered by allowing for various demand patterns (peak, average etc.). In this study only single demand pattern is considered. The distances between the nodes are also known. Circular pipes are to be used to join the nodes.

2. Location of reservoirs and potential locations of pumping stations are known. The cost of elevating a reservoir and the cost of pumping for unit head at a specified discharge is also known.

3. The reliability criterion should be specified. In the present study, the criterion used is that every demand node should have at least two paths from source node(s).

The problem is to select a near optimal layout based on the above information i.e., to select near optimal links from a set of candidate links. For such a layout the decision variables of the problem are the diameters of the pipes, the flow in each link, head at each node, the reservoir heads and the pump heads. The general optimization problem is formally stated as follows :

Problem P1

$$\text{Minimize} \quad C = \sum_{(i,j)\in\mathscr{L}} F_L[D_{(i,j)}\,,\, L_{(i,j)}] + \sum_{(i,j)\in\mathscr{L}} F_P[H_{P(i,j)}\,,\, Q_{P(i,j)}] + \sum_{i\in S} F_s[H_{Si}] \tag{2.1}$$

subject to

$$\sum_{k\in k_i} Q_{(k,i)}(t) + q_i(t) = 0 \qquad \forall i \in N, \ \ \forall t \tag{2.2}$$

$$H_i - H_j = H_{f(i,j)}(t) + H_{m(i,j)}(t) - H_{p(i,j)}(t) \qquad \forall (i,j) \in \mathscr{L}, \ \ \forall t \tag{2.3a}$$

$$H_{f(i,j)}(t) = K_{f(i,j)} L_{(i,j)} Q_{(i,j)}(t)\,|\,Q_{(i,j)}(t)\,|^{\alpha-1}/D^{\beta}_{(i,j)} \qquad \forall (i,j) \in \mathscr{L} \tag{2.3b}$$

$$H_{m(i,j)}(t) = K_{m(i,j)}\,|\,Q_{(i,j)}(t)\,|\,Q_{(i,j)}(t)/D^{4}_{(i,j)} \qquad \forall (i,j) \in \mathscr{L} \tag{2.3c}$$

$$H_{p(i,j)}(t) = H_0 + A\,|\,Q_{(i,j)}(t)\,| + B\,|\,Q_{(i,j)}\,|\,Q_{(i,j)}(t) \qquad \forall (i,j) \in \mathscr{L} \tag{2.3d}$$

$$H_i = F_i + H_{Si} \qquad \forall i \in S \tag{2.4}$$

$$\underline{H}_i \le H_i \le \overline{H}_i \qquad \forall i \in N \tag{2.5}$$

$$R_s \ge R_{min} \tag{2.6}$$

$$D_{(i,j)} \in D_I \qquad \forall (i,j) \in \mathscr{L} \tag{2.7}$$

$$H_{p(i,j)}(t) \ge 0 \qquad \forall (i,j) \in P, \ \ \forall t \tag{2.8}$$

$$H_{Si} \ge 0 \qquad \forall i \in S \tag{2.9}$$

Where :

N = set of all nodes = { 1, 2, 3,...n }
n = number of nodes
$\mathscr{L}$ = set of all potential links
S = set of reservoir nodes
P = set of links on which pumps can be located
k_i = set of nodes which are adjacent to node i
t = index indicating demand pattern = 1, 2, 3,...,N_t
N_t = number of demand patterns
D_I = set of available diameters = $\{d_1,\ d_2,\ d_3,\ldots,\ d_{N_D}\}$
N_D = number of available diameters
$D_{(i,j)}$ = diameter of link (i,j)

$L_{(i,j)}$ = length of link (i,j)
$Q_{(i,j)}(t)$ = flow in link (i,j) for pattern t with the notation $Q_{(i,j)}(t) \geq 0$ if flow is from node i to node j , $Q_{(i,j)}(t) < 0$ otherwise
$H_{f(i,j)}(t)$ = friction head loss in link (i,j) for pattern t
$H_{m(i,j)}(t)$ = minor head loss in link (i,j) for pattern t
$H_{p(i,j)}(t)$ = pump head in link (i,j) for pattern t
H_i = energy head at node i
H_{Si} = additional elevated reservoir head at node i
F_i = fixed head at (reservoir) node i
$q_i(t)$ = external flow at node i for pattern t, ≥ 0 if it is into the node (supply), < 0 otherwise (demand)
F_L = cost of link (i,j), a function of its diameter and length
F_P = cost of pumping a head $H_{p(i,j)}$ in a link (i,j), a function of the head and $Q_{(i,j)}$
F_S = cost of raising a reservoir at node i by a height of H_{Si} , a function of H_{Si}
$\underline{H}_i$ = minimum head required at node i
$\overline{H}_i$ = maximum head permitted at node i
R_s = a measure of reliability of the system; a function of the decision variables

The parameters K_f, α, and, β for the friction headloss depend on the formula and the system of units used. K_m, the minor loss coefficient, depends on the type of bend, valve etc. It is generally neglected. The parameters H_0, A and B for the pump head depend on the type of pump used.

2.2.1 Objective Function

Only the link, pump and reservoir elevation costs have been considered. If the diameters of the links are taken as decision variables then the link cost can be expressed as (Rowell and Barnes, 1982) :

$$F_L = C_1 L_{(i,j)} D_{(i,j)}^{C_2} \tag{2.10}$$

where C_1 and C_2 are cost parameters. In practice, however, the cost of the pipes are specified per unit length for a particular diameter. Thus if a link is assumed to consist of different diameters, then the lengths become decision variables and the objective function becomes linear. This will be dealt with later. The pump cost function can be expressed as (Alperovitz and Shamir, 1977) :

$$F_{P(i,j)} = C_3 \gamma | Q_{(i,j)} | H_{P(i,j)} / \eta \tag{2.11}$$

where C_3 is the unit power cost, γ is a coefficient which depends on the pump used, and η is the efficiency. γ is computed to reflect the total length of time that the specific loading condition is assumed to prevail. The cost of elevating reservoirs can be expressed as (Rowell and Barnes, 1982) :

$$F_{Si} = C_S H_{Si} \tag{2.12}$$

2.2.2 Constraints

Constraints (2.2) are linear continuity equations for all loading patterns indexed by t. Constraints (2.3a) are the energy equations for all links. The friction and minor losses and pump head are functions of flow and diameter in the links (Eqs. 2.3b to 2.3d). Constraint (2.6) is the reliability constraint. By incorporating constraint (2.7) the diameters are restricted to commercially available sizes. It is to be noted that by allowing for zero diameter the model will have the capability of eliminating uneconomical links. However, if Eqn. (2.10) is used for the link cost function, it is difficult to model constraint (2.7) unless (0,1) integer variables are used (Artina, 1973) in which case the solution becomes computationally burdensome. This problem can be overcome by using the lengths of pipes of particular diameters as decision variables.

Two main difficulties in solving the problem P1 are : (1) The difficulty of modeling the reliability constraint - if one specifies minimum diameter or minimum flows, the model does not have the capability of eliminating any link. Thus, the system becomes over redundant. (2) Even if the reliability constraint is not included, obtaining a solution to the mathematical program is a formidable task. The objective function and the energy equations are nonlinear. In general, the presence of nonlinear equality constraints makes the feasible space nonconvex. Since the energy constraints are equality constraints, the problem becomes one of minimizing a nonlinear objective function over a nonconvex constraint set. At present, no satisfactory algorithms are available which can guarantee a global optimum solution to this problem.

2.3 Relaxed Problem

In this section a methodology is presented for solving the optimization problem (P1) without the reliability constraint for *steady state single demand pattern.* It is known that the solution to this problem results in a tree layout by making the flows zero in the nontree links (Delfino, 1973). Thus the problem consists in finding the tree layout and the pipe sizes of that layout. However finding an optimal solution to Problem P1 even after omitting the reliability constraint is difficult as discussed below:

Problem P2

$$\text{Minimize} \sum_{(i,j)\in\mathscr{L}} \sum_{k=1}^{N_D} C_{Lk}X_{(i,j)k} + \sum_{(i,j)\in\mathscr{L}} C_3\gamma\,|Q_{(i,j)}|\,H_P(i,j)/\eta + \sum_{i\in S} C_s H_{Si} \quad (2.13)$$

subject to

$$\sum_{k\in k_i} Q_{(k,i)} + q_i = 0 \qquad \forall i \in N \quad (2.14)$$

$$H_i - H_j = K_{f(i,j)}Q_{(i,j)}|Q_{(i,j)}|^{\alpha-1}\sum_{k=1}^{N_D}\frac{X_{(i,j)k}}{d_k^{\beta}} + K_{m(i,j)}|Q_{(i,j)}|Q_{(i,j)}\sum_{k=1}^{N_D}\frac{1}{d_k^{\beta}}$$
$$- H_0 - A|Q_{(i,j)}| - B|Q_{(i,j)}|Q_{(i,j)} \qquad \forall (i,j)\in\mathscr{L} \quad (2.15)$$

$$H_i = F_i + H_{Si} \qquad \forall i \in S \quad (2.16)$$

$$\underline{H}_i \le H_i \le \overline{H}_i \qquad \forall i \in N \quad (2.17)$$

$$\sum_{k=1}^{N_D} X_{(i,j)k} = L_{(i,j)} \qquad \forall (i,j)\in\mathscr{L},\ \forall k \quad (2.18)$$

$$H_{p(i,j)} \ge 0 \qquad \forall (i,j) \in P \quad (2.19)$$

$$H_{Si} \ge 0 \qquad \forall i \in S \quad (2.20)$$

$$X_{(i,j)k} \ge 0 \qquad \forall (i,j)\in\mathscr{L},\ \forall k \quad (2.21)$$

Where :

N = set of all nodes = { 1, 2, 3,...n }
n = number of nodes
$\mathscr{L}$ = set of all potential links
S = set of reservoir nodes
P = set of links on which pumps can be located
k_i = set of nodes which are adjacent to node i
d_k = known diameter from the list $\{d_1, d_2, \ldots, d_{N_D}\}$
N_D = number of available diameters
$L_{(i,j)}$ = length of link (i,j)
$Q_{(i,j)}$ = flow in link (i,j) with the notation $Q_{(i,j)} \ge 0$ if flow is from node i to node j , $Q_{(i,j)} < 0$ otherwise
$H_{p(i,j)}$ = pump head in link (i,j)
H_i = energy head at node i
H_{Si} = additional elevated reservoir head at node i

F_i = fixed head at (reservoir) node i
q_i = external flow at node i ≥ 0
if it is into the node (supply), < 0 otherwise (demand)
$\underline{H}_i$ = minimum head required at node i
$\overline{H}_i$ = maximum head permitted at node i
$X_{(i,j)k}$ = length of k^{th} diameter pipe in link (i,j)
C_{Lk} = cost per unit length of k^{th} diameter pipe

The constraints of Problem P2 are the same as those of Problem P1 except the constraint (2.18). To restrict the pipe diameters to commercially available ones each link is assumed to be split into as many subsections as the number of different but known diameter pipes of unknown lengths. For example, if N_D different diameters are available, then link (i,j) is assumed to consist of N_D different pipes such that

$$\sum_{k=1}^{N_D} X_{(i,j)k} = L_{(i,j)} \tag{2.22}$$

where $X_{(i,j)k}$ is the unknown length of k^{th} diameter pipe in link (i,j). It is noted that the diameters d_k's associated with the respective pipe lengths $X_{(i,j)k}$'s are known. It is also noted that for given $Q_{(i,j)}$'s Problem P2 becomes a linear program. Such a strategy is adopted in the tree search algorithm described in section (2.4).

However for unknown $Q_{(i,j)}$'s it is difficult to obtain an optimal solution to Problem P2 because of the nonconvexity. The problem is further compounded by large number of decision variables involved. For example, for a system with 20 nodes and 30 links, 1 elevated reservoir, 1 pump and 10 diameters, there are 352 decision variables (30x10 lengths, 30 flows, 1 pump head, 1 reservoir head and 20 heads). Gradient search techniques like Generalized Reduced Gradient (GRG) method (Abadie, 1970) require specification of initial feasible solution. Since large number of variables are involved, it is difficult to guess such a solution. Linear Programming Gradient (LPG) technique (Alperovitz and Shamir, 1977 ; Quindry, et al., 1981 ; Fujiwara, et al., 1987) requires specification of flows only. However, obtaining a feasible flow distribution itself is a difficult task. This procedure also yields a local optimum. Rowell and Barnes (1982) reduced the nonconvexity of Problem P2 by assuming that hydraulic gradients remain constant in each link which converts the problem to a concave minimization problem for which only a local optimum solution, which is either a single tree or a set of disconnected trees, can be found. The set of disconnected trees called a forest may result only if the system has multiple sources. In such cases if one connects the various trees by links for better reliability, a single tree can be obtained. Therefore taking some liberty one might say that the optimal solution to Problem P2 is a tree. Based on the above remarks it is seen that there is a need for an alternative approach for obtaining the tree layout.

A naive way of obtaining an optimal tree layout is to search all possible trees. However, this can be computationally burdensome because the number of trees that can be generated from a set of potential links can be very large. For example, a 20 node, 28 link network of Fig. 3 has 135,320 possible trees (Rowell, 1979 ; see also Christofides (1975) pp.125-133, for a method to enumerate all the possible trees). Rothfarb, et al. (1970) have suggested a heuristic to obtain a near optimal tree layout without exhaustive enumeration. The tree search technique has two main advantages over the gradient based optimization methods :

1. For the tree search, shortest path tree can serve as a very good initial tree. In fact, the shortest path tree can be shown to be the optimal tree if no external energy is supplied and hydraulic gradients are assumed to be constant for each link (Rowell,

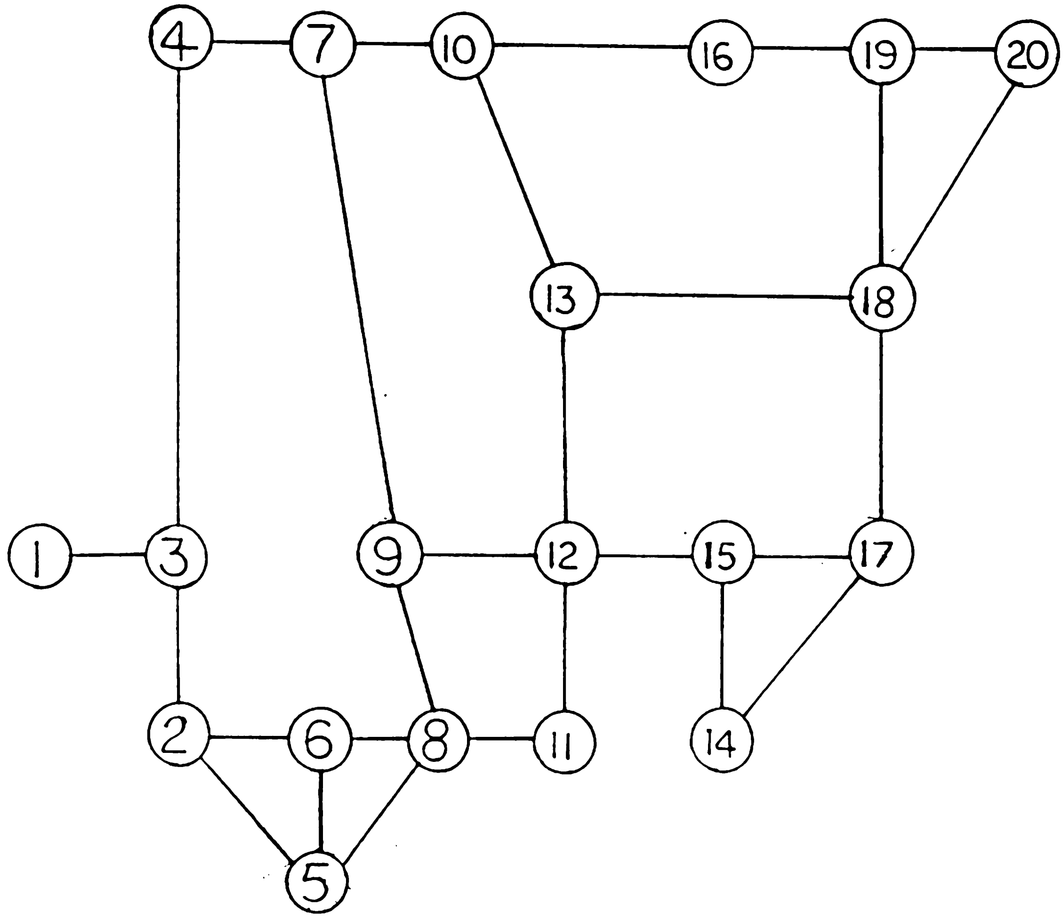


Figure 3. Example Network from Rowell (1979)

1979 ; see also Bhave, 1979). Efficient algorithms are available to obtain a shortest path tree (Dijkstra, 1959; Ravindran, et al., 1987).

2. The tree search technique does not require computation of gradients or step sizes. Therefore this approach can be applied to situations where analytical expressions for the gradients are not possible.

Also, it should be noted that for water distribution networks many potential links are eliminated from consideration because of restrictions of public right of way, topography and presence of other structures which reduce the number of trees to be examined.

2.4 Algorithm TREESEARCH

Addition of a link to a tree creates exactly one loop (Fig. 4). Different trees can be generated by deleting links from the loop one by one. Addition of a link at a node and deletion of a link from the loop to form a new tree is called an elementary transformation. Algorithm starts by performing an elementary transformation at a node. If after performing an elementary transformation the new tree is found to have a lower cost, then it is preserved and an elementary transformation at a new node is performed. If the new tree has higher cost then a different tree is generated by deleting another link from the same loop. Algorithm ends when the transformations have been performed at every node.

