A Multi-Objective Genetic Algorithm Approach to the Design of Reliable Water Distribution Networks

By T.Devi Prasad Post-Doctoral Research Fellow, Dept. of Civil Engineering., Dong-A Univ., 840, Hadan-Dong, Saha-Gu, Busan, Korea, 604-714. **and Nam-Sik Park** Associate Professor, Dept. of Civil Engineering., Dong-A Univ., 840, Hadan-Dong, Saha-Gu, Busan, Korea, 604-714.

Abstract: The paper presents a multi-objective genetic algorithm approach to the design of a water distribution network. The objectives considered are minimization of network cost and maximization of a reliability measure. In this study, a new reliability measure, called network resilience, is introduced. This measure mimics a designer's desire of providing excess power at nodes and designing reliable loops with practicable pipe diameters. The proposed method produces a set of Pareto-optimal solutions in the search space of cost and network resilience. Genetic algorithms are observed to be poor in handling constraints. To handle constraints in a better way, a constraint handling technique that does not require a penalty coefficient and applicable to water distribution systems is presented. The present model is applied to two example problems, which were widely reported. Pipe failure analysis carried out on some of the solutions obtained revealed that the network resilience based approach gave better results in terms of network reliability.

INTRODUCTION

When a source of water is far off demand points, water has to be transmitted through a network of pipes from source to demand points at specified heads and flow rates. The present day water distribution networks are complex and require huge investments in their construction and maintenance. Due to these reasons a need to improve their efficiency by way of minimizing their cost and maximizing the benefit accrued from them is strongly felt. In the last three decades significant number of methods have been developed using liner programming, dynamic programming, enumeration techniques, heuristic methods, and evolutionary programming. A review of these can be found in Simpson et. al. (1994) and Savic and Walters (1997). Most of these methods consider the minimization of cost of a pipe network, although some reliability studies and stochastic modeling of demands have been attempted. In addition to cost, obviously, there are other possible objectives like reliability, redundancy and/or water quality that can be included in the optimization process.

When more than one objective is present in an objective function, there may not exist one solution, which is best with respect to all objectives. Instead, in a multi-objective optimization problem there exist a set of solutions called Pareto-optimal solutions or nondominated solutions. ThePareto set gives an engineer more flexibility in the selection of a