Let T denote the set of links which form a tree. Let $\mathscr{L}$ denote the set of potential links. Let CT denote the set of links (cotree) which are not in T (i.e., CT = $\mathscr{L}$ - T). Let $k_{i,CT}$ denote a set of nodes such that

$$k_{i,CT} = \{j \mid (i,j) \in CT\}$$

The algorithm can be summarized as follows:

Algorithm TREESEARCH

Purpose: To find a near optimal tree layout.

Input: initial tree, T; set of nodes, N; Set of sources, S;

set of links with pumps, P; set of diameters, $\{d_1, \dots, d_{N_D}\}$;

set of potential links, $\mathscr{L}$; link lengths, $\{L_{(i,j)}, (i,j) \in \mathscr{L}\}$;

demands, maximum and minimum heads, $\{q_i, \underline{H}_i, \overline{H}_i, i \in N\}$;

fixed heads, $\{F_i, i \in S\}$; unit link costs for each diameter, $\{C_{Lk}, k = 1, \dots, N_D\}$;

pumping cost, C_P; reservoir elevation cost, C_S

Output: Optimal tree layout, T^*; link flows, $\{Q_{(i,j)}, \forall (i,j) \in T^*\}$;

pumping heads, $\{H_{p(i,j)}, \forall (i,j) \in P\}$; pipe lengths, $X_{(i,j)k}$, of diameter d_k;

heads, $\{H_i, \forall i \in N\}$; additional reservoir heads, $\{H_{Si}, i \in S\}$

Method:

1. Initialization

   Set $i = 0$

   Find cost C_T of the tree T by optimization (see section 2.4.1)

   $T^* = T$, best $= C_T$

2. Visit a new node; if all the nodes have been visited then declare the best tree found as the optimal tree

   $i = i + 1$

   if $i > n$ then

declare T^* as the optimal tree; STOP

3. If possible, add a link from the cotree to node i to form a loop. If not, visit a new node.

   If $k_{i,CT} = \Phi$ then {cannot add a link at node i}

   go to step 2

   Else

   a. Find a link (i,j) such that $j \in k_{i,CT}$

   b. $T = T \cup \{(i,j)\}$

   c. $k_{i,CT} = k_{i,CT} - \{j\}$

4. Detect the loop formed by the addition of (i,j) by using procedure LOOP. Let L_P denote the set of links on the loop except link (i,j)

5. If possible, delete a link from L_P to form a new tree. If not, then visit a new node after deleting the added link (i,j)

   If $L_P \neq \Phi$ then

   a. Find a link $(r,s) \in L_P$

   b. $L_P = L_P - \{(r,s)\}$

   c. $T = T - \{(r,s)\}$

   Else

   a. $T = T - \{(i,j)\}$

   b. go to step 3

6. Optimize the current tree, T; find the cost C_T. If the cost is less than the best cost so far, then save this tree and visit a new node; else try deleting another link from the loop

Optimize T to find C_T

If $C_T \leq$ best then

a. best = C_T

b. $T^* = T$

c. go to step 2

Else

a. T = T + {(r,s)} {replace the deleted link}

b. go to step 5

End of TREESEARCH

Example (Fig. 4)

Step 1 (Fig. 4a)

$$
\begin{aligned}
T &= \{(1,2), (2,3), (2,5), (1,4)\} \\
CT &= \{(1,3), (3,4), (3,5), (4,5)\} \\
i &= 0 \\
C_T &= 500 \text{ (say)} \\
T^* &= T \\
\text{best} &= 500
\end{aligned}
$$

Step 2

$i = 1 < 5$

Step 3

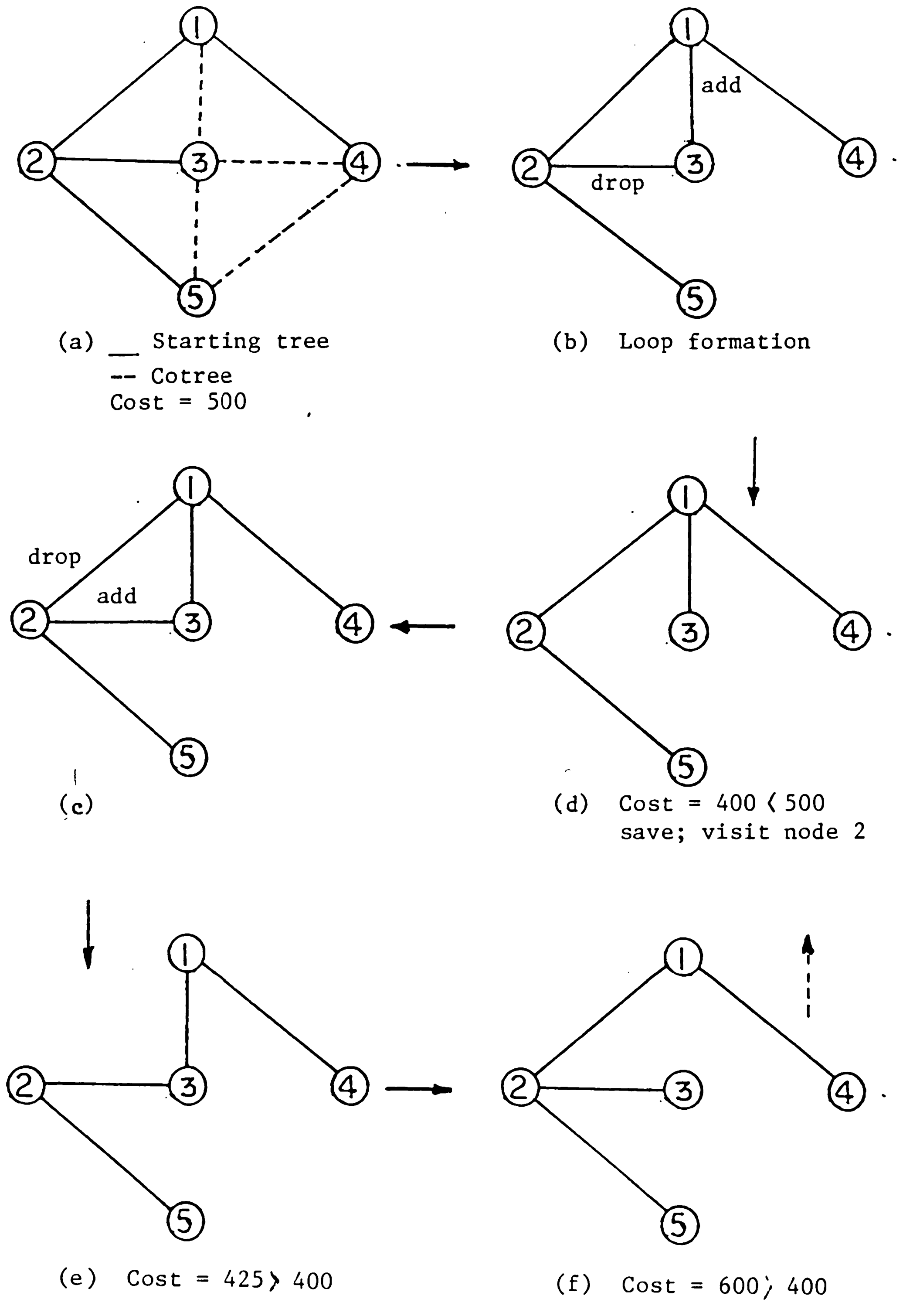


Figure 4. Illustration of Algorithm TREESEARCH

$k_{1,CT} = \{3\} \neq \Phi$

$(1,3) \in CT$; $3 \in k_{1,CT}$

$T = \{(1,2), (2,3), (2,5), (1,4), (1,3)\}$

$k_{1,CT} = \Phi$

Step 4

$L_P = \{(1,2), (2,3)\}$

Step 5

$(2,3) \in L_P$

$L_P = \{(1,2)\}$

$T = \{(2,3), (2,5), (1,4), (1,3)\}$ {Fig. 4c}

Step 6

$C_T = 400$ (say) $<$ best

best $= 400$

$T^* = T$ {Fig. 4c}

Step 2

$i = 2$

Step 3

$k_{2,CT} = \{3\} \neq \Phi$

$(i,j) = (2,3)$

$T = \{(1,3), (1,4), (2,5), (1,2), (2,3)\}$

$k_{2,CT} = \Phi$

Step 4

L_P $\{(1,3), (1,2)\}$ {Fig. 4d}

Step 5

$(r,s) = (1,2)$
$L_P = \{(1,3)\}$
$T = \{(1,3), (1,4), (2,3), (2,5)\}$ {Fig. 4e}

Step 6

$C_T = 425 > best$
$T = \{(1,3), (1,4), (2,5), (1,2), (2,3)\}$

Step 5

$L_P = \{(1,3)\}$
$(r,s) = (1,3)$
$L_P = \varphi$
$T = \{(1,4), (1,2), (2,5), (2,3)\}$ {Fig 4f}

Step 6

$C_T = 600 > best$
$T = \{(1,3), (1,4), (1,2), (2,5), (2,3)\}$

Step 5

$L_P = \Phi$
$T = \{(1,3), (1,4), (1,2), (2,5)\}$ {Fig. 4c}

Algorithm proceeds till node 5 is visited.

Loop Detection

Let (p,q) be a link from the cotree CT which is added to the tree T. This creates exactly one loop. Clearly, link (p,q) is one of the links on the loop. The procedure starts at node

q. Nodes which are connected to node q are visited. If a pendant node (a node connected to only one link) is reached then we move back to the previously visited node and visit other nodes connected to it. The visited nodes are stored sequentially in a set C_Y. Those nodes which become pendant are removed from C_Y. The procedure ends when node p is visited and at that time C_Y contains only those nodes which are on the loop.

Let adj(i) denote a set of nodes which are adjacent to node i in T. adj(i) can be found easily from adjacency matrix A of T. The element a_{ij} of A = 1 if the nodes i and j are adjacent; else a_{ij} = 0. Thus

$$\text{adj}(i) = \{j \mid a_{ij} = 1\}$$

The cardinality of set adj(i) is called the degree of node i. Let deg(i) denote the degree of node i.

$$\deg(i) = |\text{adj}(i)|$$

Procedure LOOP

Purpose: To detect the set of links L_P in the loop formed by addition of a link (p,q) to a tree T.

Input: (p,q), T

Output: Set of links on the loop, L_P

Method:

1. Initialization: C_Y stores the set of nodes on the loop. prev(i) is a node which has been visited previous to node i.

   $C_Y = \{p,q\}$

$j = q$

$prev(i) = p$

$T = T \cup \{(p,q)\}$

2. Select a node k adjacent to the current node, other than the previous node. If k is not a pendant node, then store it in C_Y ; else remove it from C_Y and delete the link connected to it. Loop is found if node p is visited.

   Until $j \neq p$ do:

   (i) Find $k \in adj(j)$ but $k \neq prev(j)$

   (ii) $prev(k) = j$

   (iii) If $deg(k) \neq 1$ then

   $C_Y = C_Y \cup \{k\}$

   Else {remove all pendant nodes and links}

   Until $deg(i) = 1$ do:

   a. $T = T - \{(prev(k), k)\}$

   b. $C_Y = C_Y - \{k\}$

   c. $k = prev(k)$

   (iv) $j = k$

3. C_Y contains the set of nodes on the loop. Convert it to a set of links.

   $L_P = \{(y_1,y_2), (y_2, y_3),...., (y_{r-1}, y_r)\}$

   where $y_1, y_2,..., y_r$ are the members of C_Y in that order. Note that $y_1 = p, y_2 = q,$ and $y_r = p$.

End of LOOP

Example (Fig. 5)

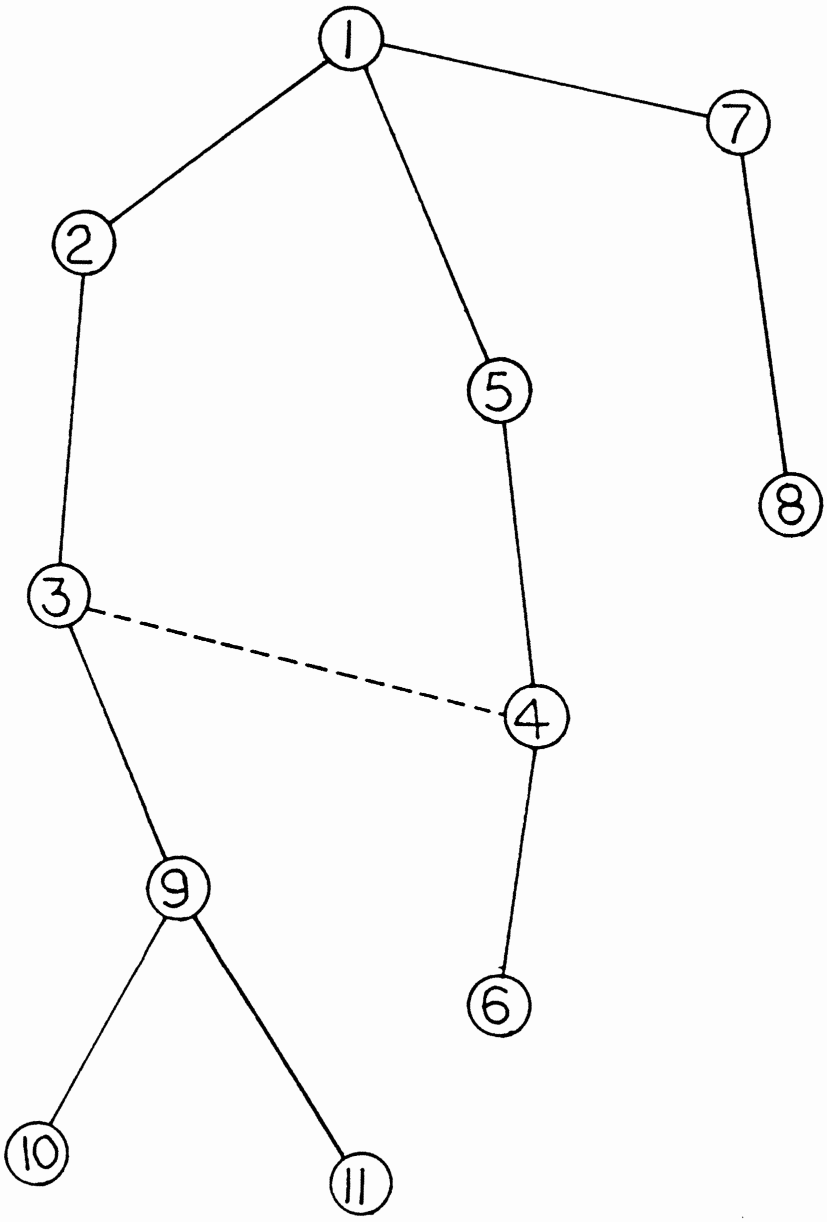


Figure 5. Loop Detection

Step 1

$C_Y = \{3,4\}$

$j = 4$

$prev(4) = 3$

$T = T \cup \{(3,4)\}$

Step 2

(i) $adj(4) = \{3,5,6\}$

$k = 6$ (say)

(ii) $prev(6) = 4$

(iii) $deg(6) = 1$

a. $T = T - \{(4,6)\}$ [(4,6) deleted]

b. $C_Y = \{3,4\}$

c. $k = 4$

(iv) $j = 4$

(i) $adj(4) = \{3,5\}$

$k = 5$

(ii) $prev(5) = 4$

(iii) $deg(5) = 2$

$C_Y = \{3,4,5\}$

(iv) $j = 5$

(i) $adj(5) = \{1,4\}$

$k = 1$

(ii) $prev(1) = 5$

(iii) $deg(1) = 3$

$C_Y = \{3,4,5,1\}$

(iv) j = 1

(i) adj(1) = {2,5,7}

k = 7 (say)

(ii) prev(7) = 1

(iii) deg(7) = 2

$C_Y = \{3,4,5,1,7\}$

(iv) j = 7

(i) adj(7) = {1,8}

k = 8

(ii) prev(8) = 7

(iii) deg(8) = 1

a. T = T - {(7,8)} [(7,8) deleted]

b. $C_Y = \{3,4,5,1,7\}$

c. k = 7

deg(7) = 1

a. T = T - {(1,7)} [(1,7) deleted]

b. $C_Y = \{3,4,5,1\}$

c. k = 1

deg(1) = 2

(iv) j = 1

(i) adj(1) = {2,5}

k = 2

(ii) prev(2) = 1

(iii) deg(2) = 2

$C_Y = \{3,4,5,1,2\}$

(iv) j = 2

(i) adj(2) = {1,3}

k = 3

(ii) prev(3) = 2

(iii) deg(3) = 2

$C_Y = \{3,4,5,1,2,3\}$

(iv) j = 3 = p

Step 3

$$L_P = \{(3,4), (4,5), (5,1), (1,2), (2,3)\}$$

2.4.1 Optimization of Tree Networks

Tree Network Optimization is an integral part of Algorithm TREESEARCH. It is noted that Problem P2 becomes a linear program if the link flows are known. The methods for optimization either iterate between a flow finding step and the linear programming step or perform these steps independently.

2.4.1.1 Single Source Problem

(1) solve the continuity equations (this can be done exactly for single source tree networks)

(2) use the linear program

2.4.1.2 Multiple Source Problem

For the multiple source networks the continuity equations cannot be solved exactly if the supply rates at the sources are unknown. Hence Problem P2 remains nonlinear. Alternative strategies are discussed to overcome this problem :

(a) Approach based on that of Goulter and Morgan (1985) :

(1) assume the diameters of each link in the tree, use a network analyzer to solve for the flows in each link

(2) using the flows from step 1 optimize the diameters by a linear program. If there is no improvement in the cost over the previous cost, stop; else go to step 1.

This procedure, however, does not have any theoretical support.

(b) The Linear Programming Gradient (LPG) method (Fujiwara, et. al., 1987) :

(1) assume a feasible flow distribution

(2) use the linear program to optimize the cost

(3) use the dual solution of the linear program to construct the gradient with respect to the perturbed flows. Take a step in the direction of the negative gradient to obtain a new feasible flow distribution. If the new flow distribution is close to the previous one or, if the gradient is zero, stop; else go to step 2.

Fujiwara, et al. provide a mathematical basis for such a procedure and it is shown that the procedure should terminate at a local optimum.

(c) Approach based on Nonlinear Minimum Cost Flow Model (NMCF) (Rowell, 1979) :

(1) use NMCF model to identify the supplies at the sources. Once the supplies are known, solve the continuity equations to obtain the link flows.

(2) use the linear program to optimize the cost.

It is shown that the Problem P2 for a potential looped network can be simplified to the NMCF model if the hydraulic gradients are assumed to be constant in each link. A statement of the NMCF model is the following:

$$\text{Minimize} \sum_{(i,j)\in\mathcal{L}} \overline{K}_{f(i,j)} L_{(i,j)} | Q_{(i,j)} |^{C} \tag{2.23}$$

subject to

$$\sum_{\forall k \in k_i} Q_{(k,i)} + q_i = 0 \tag{2.24}$$

where $\overline{K}_{f(i,j)} = C_1 [K_{f(i,j)}/J]^{C_2/\beta}$, $C = \alpha C_2/\beta$, C_1 and C_2 are the link cost parameters (Eqn. 2.10), β is the exponent in the friction head loss equation (Eqn. 1.4) and J is the hydraulic gradient in each link. Minor losses are neglected. The objective function is concave and the constraints are linear. The solution to this problem not only gives the unknown supply rates at the sources but also a near optimal tree layout.

(d) Linear Minimum Cost Flow (LMCF) model (Rowell, 1979) :

(1) use LMCF model to identify the supplies at the sources and solve the continuity equations to obtain the link flows.

(2) use the linear program to optimize the cost.

LMCF is used to obtain a shortest path tree for multiple source networks. This model assigns each demand node to a particular source, as a result of which, the supply at each source is determined. The assignment is based on source capacities, nodal demands, and the distances between each source and demand node. A statement of the model is presented below :

$$\text{Minimize} \sum_{(i,j)\in\mathscr{L}} L_{(i,j)} \left| Q_{(i,j)} \right| \tag{2.25}$$

subject to

$$\sum_{\forall k \in k_i} Q_{(k,i)} + q_i = 0 \tag{2.26}$$

This is a linear program in which the decision variables are the flows in each link (i,j) of the full potential network and the supply rates (q_i, $\forall i \in S$) at the sources. This model in essence, is the linear approximation to the NMCF model where the exponent C in the objective function is taken to be unity. Since the shortest path tree is a near optimal tree (in terms of layout) , the use of this approximate model seems appropriate to allocate the various sources to demand nodes. Moreover the problem can be efficiently solved even for a large sized problems. This procedure is used for an example problem in Algorithm TREESEARCH. The application is described in Chapter 4.

3.0 Reliability Model

3.1 Introduction

The tree network obtained from the Optimization Model in Chapter 2 is vulnerable because even if one link fails, many demand nodes can be cutoff from supply. Thus it is necessary to provide loop forming redundant links in such a way that the demand nodes remain connected to the sources in case of a tree link failure. Loop forming links also provide better circulation and help avoid accumulation of sediments. The reliability model employed in this study involves two steps : (1) selection of redundant links and (2) selection of diameters of redundant links. When a link in a tree fails, two connected components are created. Therefore any link from the cotree (complement of the given tree from the potential link set) having adjacent nodes in the respective connected components should qualify to be a candidate redundant link. Of course different such reconnecting cotree links denoted as $RC_{(i,j)}$ will qualify for each link (i,j) of the given tree. Therefore a judicious choice would involve the selection of subsets of these $RC_{(i,j)}$'s with a minimum number of reconnecting links i.e., to pick the the most common of all reconnecting links such that two different paths are provided between sources and demand nodes for each link failure in the tree. Algorithm REDUNDANCY presented in section 3.3 generates such reconnecting links from the cotree. Section 3.4 provides some general guidelines for selecting diameters for the redundant links.

3.2 Selection of Redundant Links

Optimal tree layout is selected by the Algorithm TREESEARCH discussed in section 2.4. Let T denote the set of links in the optimal tree layout. All the links which are not in T constitute the cotree, denoted by the set CT = $\{(i,j) \in \mathscr{L} - T\}$ in which $\mathscr{L}$ is the set of all links. The problem of finding the redundant links which satisfy the reliability criterion can be formally stated as follows: Given a network with a set of nodes N, a set of source nodes S, a set of candidate links $\mathscr{L}$ and the optimal tree T, find a subset RL of redundant links from the cotree CT in such a way that RL contains minimum number of links and the new network formed by the set of links ($T \cup RL$) and the associated nodes has the following property :

> All (or most) nodes of the new network which are disconnected from the sources because of a failure of any link in T are reconnected to the source(s) because of addition of the set RL from the cotree CT.

The solution procedure for this problem is explained with an example. Consider a network shown in Fig. 6. The network has N = {1, 2, 3,...,8 } ; $\mathscr{L}$ = {(1,2), (1,5), (1,6), (2,3), (2,5), (3,4), (3,5), (4,8), (5,6), (5,7), (5,8), (6,7), (7,8)} ; T = {(1,2), (1,5), (1,6), (2,3), (3,4), (6,7), (7,8)} as shown by the dark lines and CT = {(2,5), (3,5), (4,8), (5,6), (5,7), (5,8)} as shown by the dotted lines. Node 2 is the source. Fig. 7a shows the failure of link (1,6) of T. Clearly this disconnects the set of nodes N_1 = {6, 7, 8} from

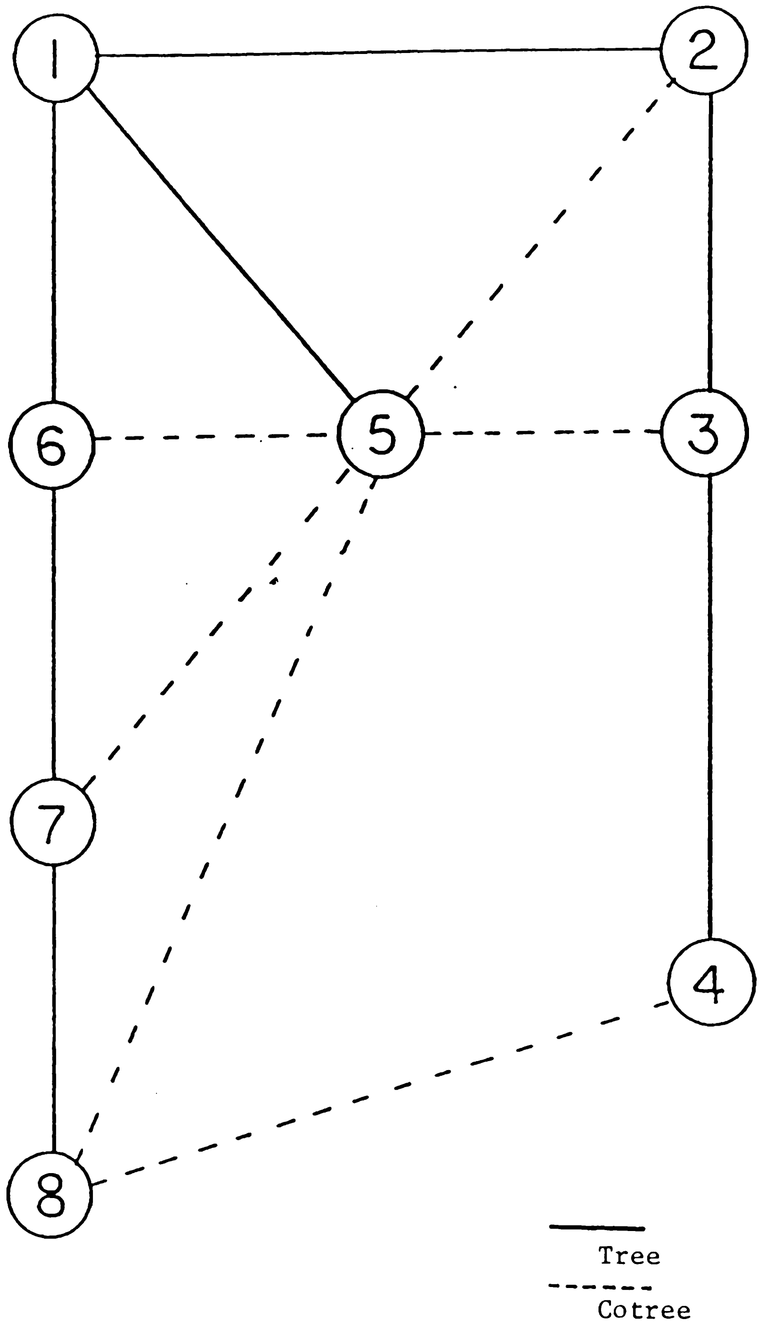
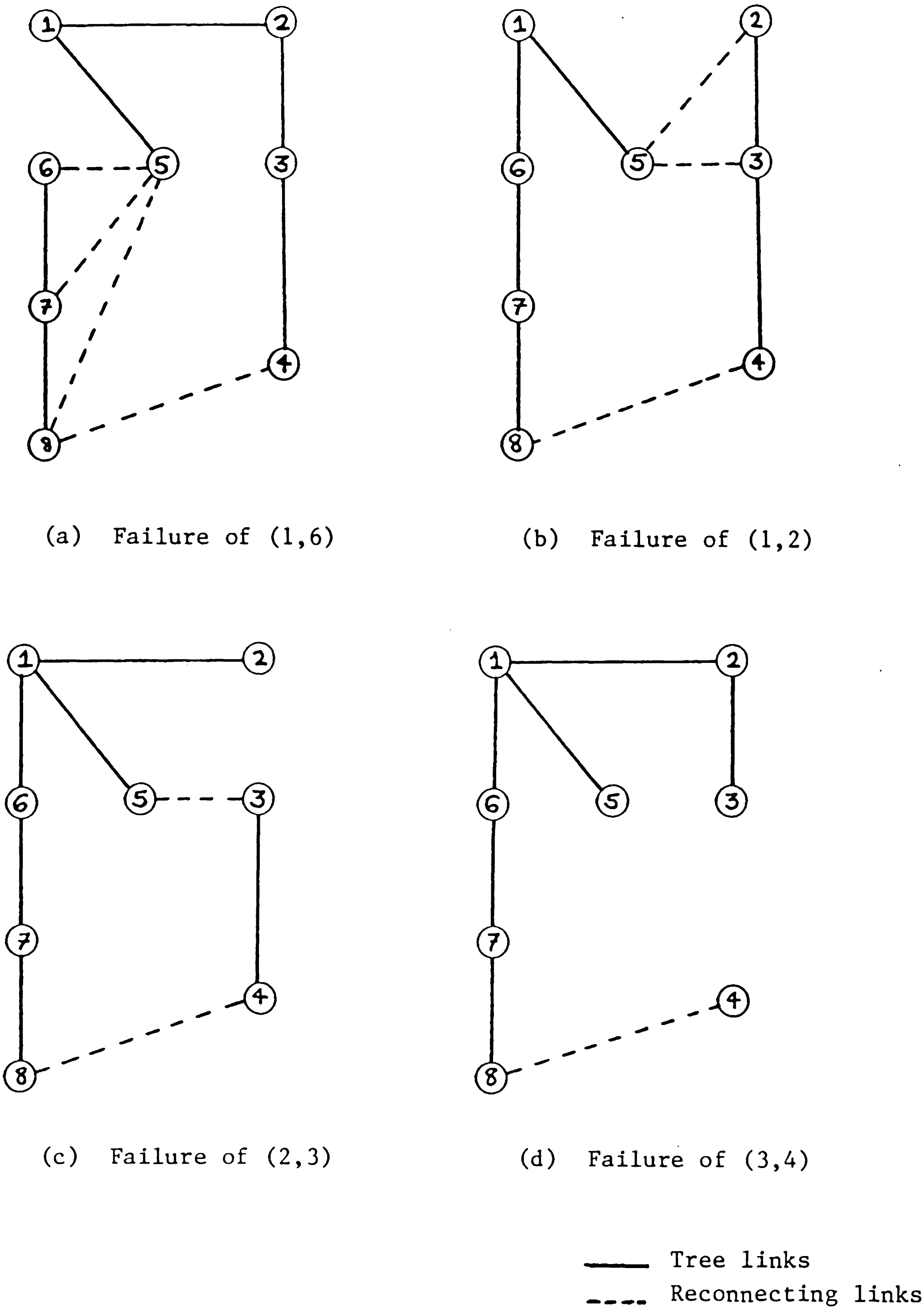


Figure 6. Network Showing Tree and Cotree

the set of nodes $N_2 = \{1, 2, 3, 4, 5\}$ which contains the source. Inspection of cotree links show that any one of the links {(4,8), (5,6), (5,7), (5,8)} will reconnect all the nodes back to the source. Such reconnecting set of links, $RC_{(i,j)}$, can be found for failure of each link (i,j) of T (Figs. 7a through 7d). It may so happen that corresponding to the failure of a link, it is not possible to find any reconnecting link. For example, if link (4,8) were not present in the cotree, for the failure of (3,4) we will not be able to find any link which reconnects node 4 to the source. Also, situations may arise when no link is required to reconnect. For example, if node 7 were also a source node, then corresponding to failure of link (1,6), all the nodes remain connected to one source or the other. Reconnecting sets corresponding to such situations are empty and contain no useful information. A trivial situation may arise when all the reconnecting sets are empty. Clearly, such situations arise when all the nodes are sources or when the cotree is empty. In such cases either no redundant link is required (all nodes sources) or no solution can be found (cotree empty) in which case, alternative strategies like adding links parallel to the existing links have to be resorted to.

For the present example, the reconnecting sets for each of the tree links are shown in Table 1. If any empty set were present, it should be discarded from further analysis. First, sets of lowest cardinality are selected. The set $RC_{(3,4)}$ has the lowest cardinality 1. Since (4,8) is the only member of this set, it must be selected. If more than one reconnecting link is present, then the link which has more number of occurrences in other reconnecting sets can be chosen. If a tie occurs in the number of occurrences, the link with smaller length is selected. Sets of higher cardinality are then examined to see if they contain any link which has been selected before from the sets of lower cardinality. If they do, then no link needs to be selected from those sets. Sets $RC_{(1,2)}$, $RC_{(1,6)}$, $RC_{(2,3)}$, $RC_{(6,7)}$ and $RC_{(7,8)}$ contain (4,8) hence no link is selected



(a) Failure of (1,6)

(b) Failure of (1,2)

(c) Failure of (2,3)

(d) Failure of (3,4)

Figure 7. Failure Analysis of Tree Links

Table 1. Reconnecting Sets From the Cotree

Failing Link (i,j)	Reconnecting set $RC_{(i,j)}$	Cardinality $\lvert RC_{(i,j)} \rvert$
(1,2)	{(2,5), (3,5), (4,8)}	3
(1,6)	{(5,6), (5,7), (5,8), (4,8)}	4
(1,5)	{(5,6), (5,7), (5,8), (3,5), (2,5)}	5
(2,3)	{(3,5), (4,8)}	2
(3,4)	{(4,8)}	1
(6,7)	{(5,7), (5,8), (4,8)}	3
(7,8)	{(5,8), (4,8)}	2

Reconnecting link (i,j)	Number of occurences $N_{(i,j)}$
(2,5)	2
(3,5)	3
(4,8)	6
(5,6)	2
(5,7)	3
(5,8)	4

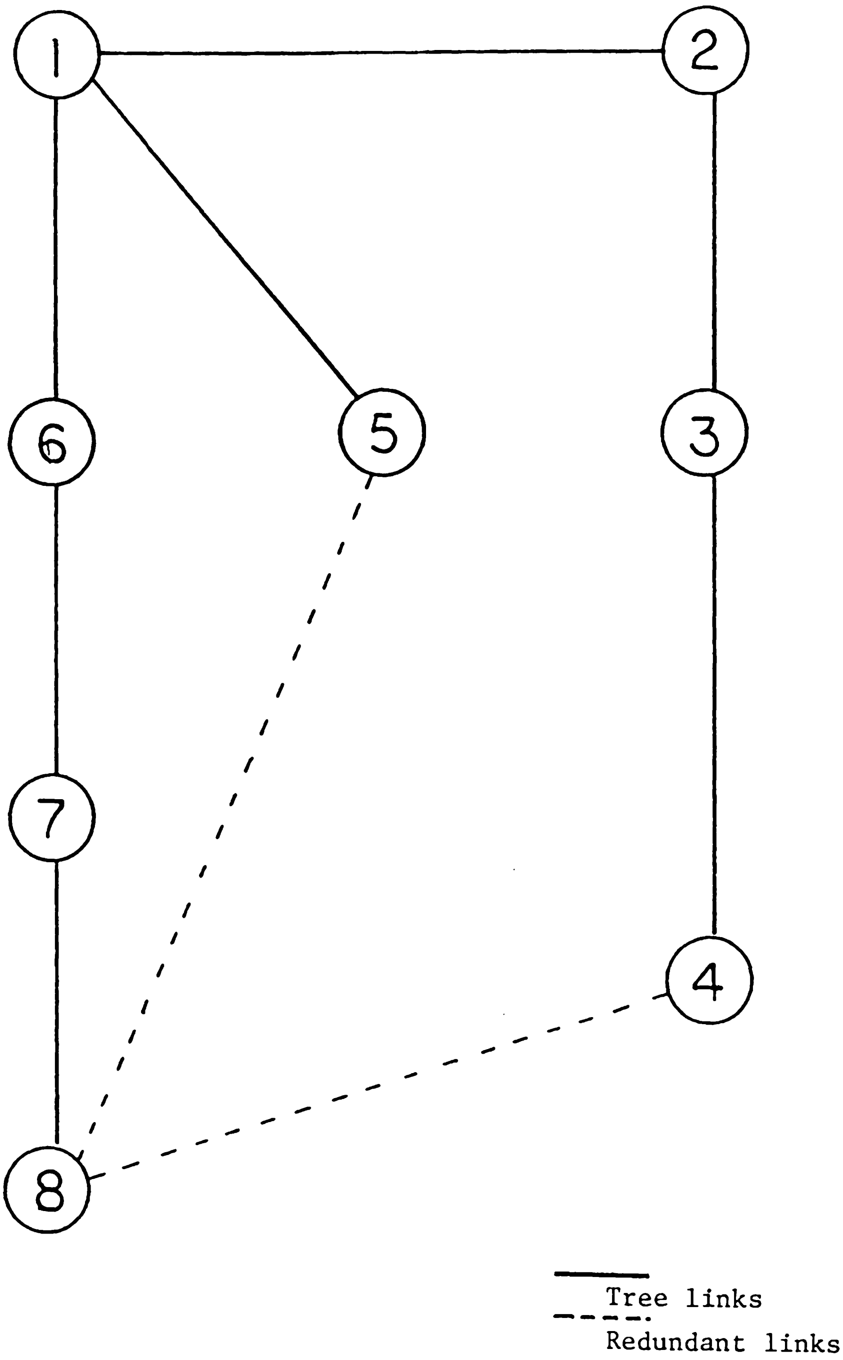


Figure 8. Network With Redundant Links

from these sets. The set $RC_{(1,5)}$ does not contain (4,8). Hence a link has to be selected from this set. Because the number of occurrences of (5,8) is the most in all the other sets, it is chosen. Since all the sets are exhausted, the procedure stops here. The solution for this problem is RL = {(4,8), (5,8) }, as shown in Fig. 8.

3.3 Algorithm REDUNDANCY

This section gives a step by step algorithm for obtaining redundant links. The algorithm uses procedure RECONNECT for obtaining reconnecting sets which is described in the subsequent section.

Algorithm REDUNDANCY

Purpose: To find a set of redundant links, RL.

Input: tree, T; set of potential links, $\mathscr{L}$; link lengths, $\{L_{(i,j)}, \ \forall (i,j) \in \mathscr{L}\}$

Output: RL

Method:

1. Initialization

   RL = Φ

2. For each reconnecting set $RC_{(i,j)}$ of tree link (i,j) record the cardinality $C_{(i,j)}$

   $\forall (i,j) \in T$ do :

   a. Find $RC_{(i,j)}$ using procedure **RECONNECT**

   b. set $C_{(i,j)} = |RC_{(i,j)}|$

3. Find the set of all reconnecting links, R. Check for the case of all reconnecting sets being empty; remove empty reconnecting sets from further analysis (this is equivalent to removing the corresponding tree links). Initialize number of occurrences, $N_{(p,q)}$, of each reconnecting link $(p,q) \in R$.

$$R = \bigcup_{(i,j) \in T} RC_{(i,j)}$$

If $R = \Phi$ then STOP

$\forall (p,q) \in R$ do:

$N_{(p,q)} = 0$

$\forall (i,j) \in T$ do:

If $RC_{(i,j)} = \Phi$ then $T = T - \{(i,j)\}$

4. Let $N_{(p,q)}$ be equal to the number of reconnecting sets in which (p,q) occurs.

$\forall (p,q) \in R$ do:

$\forall (i,j) \in T$ do:

If $(p,q) \in RC_{(i,j)}$ then $N_{(p,q)} = N_{(p,q)} + 1$

5. Find the redundant links by checking reconnecting sets in ascending order of cardinality. If two or more links in a reconnecting set have the same (max) number of occurrences, then choose the one amongst them having the minimum length. If the lengths also happen to be the same, then the selection is arbitrary.

Until $T = \Phi$ do:

a. select $(i', j') \in T$, such that $C_{(i', j')} = \min_{(i,j) \in T} \{C_{(i,j)}\}$

b. $n_m = 0$; $X = \Phi$

c. $\forall (p,q) \in RC_{(i', j')}$ do:

If $N_{(p,q)} = \max\limits_{(r,s)\in RC_{(i',j')}} \{N_{(r,s)}\}$ then

(i). $n_m = n_m + 1$

(ii). $X = X \cup \{(p,q)\}$

d. If $n_m = 1$ then {there is a unique link $X = \{p,q\}$ having max occur.}

(i). $RL = RL \cup X$

(ii). $\forall (i,j) \in T$ do:

If $X \in RC_{(i,j)}$ then $T = T - \{(i,j)\}$

Else {there are two or more links with max occur.}

(i). $RL = RL \cup \{(u,v) \mid L_{(u,v)} = \min\limits_{(r,s)\in X} \{L_{(r,s)}\}\}$

(ii). $\forall (i,j) \in T$ do:

if $(u,v) \in RC_{(i,j)}$ then $T = T - \{(i,j)\}$

End of REDUNDANCY

3.3.1 Construction of Reconnecting Sets

Deletion of link (p,q) from a tree T disconnects N, the set of all nodes, into two connected components N_1 and $N_2 = N - N_1$ (Fig. 7). Let the nodes be numbered from 1 to n. To find N_1 (and hence N_2) it is convenient to represent T by an (n x n) adjacency matrix A whose elements $a_{ij} = 1$ if link (i,j) is in T, $a_{ij} = 0$ otherwise (Smith, 1982). N_1 is obtained based on the following principle : if i is connected to j and j to k, then i is connected to k. This can be expressed in terms of logical addition of the j^{th} row to the i^{th} row and j^{th} column to the i^{th} column of the adjacency matrix. The rules for the logical addition are : $(0 + 0) = 0$, $(1 + 0) = (0 + 1) = (1 + 1) = 1$. Once N_1 is found the reconnecting set is obtained by recording links from the cotree, CT, whose one end (node) lies in N_1 and the other in N_2. It is noted that if both N_1 and N_2 contain a source, then no reconnecting link is required. If a reconnecting link cannot be found then a link parallel to link (p,q) (if possible) is taken as a reconnecting link.

Procedure RECONNECT

Purpose: To find reconnecting set $RC_{(p,q)}$, for the failing link (p,q) of a given tree T

Input: Tree, T; cotree, CT; node set, N; link (p,q); set of sources, S;

parallel link indicator, P. If P = 1 then link $(\overline{p}, \overline{q})$ parallel to (p,q) is possible else it is not.

Output: $RC_{(p,q)}$

Method:

1. Construct the adjacency matrix of T and denote it as A; initialize N_1 and $RC_{(p,q)}$ and remove the failing link (p,q)

   $a_{ij} = a_{ji} = 1$ if $(i,j) \in T$

   $a_{ij} = a_{ji} = 0$ otherwise

   set $a_{pq} = a_{qp} = 0$

   set $i = 1$, $N_1 = \{i\}$ and $RC_{(p,q)} = \Phi$

2. Find all the nodes which are connected to node i and add them to the component N_1

   $j = i + 1$

   Until $j > n$ do:

   a. If $a_{ij} \neq 0$ then

      (i). logically add row j to row i and column j to column i and delete row j and column j from A

      (ii). $N_1 = N_1 \cup \{j\}$

b. $j = j + 1$

$N_2 = N - N_1$

3. Check if both the components contain a source. If they do, then no reconnection is required; else find links from the cotree whose nodes are adjacent in N_1 and N_2

   If $\{N_1 \cap S = \Phi$ or $N_2 \cap S = \Phi\}$ then

   $$RC_{(p,q)} = \{ \bigcup_{(i,j)\in CT} (i,j) \mid (i \in N_1 \text{ and } j \in N_2) \text{ or } (i \in N_2 \text{ and } j \in N_1)\}$$

   If $\{ RC_{(p,q)} = \Phi$ and $P = 1 \}$ then

   $$RC_{(p,q)} = RC_{(p,q)} \cup \{(\overline{p}, \overline{q})\}$$

End of RECONNECT

Example Consider the network shown in Fig. 9. Let the failing link, $(p,q) = (1,2)$. $S = \{1\}$.

Step 1:

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

setting $a_{12} = a_{21} = 0$,

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

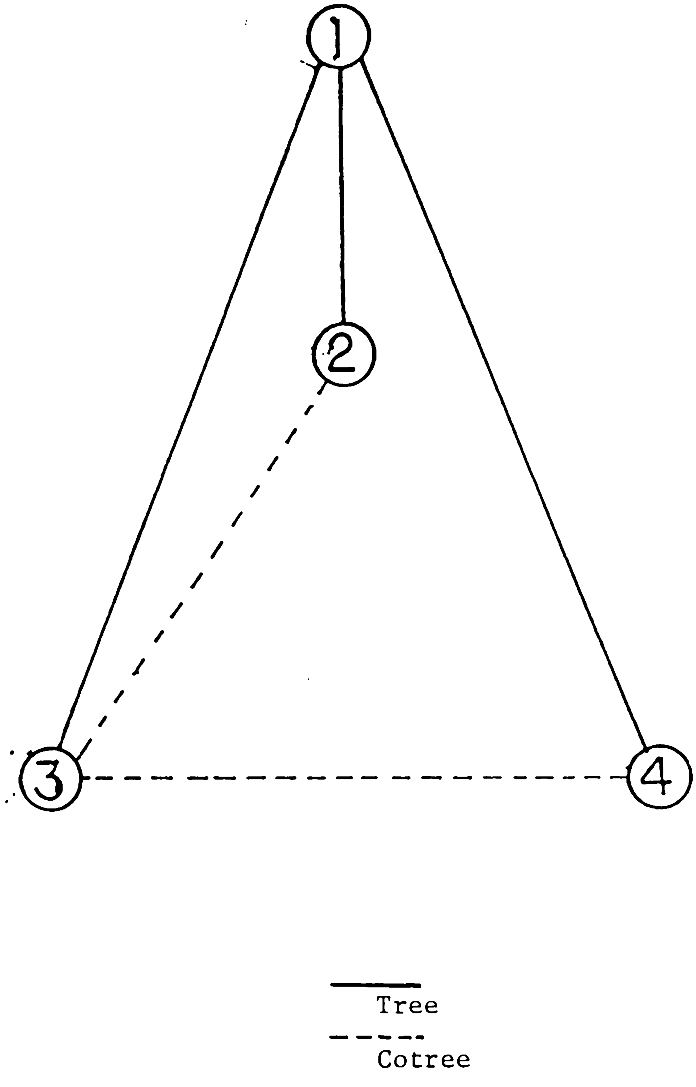


Figure 9. Network for Procedure RECONNECT

$$i = 1,\ \ N_1 = \{1\},\ \ RC_{(1,2)} = \{\Phi\}$$

Step 2:

$$a_{13} = 1 \neq 0$$

$$A = \begin{bmatrix} 1 & 0 & - & 1 \\ 0 & 0 & - & 0 \\ - & - & - & - \\ 1 & 0 & - & 0 \end{bmatrix}$$

$$N_1 = \{1, 3\}$$

$$a_{14} = 1 \neq 0$$

$$A = \begin{bmatrix} 1 & 0 & - & - \\ 0 & 0 & - & - \\ - & - & - & - \\ - & - & - & - \end{bmatrix}$$

$$N_1 = \{1, 3, 4\};\ \ N_2 = \{2\}$$

Step 3:

N_2 does not contain a source.

$(2,3) \in CT$; $2 \in N_2$ and $3 \in N_1$

$(3,4) \in CT$; but 3 and 4 $\notin N_2$

therefore $RC_{(1,2)} = \{(2,3)\}$

3.4 Selection of Redundant Link Diameters

Since the role of redundant links is to provide an alternate path in case of a tree link failure, it seems reasonable to provide minimum diameter for them. However, since higher diameter pipes have longer life (Clark, et al., 1982 ; Andreou, et al., 1987a, 1987b) it may be desirable to provide higher diameters. Again, the diameter selection depends on how much of unsatisfied demand in case of failure needs to be met. A high level of service may require large diameters for the redundant links.

Moreover, since the role of the redundant links is also to provide better circulation and avoid sedimentation (Rowell, 1979), it is desirable to keep them in operating condition most of the time. Thus when the redundant links are added, the whole network should be able meet the demands at specified pressure levels. In this study, a trial and error procedure is used to obtain the diameters of the redundant links :

1. Assume a minimum diameter for all the redundant links.

2. Analyze the full network. If all the pressure requirements are met then STOP.

3. Check the number of trials for adjusting the diameters of the redundant links; if the number of trials are less than that specified, revise the diameters, go to step 2. Otherwise go to step 4.

4. Redesign the optimal tree using the linear program (Problem P2) by increasing the minimum head requirements at the nodes at which the requirements are violated. Go to step 1.

This type of procedure was found to converge to a feasible solution in about five to six trials for the networks studied by this author. Chapter 4 discusses the application of the this procedure for two networks. In the first problem, a small single source network is considered wherein the minimum diameter for the redundant links suffice. In the second large size multiple source network problem, tree network needed to be redesigned using the linear program. However the additional cost of raising the minimum head requirement is minimal.

4.0 Application

4.1 Introduction

In this chapter two example problems are solved using the methodology developed in Chapters 2 and 3. The first one is a single source problem and the second is a multiple source problem. The single source problem is taken from Alperovitz and Shamir (1977) who solved the problem using the Linear Programming Gradient (LPG) technique. Quindry, et. al. (1981) and Fujiwara, et al. (1987) have successively improved the LPG technique and have obtained better solutions with lesser costs. The solution obtained by Algorithm TREESEARCH is even better than the last two. The second example is adapted from Rowell and Barnes (1982). The size of this problem is representative of a real life water distribution system. The optimal tree obtained from Algorithm TREESEARCH has lower cost than the shortest path tree and the tree obtained by the Nonlinear Minimum Cost Flow model of Rowell and Barnes (1982) which is briefly described in section 2.4.1.2. The selection of redundant links and their diameters for these two networks is demonstrated. The results and Figures are given in Appendix D.

4.2 Single Source Network

The network has 8 candidate links and 7 nodes (Fig. 14). It is fed by gravity from a constant head reservoir at node 1. It is assumed that parallel links are not possible. The minimum head requirements and flow demands at each node are shown in Table 3. Available diameters and their costs are given in Table 4. Length of each link is 1000 m. A Hazen-Williams coefficient of 130 is to be used for each link. No pumps are to be installed in the system. Minor losses are neglected.

From this network fifteen different spanning trees can be generated as shown in Fig. 15 (Rowell, 1979). The first four trees are the shortest path trees. First two trees were used as initial guesses for the algorithm TREESEARCH. The costs of these two trees were found to be 413,918 and 430,386 respectively. The implementation of algorithm TREESEARCH for each initial guess resulted in the same optimal tree layout, tree number 3 of Fig. 15, which has a cost of 399,667 units. The results for this tree are shown in Tables 5 and 6. For this tree layout the failure analysis gives reconnecting sets for each link as shown in Table 7. Clearly, link (5,7) from $RC_{(6,7)}$ has to be chosen. Since all other reconnecting sets contain (5,7) (except $RC_{(1,2)}$ which is empty and hence discarded), link (5,7) is the only redundant link required. A 1 inch diameter was assumed for this link and the network was analyzed using the computer program KYPIPEF (Wood, 1985). All the constraints were satisfied. The complete network is shown in Fig. 16 and the results are shown in Tables 8 and 9. The network has a total cost of 401,667 units which is less than 415,271 which was obtained by Fujiwara, et. al. (1987).

Fujiwara, et al. imposed the requirement that the hydraulic gradients for each link should lie between 0.0005 and 0.05 and that every link should have a minimum flow of 1 $m^3/$ hr.. Thus no link could be eliminated. To make a better comparison with Fujiwara's results, both the remaining links (4,5) and (5,7) were added to the tree. A minimum diameter of 1 inch was assumed for each link. the resulting network was analyzed using KYPIPEF and the pressures were found to be within the bounds. Table 10 gives a comparison of flows with those of Fujiwara et. al. and also shows the optimal link lengths. Optimal heads are given in Table 11. A comparison of the cost of this solution with the costs obtained by previous researchers is given in Table 12. The cost of 403,667 units is less than that obtained by Fujiwara, et. al.

For each initial tree, 8 different trees were visited. This means solution of 8 linear programs, each time. Fujiwara's results show that 33 linear programs were required for the LPG method for a particular initial feasible flow distribution. However, for this example it is not possible to show the full power of algorithm TREESEARCH because there are only 15 possible trees and almost any starting tree will suffice to arrive at the optimal tree. A large size network is therefore considered in the next example. The purpose of the next example is also to show the application to a multiple source network.

4.3 Multiple Source Network

An application of algorithm TREESEARCH using Linear Minimum Cost Flow Model (LMCF) described in section 2.4.1.2 is considered here. The network consists of two elevated reservoirs, nodes 1 and 2. There are 24 demand nodes and 51 potential links

(Fig. 17). It is assumed that parallel links are not possible. The nodal demands and minimum head requirements at the nodes are given in Table 13. Link lengths and cost data are given in Tables 14 and 15 respectively. The supply rates at the reservoirs are treated as unknowns. It is not possible to elevate the reservoir at node 2. There are more than 6.5×10^{10} trees for this network (Rowell, 1979) which makes the exhaustive enumeration impossible.

(a) Identification of supplies and the shortest path tree:

The LMCF model (section 2.4.1.2) was solved by using the linear programming code LINDO (Schrage, 1986). This identifies supply rates of 4,500 GPM at source node 1 and 10,500 GPM at the source node 2. The solution also allocates nodes 16-24 to source node 1 and the remaining nodes to source node 2 thus resulting in two disconnected shortest path trees as shown in Fig. 18. The two disconnected trees so formed are connected by link (14,18) which forms the shortest path between the two sources (Rowell, 1979). This is taken as the shortest path tree for the whole network (Fig. 18). The tree was optimized using the linear program (Problem P2) which gave a cost of $64,300.

(b) Solution from Nonlinear Minimum Cost Flow (NMCF) model (section 2.4.1.2):

The solution of NMCF depends on

(i). nodal demands and supplies $\{ q_i, \ \forall i \in N \}$

(ii). link lengths $\{ L_{(i,j)}, \ \forall (i,j) \in \mathscr{L} \}$

(iii). parameters C and $\overline{K}_{f(i,j)}$

The values of the above parameters for the present problem are the same as in Rowell (1979), except the parameter $\overline{K}_{f(i,j)}$. Rowell used a constant value of

$$\overline{K}_{f(i,j)} = \frac{1.286 \times 10^{-3}}{|J|} \quad \forall (i,j) \in \mathscr{L}$$

where J is the constant hydraulic gradient. In the present case,

$$\overline{K}_{f(i,j)} = \frac{8.914 \times 10^{-6}}{|J|} \quad \forall (i,j) \in \mathscr{L}$$

Since this value is simply a constant multiple of that of Rowell, the objective function of NMCF model is simply scaled by a positive number. Thus the solution remains the same as given in Rowell. This tree is shown in Fig. 19. Rowell used Linear Programming Gradient (LPG) technique to optimize the tree. Since the link flows for this tree are determined from continuity equations (supplies and demands are known), the LPG technique (section 2.4.1.2) is equivalent to solving a single linear program, Problem P2. The tree was optimized using this linear program and the cost was found to be \$58,500.

(c) Solution from TREESEARCH:

The shortest path tree (Fig. 18) was used as an initial guess for the TREESEARCH and a tree was found (Fig. 20) which has a cost of \$ 58,400. This tree was used as an initial guess and a still cheaper tree was found (Fig. 21) with a cost of \$58,100. The results for this tree are given in Tables 16 and 17.

(d) Superposition of redundant links:

The tree obtained from TREESEARCH was adopted as the optimal tree and the failure analysis was performed using the algorithm REDUNDANCY. The results of the computer program written for this purpose gives RL = {(9,15), (10,25), (11,12), (20,22), (22,23), (3,26)}. First a minimum diameter of 6 inches was assumed for all the redundant links and the resulting network was analyzed using KYPIPEF and the minimum head requirement at node 9 was violated. Various diameters were subsequently tried for all the redundant links but minimum head requirement at node 9 was still violated. To overcome this problem, the minimum head requirement at node 9 was increased to 1309 ft. from 1302 ft. and the optimal tree was redesigned using the linear program. The cost of the tree increased by $400. Again a minimum diameter of 6 inches was assumed for all the redundant links and the network was analyzed. All the head requirements were met. The final network is shown in Fig. 22. The results are shown in Tables 18 and 19. It has a cost of $72,724 ($58,500 for the tree which includes $8,955 for the reservoir elevation at node 1, and $14,224 for the redundant links).

5.0 Conclusions and Recommendations

The principal contribution of this thesis is the heuristic optimization procedure made up of algorithms TREESEARCH and REDUNDANCY. Even in large water distribution systems, the number of redundant pipes is not large. Right of way decisions and increased cost tend to reduce the number of redundant pipes. In such networks the cost is dominated by the tree links which is ideal for the TREESEARCH algorithm. The results from the example problems from literature tend to accentuate this conclusion. The REDUNDANCY algorithm provides alternate paths between source nodes and demand nodes in case of failures. Both the algorithms combined together form a powerful methodology for water distribution system design.

With regard to future research two suggestions are offered. The first one is related to the multiple source application in which a Linear Minimum Cost Flow (LMCF) model is used to identify the supply quantities at the sources. The LMCF model does not take into account the pressure constraints. One may consider using Linear Programming Gradient (LPG) technique in place of LMCF within Algorithm TREESEARCH. This will result in increased computational burden and numerical experimentation is required to evaluate the trade-off between computational effort and the improvement in the optimal solution obtained.

Secondly, Algorithm REDUNDANCY treats link failures deterministically and independent of pipe and soil properties. Some of the recent studies (Andreou et. al., 1987a, 1987b) address the pipeline failures probabilistically and distributional properties are derived based on pipe and soil properties and the age of the pipes. Such a random failure model should be included within the methodology.

References

Abadie, J., ed., *Integer and Nonlinear Programming*, North Holland, 1970.

Alperovitz, E., and Shamir, U., "Design of Optimal Water Distribution Systems," *Water Resources Research*, Vol. 13, Dec., 1977, pp. 885-900.

Andreou, S. A., Marks, D. H. and Clark R. M., "A New Methodology for Modelling Break Failure Patterns in Deteriorating Water Distribution Systems : Theory," *Advances in Water Resources*, Vol 10, March, 1987a, pp. 1-10.

Andreou, S. A., Marks, D. H. and Clark R. M., "A New Methodology for Modelling Break Failure Patterns in Deteriorating Water Distribution Systems : Application," *Advances in Water Resources*, Vol 10, March, 1987b, pp. 10-20.

Artina, S., "The Use of Mathematical Programming Techniques in Designing Hydraulic Networks" *Meccanica*, Sept., 1973, pp. 158-167.

Berge, C., *Graphs*, Elsevier Science, 1985.

Bhave, P. R., "Selecting Pipe Sizes in Network Optimization by LP," *Journal of the Hydraulics Division*, ASCE, Vol. 105, No. HY8, Aug., 1979, pp. 1019-1025.

Bhave, P. R., "Optimization of Gravity-Fed Water Distribution Systems : Theory," *Journal of Enviornmental Engineering*, ASCE, Vol. 109, No. 1, Feb., 1983a, pp. 189-205.

Bhave, P. R., "Optimization of Gravity-Fed Water Distribution Systems : Application," *Journal of Enviornmental Engineering*, ASCE, Vol. 109, No. 2, April, 1983b, pp. 383-395.

Billinton, R. and Allan, R. N., *Reliability Evaluation of Engineering Systems : Concepts and Techniques*, Pitman, London, 1983.

Boland, J. J., "The Requirement of Disaggregate Forecasts," *ASCE Fall Mtg.*, Hollywood, Fla., Oct. 27-31, 1980.

Boland, J. J., "A Research Agenda for Municipal Water Demand Forecasting," *Journal of American Water Works Association*, 1983, pp. 20-23. (as cited in Boland, 1980).

Choate, P., and Walter, S., *American in Ruins*, Council State Planning Agencies, Washington, D. C., 1981. (as cited in Mays and Cullinane, 1986).

Christofides, N., *Graph Theory*, Academic Press, 1975.

Clark, R. M., Stafford, C. L. and Goodrich, J. A., "Water Distribution Systems : A Spatial Cost Evaluation, " *Journal of Water Resources Planning and Management* , ASCE, Vol. 108, WR3, Oct., 1982, pp. 243-257.

Collins, M. A., "Pitfalls in Pipe Network Analysis Techniques," *Transportation Engineering Journal*, TE5, Sept., 1980, pp. 507-521.

Collins, M. A., Cooper, L. and Kennington, J. L., "Multiple Operating Points in Complex Pipe Networks," *Journal of the Hydraulics Division*, ASCE, HY3, March, 1979, pp. 229-245.

Collins, A. G., and Johnson, R. L., "Finite Element Method for Water Distribution Networks," *Journal of the American Water Works Association*, Vol. 67, July, 1975 , pp. 385-389.

Damelin, E., Shamir, U. and Arad, N., "Engineering and Economic Evaluation of the Reliability of Water Supply," *Water Resources Research*, Vol 8, No. 4, Aug., 1972, pp. 861-877.

Davis, C. V., and Sorensen, K. E., eds., chs. 36-37, *Handbook of Applied Hydraulics*, McGraw-Hill, NY, 1969.

Delfino, W. C. D., "Optimal Design of Water Distribution Pipeline Networks," Ph. D. thesis, Case Western Reserve University, June, 1973. (as cited in Rowell, 1979).

Dijkstra, E. W., "A Note on Two Problems in Connection with Graphs," *Numeriche mathematik*, Vol. 1, 1959, pp. 269-271.

Dillingham, J. H., and Cleasby, J. L., "New Computer Program for Hardy Cross Analysis" CIRAS, Iowa State University, Ames, Iowa, June, 1965. (as cited in Collins and Johnson, 1975).

Fiacco, A. V., *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, 1983, pp. 69-83.

Fujiwara, O., Jenchaimahakoon, B., and Edirisinghe, N. C. P., "A Modified Linear Programming Gradient Method for Optimal Design of Looped Water Distribution Networks," *Water Resources Research*, Vol. 23, No. 6, June, 1987, pp. 977-982.

Gessler, J., "Optimization of Pipe Networks," Proceedings 1982 International Symposium on Urban Hydrology, Hydraulics and Sediment Control, University of Kentucky, Lexington, Kentucky, July, 27-30, 1982, pp. 165-171.

Goulter, I. C. and Morgan, D. R., "Discussion of 'Obtaining Layout of Water Distribution Systems,' by W. F. Rowell and J. W. Barnes," *Journal of Hydraulics Division*, ASCE, Vol. 109, HY1, 1984, pp. 67-68.

Goodman, A. S., *Principles of Water Resources Planning*, Prentice Hall, 1984.

Harary, F., *Graph Theory*, Addision-Wesley, 1972.

Jacoby, S. L. S., "Design of Optimal Hydraulic Networks," *Journal of Hydraulics Division*, ASCE, Vol. 94, HY3, 1968, 641-666.

James, W. and Robinson, M. A., "A Program Package for Interactive Design of Optimal Pipe Networks for Any Climatic Region in Canada," *Canadian Journal of Civil Engineering*, Vol 6, 1979, pp. 365-372.

Jeppson, R. W., *Analysis of Flow in Pipe Networks*, Butterworth, 1976.

Kettler, A. J. and Goulter, I. C., "An Analysis of Pipe Breakage in Urban Water Distribution Networks," *Canadian Journal of Civil Engineering*, Vol. 12, 1985, pp. 286-293.

Lasdon, L. S., and Waren, A. D., *GRG2 User's Guide*, The University of Texas at Austin, May, 1982.

Linsley, R. K., and Franzini, J. B., *Water Resources Engineering*, McGraw-Hill, 1979.

Luenberger, D. G., *Introduction to Linear and Nonlinear Programming*, Addision-Wesley, 1973.

Mays, L. W. and Cullinane, M. J., "A Review and Evaluation of Reliability Concepts for Design of Water Distribution Systems," Report presented for Department of Army under Water Supply Design Work Unit 31733, U. S. Army Corps of Engineers, Washington, DC, Jan., 1986.

Millette, R. J., Hammonds, A. F., Pansing, M. F., Hansen, E. C., and Clark, P. J., "Aggressive Water : Assesing the Extent of the Problem," *Journal of American Water Works Association*, May, 1980, pp. 262-266.

Morgan, D. R., and Goulter, I. C., "Optimal Urban Water Distribution Design," *Water Resources Research*, Vol. 21, No. 5, May 1985, pp. 642-652.

Ormsbee, L., and Contractor D. N., "Optimization of Hydraulic Networks," Proceedings 1981 International Symposium on Urban Hydrology, Hydraulics and Sediment Control, University of Kentucky, Lexington, Kentucky, July 27-30, 1981, pp. 245-261.

Quindry, G., Brill, E. D., and Liebman, J. C., and Robinson, A., "Comments on 'Design of Optimal Water Distribution Systems' by E. Alperovitz and U. Shamir," *Water Resources Research*, Vol. 15, No. 6, 1979, pp. 1651-1654.

Quindry, G., Brill, E. D., and Liebman, J. C., "Optimization of Looped Water Distribution Systems," *Journal of Enviornmental Engineering Division*, ASCE, Vol. 107, EE4, 1981, pp. 665-679.

Ravindran, A., Philips, D. T., Solberg, J. J., *Operations Research : Principles and Practice*, John Wiley and Sons, 1987.

Rothfarb, B., Frank, H., Rosenbaum, D. M., Steiglitz, K., and Kleitman, D. J., "Optimal Design of Offshore Natural-Gas Pipeline Systems," *Operations Research*, Vol. 18, Nov-Dec., 1970, pp. 992-1020.

Rowell, W. F., "A Methodology of Optimal Design of Water Distribution Systems," thesis presented to the University of Texas at Austin, Tex., 1979, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Rowell, W. F., and Barnes J. W., "Obtaining Layout of Water Distribution Systems," *Jornal of Hydraulics Division*, ASCE, Vol. 108, HY1, 1982, pp. 137-148.

Schrage, L., *Linear, Integer and Quadratic Programmimg with LINDO*, Scientific Press, 1986.

Shamir, U. and Howard, C. D. D., "An Analytic Approach to Scheduling Pipe Replacement," *Journal of American Water Works Association*, May, 1979, pp. 248-258.

Shamir, U. and Howard, C. D. D., "Water Supply Reliability Theory," *Journal of American Water Works Association*, July, 1981, pp. 379-384.

Smith, D. K., *Network Optimization Practice: A Computational Guide*, Ellis Harwood, 1982.

Templeman, B., "Discussion of 'Optimization of Looped Water Distribution System' by Quindry et. al.," *Journal of Enviornmental Division, ASCE*, Vol. 108, EE3, 1982, pp. 599-602.

Texas Water Utilities Association, *Manual of Water Utility Operations*, 1979.

Walski, T. M., *Analysis of Water Distribution Systems* Van Nostrand Reinhold Company, 1984.

Washington Suburban Sanitary Commission, Maryland, *Design Manual*, Dec., 1981.

Wood, D. J. and Charles C. O., "Hydraulic Network Analysis Using Linear Network Theory," *Journal of the Hydraulics Division*, ASCE, Vol 98, HY7, July, 1972, pp. 1157-1171.

Wood, D. J., *Kentucky Pipe Hydraulic Network Analysis Program Manual*, Department of Civil Engineering, University of Kentucky, 1985.

Appendix A. Convexity and Optimality

Convex Sets

A set S is said to be a **convex set** if for any two points in the set the line joining those points is also contained in the set. Mathematically, S is a convex set if for any two vectors x^1 and x^2 in S, the vector $x = \lambda x^1 + (1 - \lambda)x^2$ is also in S for any number λ between 0 and 1. Fig. 10(a) shows a convex set while Fig. 10(b) illustrates a **nonconvex** set.

Convex and Concave Functions

A function $f(x)$ is said to be a **convex function** if and only if for any two points x^1 and x^2 and $0 \le \lambda \le 1$

$$f[\lambda x^1 + (1 - \lambda)x^2] \le \lambda f(x^1) + (1 - \lambda)f(x^2)$$

A function g is a **concave** function if and only if -g is a convex function. Figs. 10(c) and 10(d) illustrate a convex function and a concave function respectively. A linear function shown in Fig. 10(e) is both concave and convex.

Optimality

Mathematical programming is concerned with solving an optimization problem of the form

$$\text{Problem P0: minimize} \quad f(x)$$
$$\text{subject to : } x \in X$$

where x is an n-vector of decision variables, f is a real-valued function of x and X is a constraint set. Solving P0 means, finding $x^0 \in X$ such that

$$f(x^0) \leq f(x) \quad \forall x \in X \tag{1}$$

The point x^0 is said to be a **global optimum** of Problem P0. If strict inequality holds in (1), x^0 is the **unique** global optimum of P0. If (1) holds only for some neighborhood of x^0, then x^0 is a **local** optimum of P0. If (1) holds for several $x^0 \in X$, then P0 has **alternate optima**. Examples are shown in Figs. 10(f)-10(h).

Convex Programming Problem : Minimizing a convex function or maximizing a concave function over a convex constraint set is called a convex programming problem.

Theorem For a convex program, every local optimum is also a global optimum.

From the Theorem it follows that for a bounded convex program, any iterative optimization procedure, which in each iteration improves the objective function value, will eventually reach a global optimum. Nonconvex optimization problems may have local optima. In that case, if an iterative procedure reaches a local optimum all the neighboring points will show an inferior value of the objective function and the procedure would end. In most cases there is no criterion to find out whether the procedure ended at a global or local optimum.

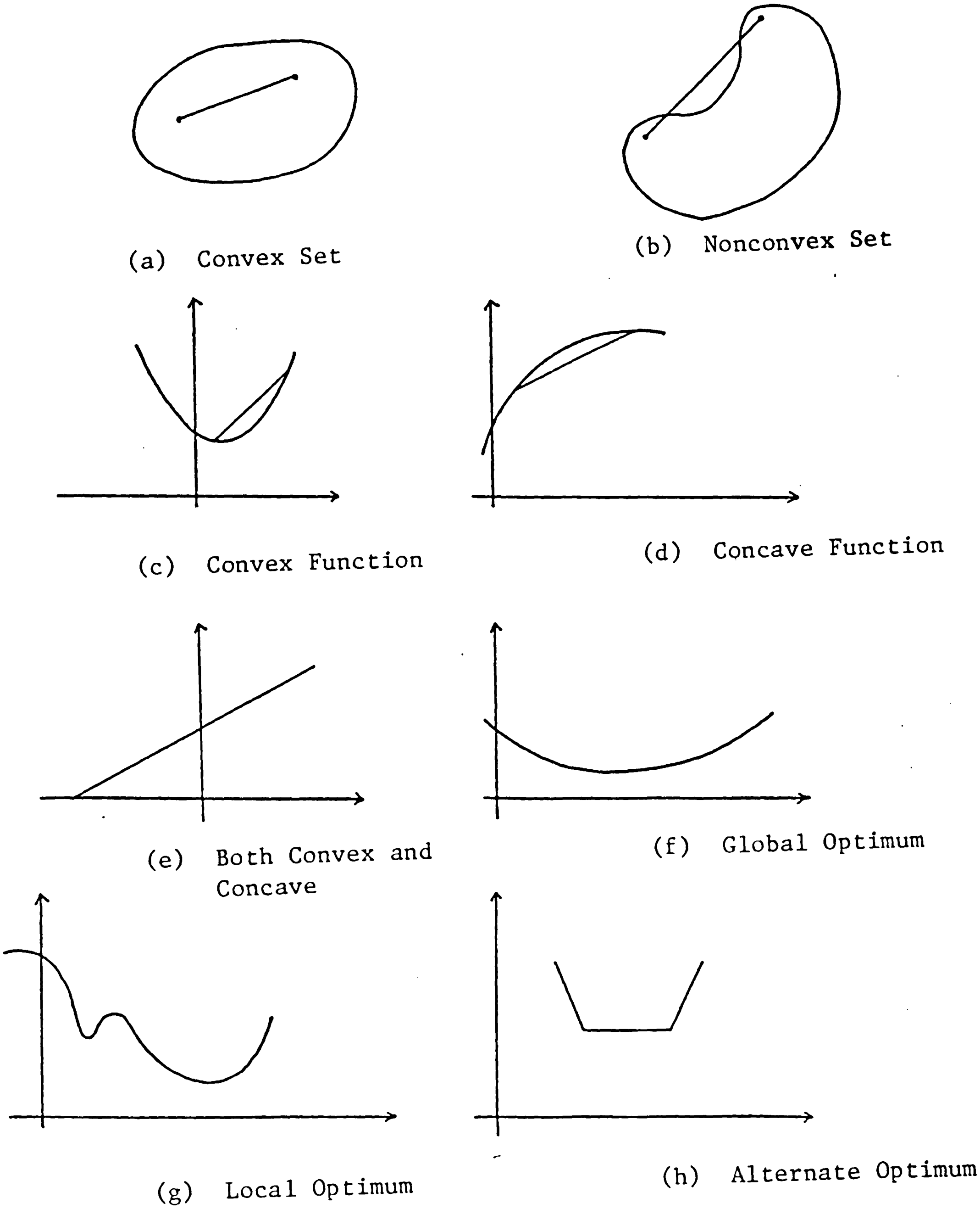


Figure 10. Convexity and Optimality

Example 1 Consider the following optimization problem:

Minimize: $f(x) = x_1 + x_2$

subject to:

$$f_1(x) = x_1 + x_2 \leq 1$$

$$f_2(x) = 2x_1 - x_2 \leq 0$$

$$f_3(x) = x_1 \geq 0$$

The constraint set formed by these linear constraints is convex (see Fig. 11a). The point A is the global optimum. In general, linear constraints form a convex feasible region.

Example 2 Objective function is convex but constraint set is nonconvex (see Fig. 11b).

Minimize: $f(x) = -x_1$

subject to:

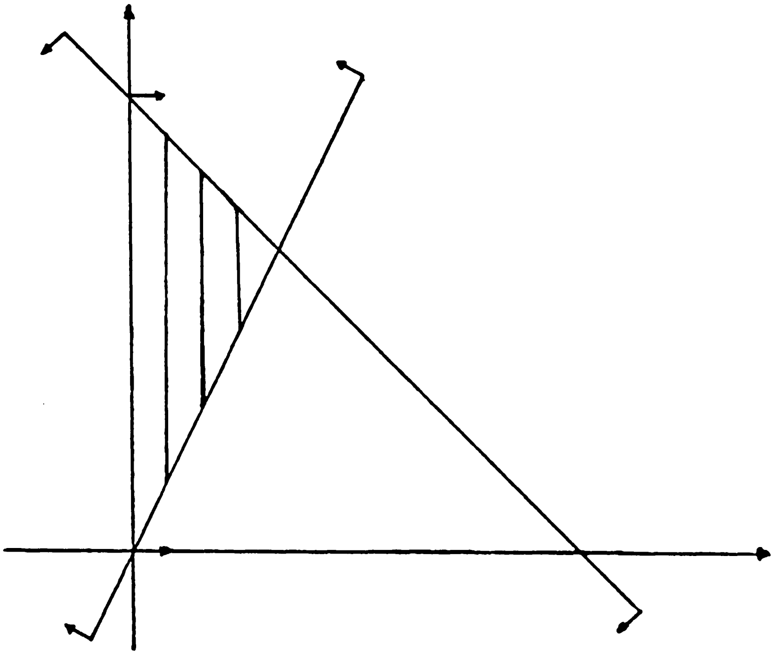
$$f_1(x) = x_2^2 - x_1 \geq 1$$

$$f_2(x) = x_2 \leq 1$$

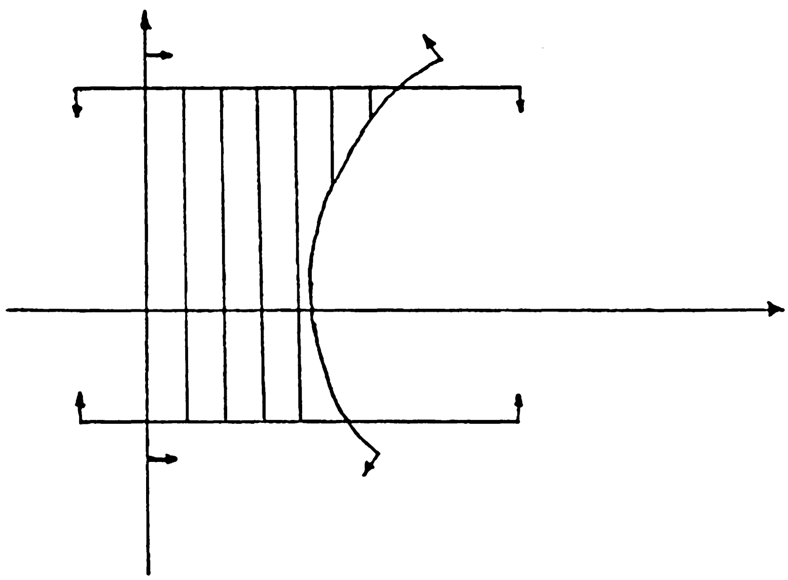
$$f_3(x) = x_2 \geq -1/2$$

$$f_4(x) = x_1 \geq 0$$

Point A is a global optimum but point B is a local optimum.



(a) Convex Constraint Set



(b) Nonconvex Constraint Set

Figure 11. Convex and Nonconvex Constraint Sets

Appendix B. Graph Theory : Some Definitions

A **simple graph** (or a network; referred to as a graph subsequently) G is a pair $(N, \mathscr{L})$ where N is a non-empty finite set of elements called **nodes** (or vertices or points) and $\mathscr{L}$ is a finite set of unordered pairs of distinct elements of N called **links** (or edges or lines). Two nodes u, v of a graph are said to be **adjacent** if the link (u,v) is in $\mathscr{L}$. A graph in which every pair of distinct nodes are adjacent is called a **complete graph**. A graph $G_s(N_s, \mathscr{L}_s)$ where N_s is a subset of N and $\mathscr{L}_s$ is a subset of $\mathscr{L}$, is called a **subgraph** of $G(N, \mathscr{L})$. For a **partial graph**, $N_s = N$ and $\mathscr{L}_s$ is a subset of $\mathscr{L}$.

A **path** between nodes i_1 and i_2 of a graph is an alternating sequence of nodes and links $\{ i_1, (i_1, i_2), i_2, \ldots ,(i_{m-1}, i_m), i_m \}$. A path becomes a **cycle** (or loop) if $i_1 = i_m$. Each link is associated with a weight (e.g., length). **Length** of a path is the sum of the lengths of all the links on the path. A graph is said to be **connected** if for every pair of nodes a path exists; otherwise it is said to be disconnected. **Union** of two graphs $G_1(N_1, \mathscr{L}_1)$ and $G_2(N_2, \mathscr{L}_2)$ is defined as a graph $G[(N_1 \cup N_2), (\mathscr{L}_1 \cup \mathscr{L}_2)]$. A disconnected graph can be expressed as a union of a finite number of connected graphs; each such connected graph is called a **connected component** (or component).

Example Consider the graph shown in Fig. 12(a). For this graph,

- $N = \{1,2,\ldots,8\}$
- $\mathscr{L} = \{(1,2), (3,4), (4,5), (5,7), (3,6), (6,7), (4,8)\}$.
- $\{3, (3,4), 4, (4,5), 5, (5,8), 8\}$ is a path between the nodes 3 and 8.
- $\{4, (4,5), 5, (5,8), 8, (8,4), 4\}$ is a cycle.
- the graph is disconnected because no path exists between any node in $\{1,2\}$ and any node in $\{3,4,\ldots,8\}$
- $G_1(N_1, \mathscr{L}_1)$ where $N_1 = \{1,2\}$ and $\mathscr{L}_1 = \{(1,2)\}$ is a component of G.

Trees

A **forest** is defined as a graph having no cycles, a connected forest is a **tree** (Fig.12b-12c). A tree has the following properties:

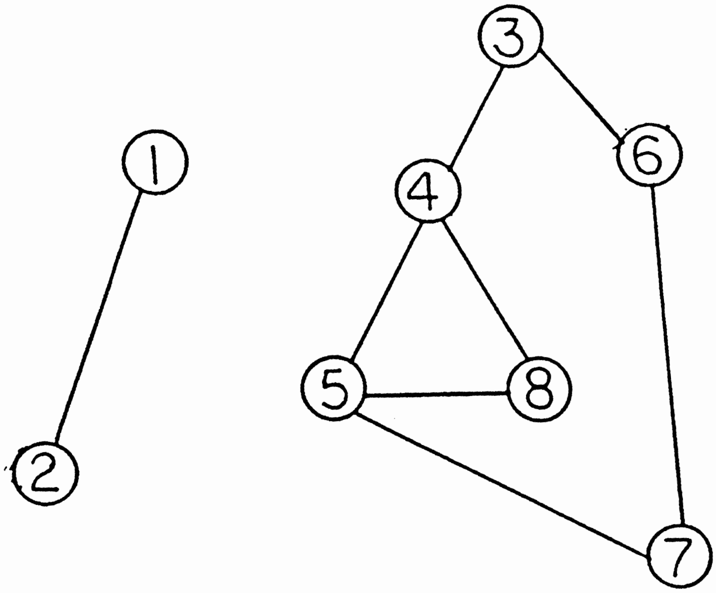
(i). a tree with n nodes has (n-1) links

(ii). a unique path exists between any two nodes of a tree

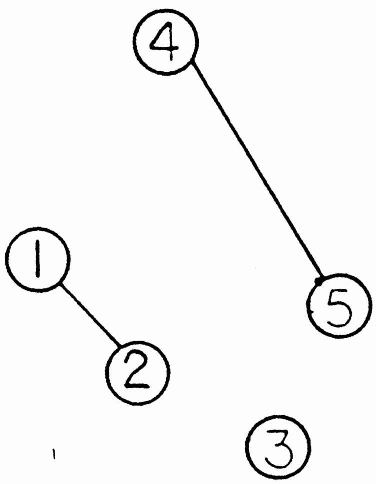
(iii). addition of a new link to a tree creates exactly one cycle.

(iv). removal of a link from a tree results in two disjoint trees.

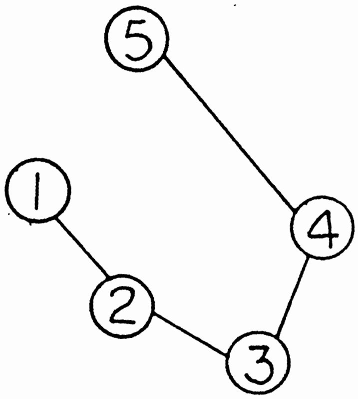
A **shortest path tree** of a graph G with a source node s is a partial graph of G formed by those links which yield the shortest path between s and every other node of G.



(a) Graph



(b) Forest



(c) Tree

Figure 12. Graphs

Appendix C. Network Optimization by Nonlinear Programming

In this section, example 1[a] of Jacoby (1968) is solved using the Generalized Reduced Gradient (GRG) (Abadie, 1970) technique. The network is shown in Fig. 13(a). It is fed through a pump at node 1. The objective is to minimize the link and pumping costs:

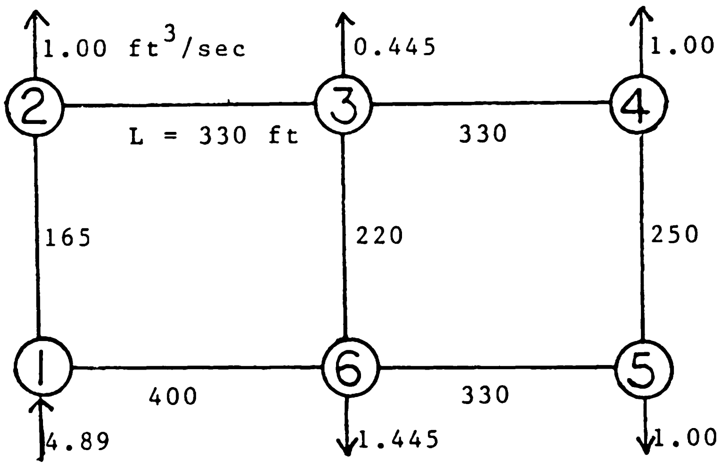
$$\text{Minimize} \quad \sum_{\forall (i,j)} (25.7D_{(i,j)} - 3.9)L_{(i,j)} + 435H_p \tag{1}$$

where $D_{(i,j)}$ and $L_{(i,j)}$ are the diameter and the length of a link (i,j); H_p is the pumping head at node 1.

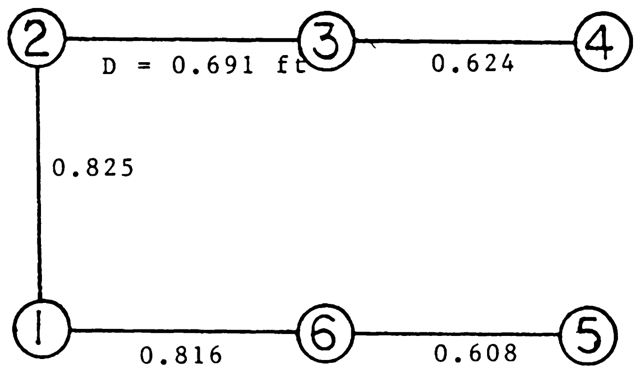
The constraints are the continuity equations and the energy equations. The Darcy-Weishbach formula is used for the headloss:

$$H_{L(i,j)} = f_{(i,j)}L_{(i,j)}\frac{Q^2_{(i,j)}}{39.7D^5_{(i,j)}} \tag{2}$$

where $f_{(i,j)}$ is given by the Colebrook-White equation:



(a) Jacoby's Network



(b) Tree from GRG2

Figure 13. Example From Jacoby (1968)

$$\frac{1}{\sqrt{f_{(i,j)}}} = 1.74 - 2.00 \log_{10}\left[\frac{2k_s}{D_{(i,j)}}\right] \tag{3}$$

a value of 0.01 is assumed for the roughness factor k_s. Since the friction factor $f_{(i,j)}$ is dependent on the decision variable $D_{(i,j)}$, Eqn. (3) is also included as a constraint. Jacoby implicitly used a minimum diameter of 0.2 ft. Here, this restriction is relaxed; diameters are allowed to take zero values.

The solution is achieved through GRG2 code (Lasdon and Waren, 1982). The tree network obtained is shown in Fig. 13(b). The total cost is $25,942 ($21,042 for links, $4,902 for pumping) as opposed to $28,703 obtained by Jacoby ($23,756 for links, $4,947 for pumping). The results are compared in Table 2.

It is possible to obtain the tree network using GRG2 for such small size networks; however, for large realistic problems it is difficult to obtain a near optimal tree because of the nonconvexity of the constraint set.

Table 2. Comparison of GRG Results With Jacoby (1968)

Link	Diameter(ft)		Flow(ft^3/s)	
	GRG	Jacoby	GRG	Jacoby
(1,2)	0.825	0.834	2.44	2.31
(2,3)	0.691	0.667	1.44	1.31
(3,6)	0.014	0.200	0.00	-0.01
(3,4)	0.624	0.584	1.00	0.88
(4,5)	0.334	0.011	0.00	-0.12
(5,6)	0.608	0.667	-1.00	-1.12
(6,1)	0.816	0.834	-2.44	-2.58

Appendix D. Data and Results for the Single and Multi-source Networks

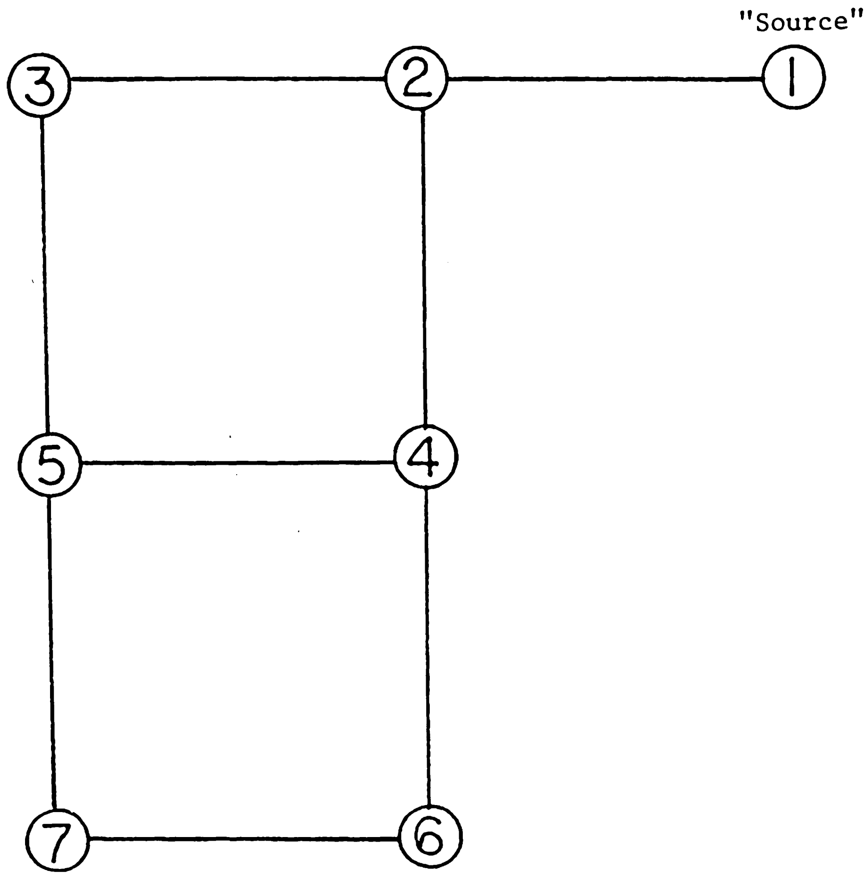


Figure 14. Single Source Network

Table 3. Nodal Data for the Single Source Network

Node i	Demand (cu. m/hr)	Min. Head (m)
1	-1120	210*
2	100	180
3	100	190
4	120	185
5	270	180
6	330	195
7	200	190

* fixed head node

Table 4. Cost Data for the Single Source Network

Diameter (inch)	Cost/meter	Diameter (inch)	Cost/meter
1	2.0	2	5.0
3	8.0	4	11.0
6	16.0	8	24.0
10	32.0	12	50.0
14	60.0	16	18.0
18	130.0	20	170.0
22	300.0	24	550.0

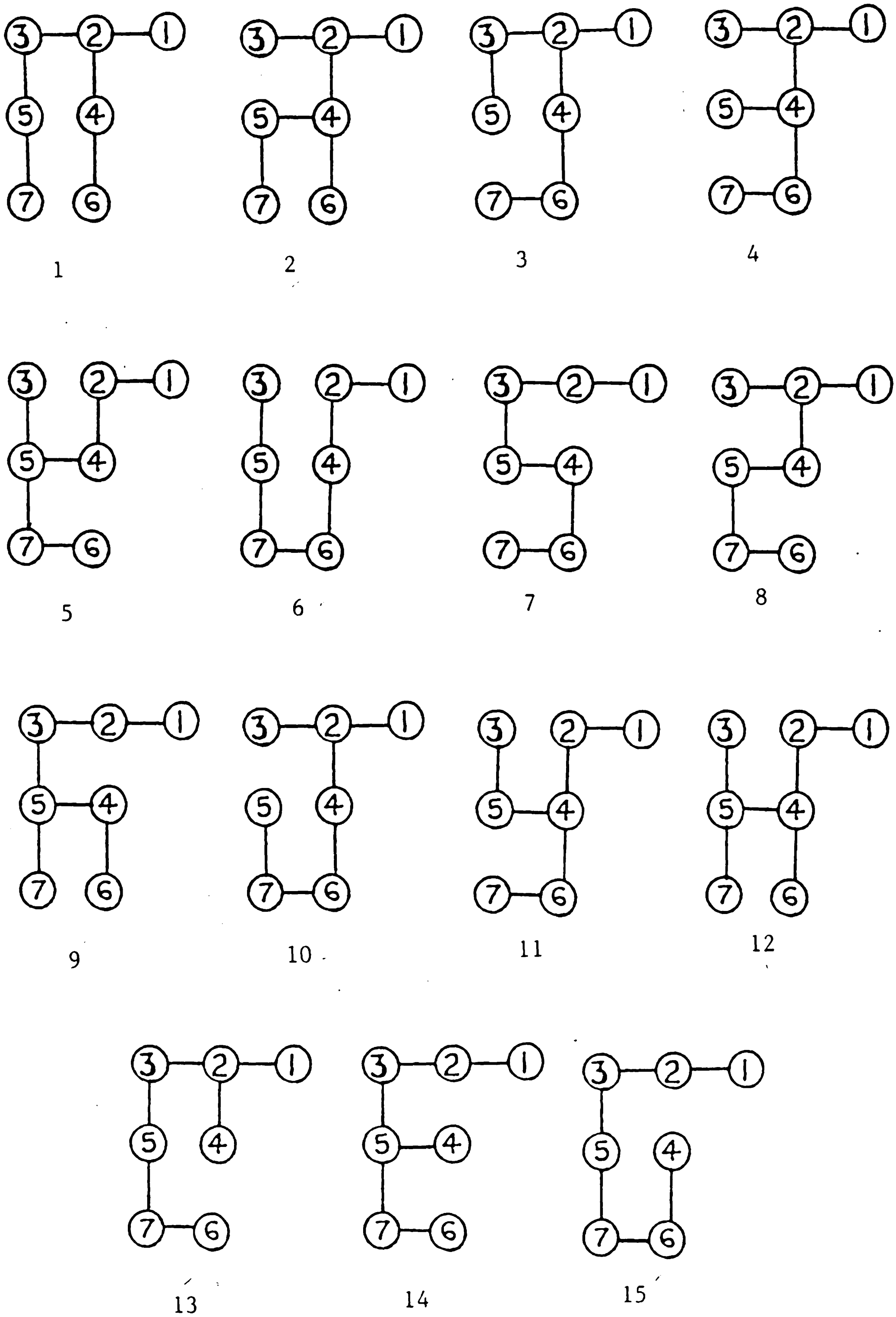


Figure 15. Spanning Trees of the Single Source Network

Table 5. Optimal Flows and Link Lengths From TREESEARCH for the Single Source Network

Link (i,j)	Flow (cu.m/hr)	Length (m)	Diameter (inch)
(1,2)	1120.00	1000.00	18.00
(2,3)	370.00	780.34	10.00
		219.66	12.00
(2,4)	650.00	1000.00	16.00
(3,5)	270.00	90.86	8.00
		909.14	10.00
(4,6)	530.00	314.96	14.00
		685.04	16.00
(6,7)	200.00	13.87	8.00
		986.13	10.00

Table 6. Optimal Heads From TREESEARCH for the Single Source Network

Node (i)	Head (m)
1	210.00
2	203.24
3	190.00
4	198.87
5	180.00
6	195.00
7	190.00

Table 7. Reconnecting Sets for the Single Source Network

Failing Link (i,j)	Reconnecting set $RC_{(i,j)}$	Cardinality $\lvert RC_{(i,j)} \rvert$
(1,2)	Φ	0
(2,3)	{(4,5), (5,7)}	2
(2,4)	{(4,5), (5,7)}	2
(3,5)	{(4,5), (5,7)}	2
(4,6)	{(5,7)}	1
(6,7)	{(5,7)}	1

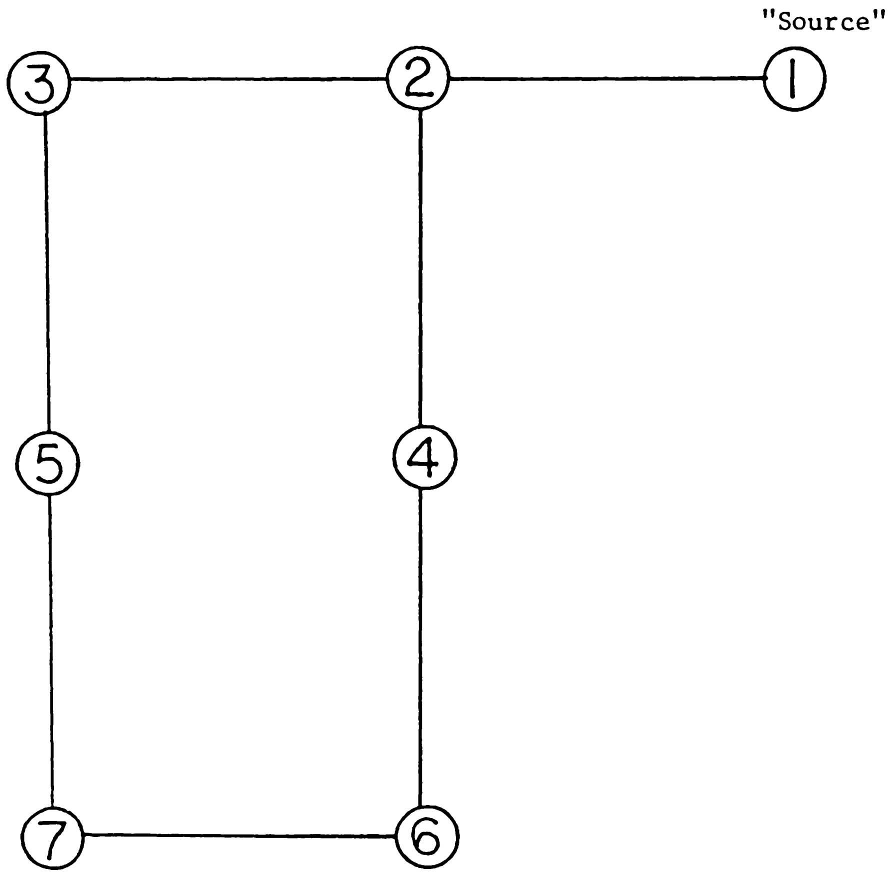


Figure 16. Application of TREESEARCH and REDUNDANCY to the Single Source Network

Table 8. Optimal Flows and Link Lengths From REDUNDANCY for the Single Source Network

Link (i,j)	Flow (cu.m/hr)	Length (m)	Diameter (inch)
(1,2)	1120.00	1000.00	18.00
(2,3)	368.92	780.34	10.00
		219.66	12.00
(2,4)	651.21	1000.00	16.00
(3,5)	269.04	90.86	8.00
		909.14	10.00
(4,6)	530.95	314.96	14.00
		685.04	16.00
(6,7)	200.76	13.87	8.00
		986.13	10.00
(5,7)*	1.02	1000.00	1.00

* Redundant link

Table 9. Optimal Heads from REDUNDANCY for the Single Source Network

Node (i)	Head (m)
1	210.00
2	203.23
3	190.03
4	198.84
5	180.06
6	195.00
7	190.00

Table 10. Comparison of Flows With Fujiwara et. al. (1987)

Link	Flow(m^3/hr)		Length(m)	Diameter(in)	Hydraulic Gradient
	Present	Fujiwara			
(1,2)	1120.00	1120	1000.00	18.00	0.0068
(2,3)	367.90	334.23	780.34	10.00	0.0132
			219.66	12.00	
(2,4)	652.23	685.68	1000.00	16.00	0.0044
(3,5)	268.03	234.32	90.86	8.00	0.0099
			909.14	10.00	
(4,6)	530.95	531.01	314.96	14.00	0.0038
			685.04	16.00	
(6,7)	200.76	201.01	13.87	8.00	0.0050
			986.13	10.00	
(7,5)*	1.02	1.01	1000.00	1.00	0.0098
(4,5)*	1.02	34.67	1000.00	1.00	0.0186

* Redundant links

Note: Link flows are above 1.00 m^3/hr and the hydraulic gradients are within the specified limits of 0.0005 to 0.05.

Table 11. Heads for the Network of Fujiwara et. al. (1987)

Node (i)	Head (m)
1	210.00
2	203.23
3	190.09
4	198.83
5	180.20
6	195.00
7	190.00

Table 12. Comparison of Cost for the Single Source Network With Previous Studies

Study	Cost
Alperovitz and Shamir (1977)	479,525
Quindry et. al. (1979)	441,522
Fujiwara et. al. (1987)	415,271
Present	403,667

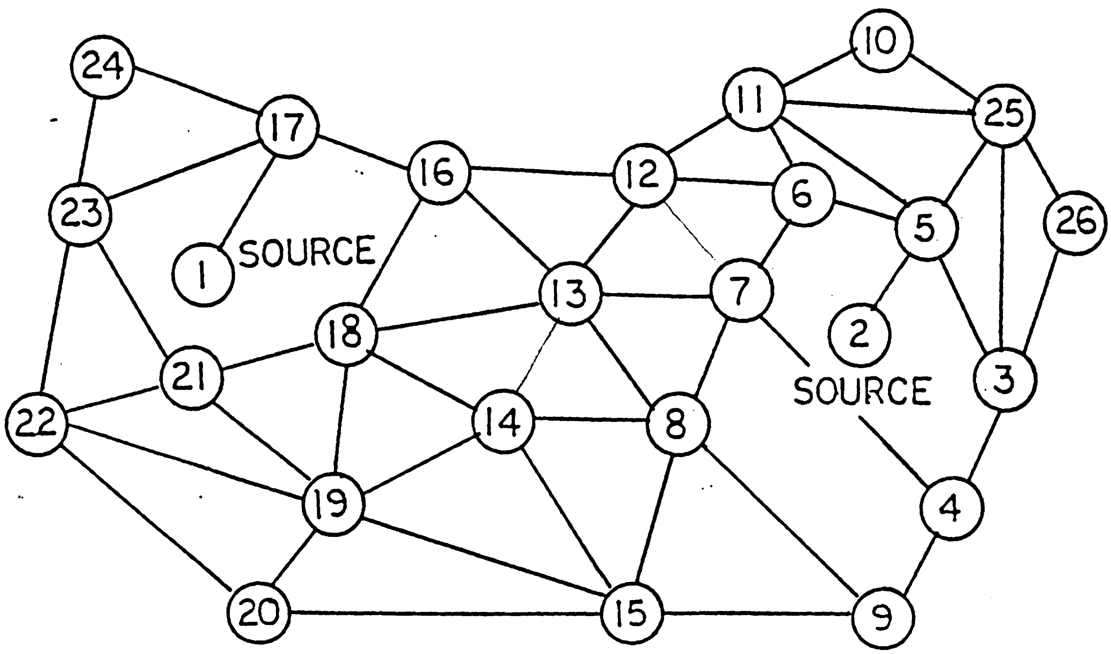


Figure 17. Multi-source Network

Table 13. Nodal Data for the Multi-Source Network

Node i	Demand (GPM)	Min. Head (ft)
1*	-	1404.0
2*	-	1347.50
3	800	1310.0
4	1150	1305.0
5	850	1339.0
6	290	1346.0
7	550	1319.0
8	750	1308.0
9	1040	1302.0
10	530	1306.0
11	560	1330.0
12	1060	1290.0
13	450	1305.0
14	530	1306.0
15	1050	1290.0
16	490	1420.0
17	610	1326.0
18	240	1317.0
19	590	1288.0
20	580	1279.0
21	770	1314.0
22	190	1312.0
23	810	1305.0
24	220	1316.0
25	520	1330.0
26	370	1324.0

* source node

Table 14. Link Data for the Multi-Source Network

Link (i,j)	Length (ft)	Link (i,j)	Length (ft)
(1,17)	450.0	(12,13)	1370.0
(2, 5)	60.0	(12,16)	2990.0
(3, 4)	2685.0	(13,14)	1560.0
(3, 5)	2400.0	(13,16)	3120.0
(3,25)	2250.0	(13,18)	4020.0
(3,26)	1555.0	(14,15)	3010.0
(4, 7)	5820.0	(14,18)	2510.0
(4, 9)	3480.0	(14,19)	3960.0
(5, 6)	1800.0	(15,19)	4490.0
(5,11)	2510.0	(15,20)	5620.0
(5,25)	1535.0	(16,17)	1380.0
(6, 7)	1260.0	(16,18)	2500.0
(6,11)	1210.0	(17,23)	5110.0
(6,12)	2920.0	(17,24)	4710.0
(7, 8)	1695.0	(18,19)	2750.0
(7,12)	2210.0	(18,21)	2840.0
(7,13)	1780.0	(19,20)	1440.0
(8, 9)	4330.0	(19,21)	2720.0
(8,13)	1660.0	(19,22)	5180.0
(8,14)	1840.0	(20,22)	5570.0
(8,15)	2500.0	(21,22)	2200.0
(9,15)	3850.0	(21,23)	4040.0
(10,11)	1790.0	(22,23)	3510.0
(10,25)	2490.0	(23,24)	1800.0
(11,12)	2510.0	(25,26)	1650.0
(1,25)	3900.0		

Table 15. Cost Data for the Multiple Source Network

Diameter (inch)	Cost/ft ($)
6	0.73
8	1.06
10	1.41
12	1.78
14	2.17
16	2.58
18	3.00
20	3.44
22	3.89
24	4.35
26	4.82
28	5.30
30	5.80

reservoir elevation cost = $508.34/ft

The tabulated unit link cost values correspond to the following relationship (eqn. 2.10):

$$F_L = C_1 L_{(i,j)} D_{(i,j)}^{C_2}$$

where $C_1 = 0.0725$ and $C_2 = 1.29$ for length in feet, diameter in inches and cost in dollars.

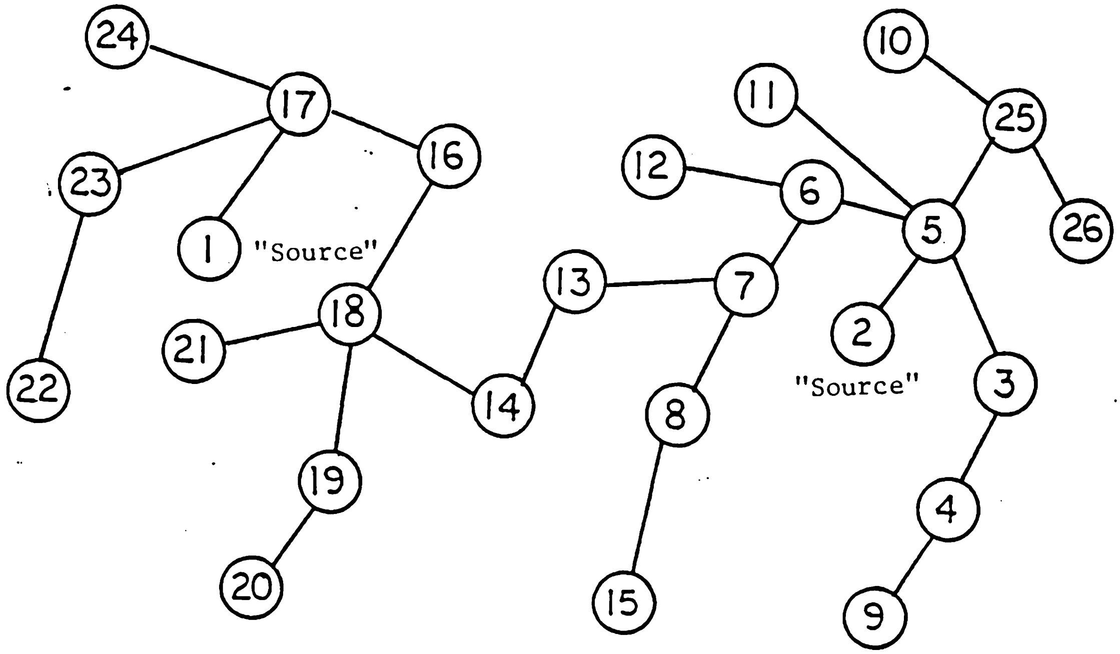


Figure 18. Shortest Path Tree for the Multi-source Network: Link (14,18) is the connecting link between the two components

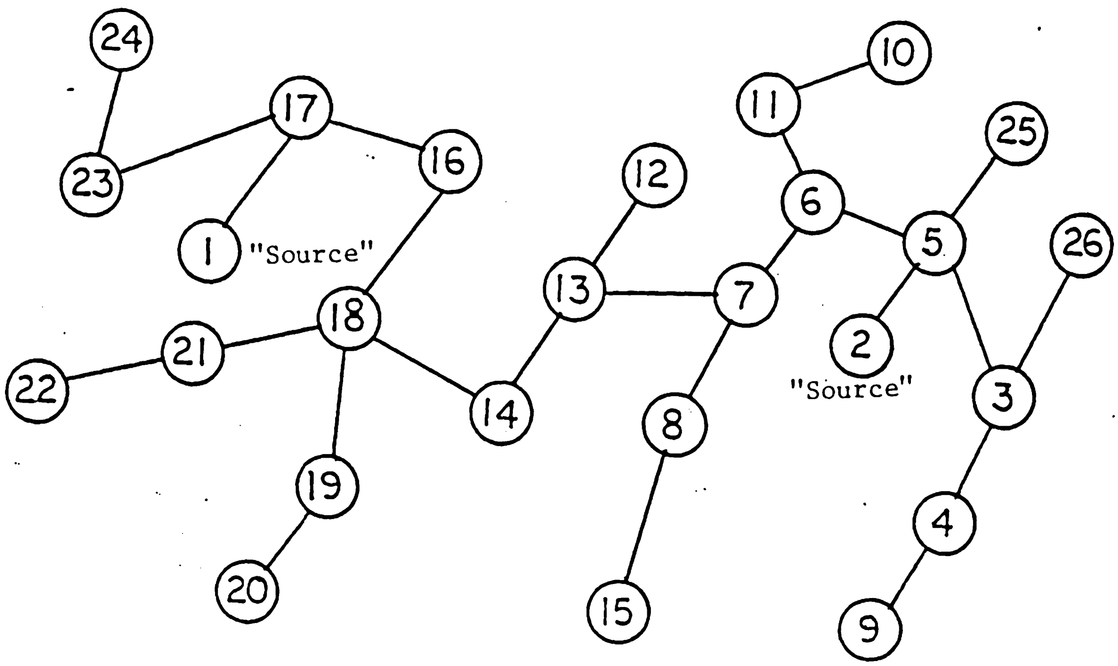


Figure 19. Optimal Tree for the Multi-source Network from Rowell (1979)

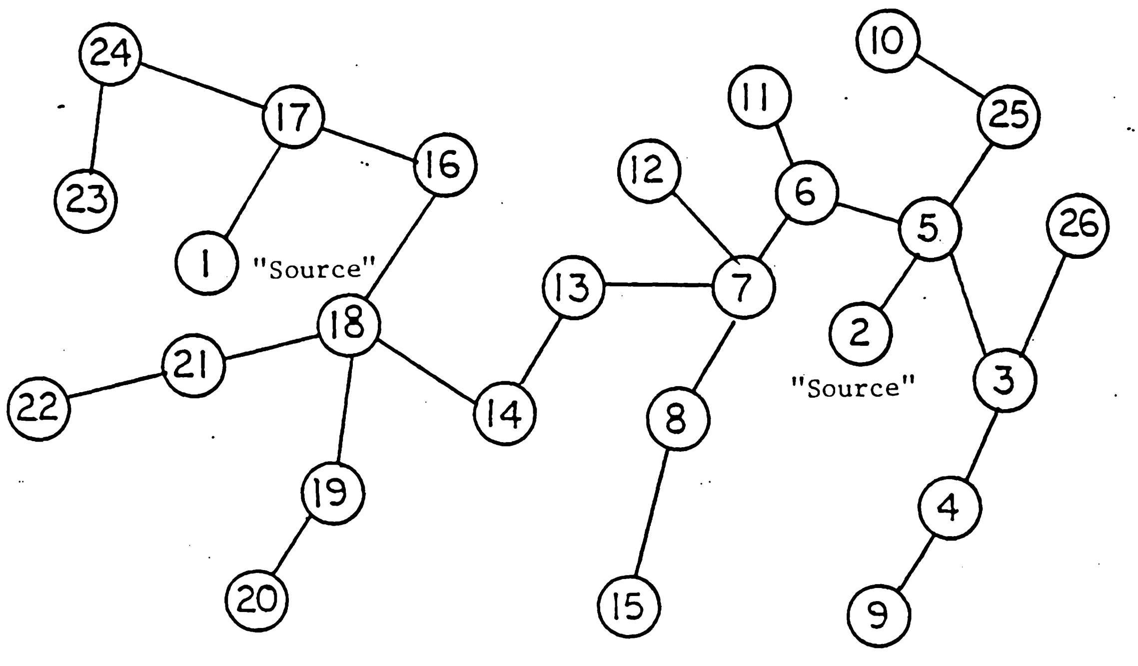


Figure 20. An Intermediate Solution for the Multi-source Network

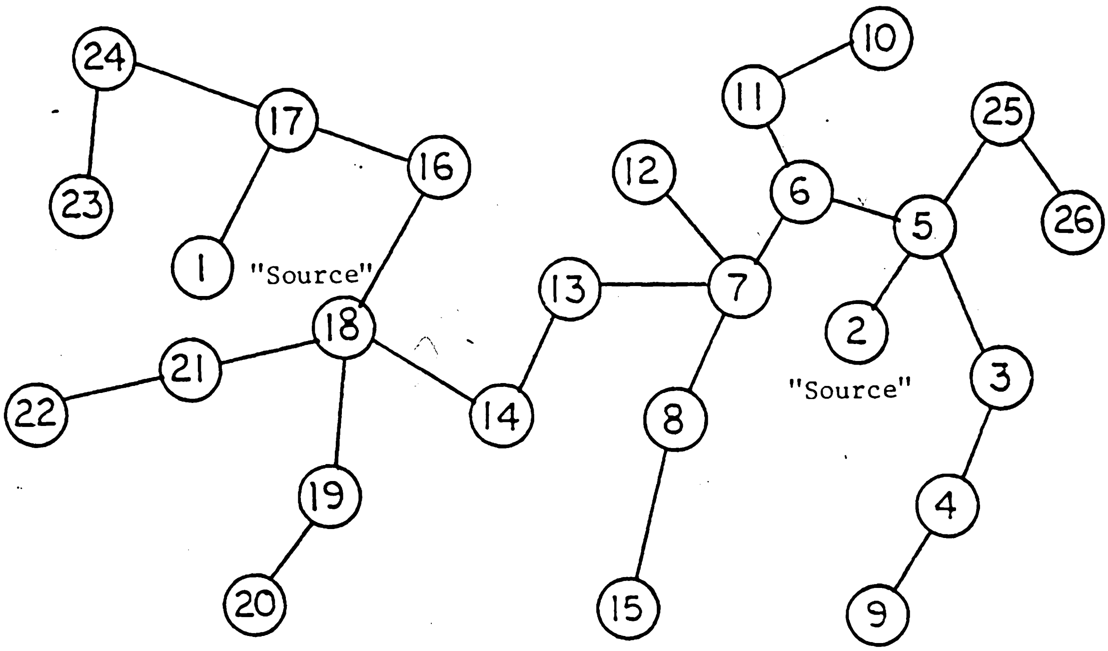


Figure 21. Optimal Tree From TREESEARCH for the Multi-source Network.

Table 16. Optimal Flows and Link Lengths From TREESEARCH for the Multi-Source Network

Link (i,j)	Flow (GPM)	Length (ft)	Diameter (inch)
(3,4)	2190.00	2685.00	8.00
(2,5)	10500.00	60.00	22.00
(3,5)	-2990.00	2400.00	10.00
(5,6)	5770.00	1266.41	18.00
		533.59	20.00
(6,7)	4390.00	1260.00	12.00
(7,8)	1800.00	1289.64	6.00
		405.36	8.00
(4,9)	1040.00	1045.40	6.00
		2434.60	8.00
(6,11)	1090.00	1210.00	6.00
(10,11)	-530.00	1790.00	6.00
(7,12)	1060.00	2210.00	6.00
(7,13)	980.00	1152.83	6.00
		627.17	8.00
(13,14)	530.00	1560.00	6.00
(8,15)	1050.00	2500.00	6.00
(1,17)	4500.00	450.00	16.00
(16,17)	2860.00	1380.00	14.00
(14,18)	0.00	2510.00	6.00
(16,18)	2370.00	2500.00	6.00
(18,19)	1170.00	2750.00	6.00
(19,20)	580.00	1440.00	6.00
(18,21)	960.00	1784.00	6.00
		1056.00	8.00
(21,22)	190.00	2200.00	6.00
(17,24)	1030.00	4710.00	6.00
(23,24)	-810.00	1800.00	6.00
(5,25)	890.00	1535.00	6.00
(25,26)	370.00	1650.00	6.00

Table 17. Optimal Heads From TREESEARCH for the Multi-Source Network

Node (i)	Head (ft)	Node (i)	Head (ft)
1	1421.62	14	1328.09
2	1347.50	15	1290.00
3	1336.16	16	1420.00
4	1315.16	17	1421.16
5	1347.44	18	1328.09
6	1346.00	19	1300.74
7	1341.04	20	1296.84
8	1310.35	21	1314.00
9	1302.00	22	1313.24
10	1331.34	23	1375.10
11	1335.44	24	1384.16
12	1322.73	25	1338.24
13	1331.67	26	1336.29

additional head at source node 1 = 17.617 ft.

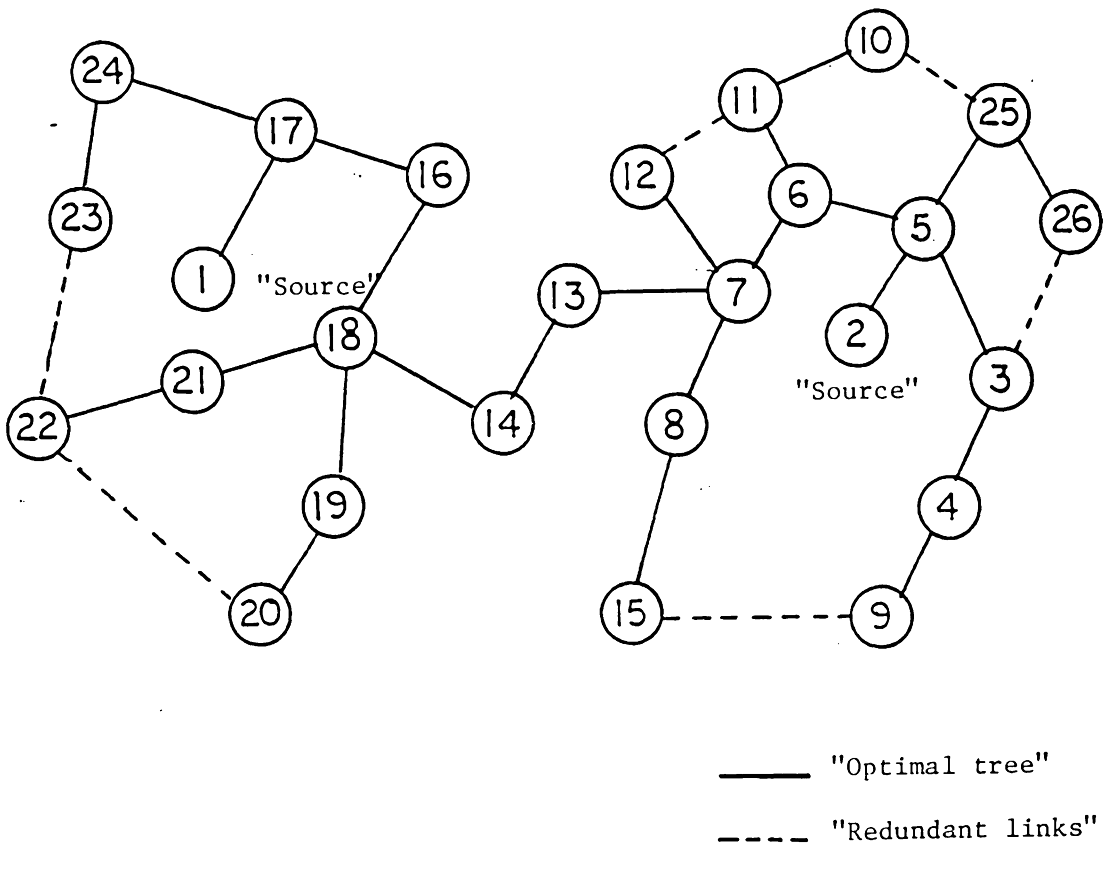


Figure 22. Application of TREESEARCH and REDUNDANCY to the Multi-source Network

Table 18. Optimal Flows and Link Lengths From REDUNDANCY for the Multi-Source Network

Link (i,j)	Length (ft)	Flow (GPM)	Diameter (inch)
(1,17)	450.0	4774.48	16.0
(2,5)	60.0	10225.52	24.0
(5,3)	2400.0	3248.53	10.0
(3,4)	2554.2	2343.26	8.0
	130.8		10.0
(4,9)	3480.0	1193.26	8.0
(5,25)	1535.0	1040.17	6.0
(5,6)	1325.3	5086.82	18.0
	474.7		20.0
(6,7)	1260.0	3661.26	12.0
(6,11)	1210.0	1135.56	6.0
(7,12)	2210.0	759.00	6.0
(7,8)	1289.6	1646.74	6.0
	405.4		8.0
(7,13)	1152.8	705.52	6.0
	627.2		8.0
(8,15)	2500.0	896.74	6.0
(10,11)	1790.0	-274.56	6.0
(13,14)	1560.0	255.52	6.0
(14,18)	2510.0	-274.48	6.0
(16,18)	2500.0	2225.82	6.0
(18,19)	2750.0	857.51	6.0
(16,17)	1380.0	-2715.82	14.0
(17,24)	4710.0	1448.66	6.0
(18,21)	1784.0	853.83	6.0
	1056.0		8.0
(19,20)	1440.0	267.51	6.0
(21,22)	2200.0	83.83	6.0
(23,24)	1800.0	-1228.66	6.0
(25,26)	1650.0	264.73	6.0
(9,15)*	3850.0	153.26	6.0
(3,26)*	1555.0	105.27	6.0
(25,10)*	2490.0	255.44	6.0
(12,11)*	2510.0	-301.00	6.0
(20,22)*	5570.0	-312.49	6.0
(23,22)*	3510.0	418.66	6.0

*Redundant Links

Table 19. Optimal Heads From REDUNDANCY for the Multi-Source Network

Node (i)	Head (ft)	Node (i)	Head (ft)
1	1421.62	14	1336.77
2	1347.50	15	1301.67
3	1334.35	16	1420.06
4	1311.37	17	1421.11
5	1347.46	18	1338.47
6	1346.31	19	1323.13
7	1342.77	20	1322.20
8	1316.82	21	1327.15
9	1302.56	22	1326.99
10	1333.74	23	1332.18
11	1334.95	24	1351.72
12	1332.94	25	1335.22
13	1337.69	26	1334.17

additional head at source node 1 = 17.617 ft.

The vita has been removed from
the scanned document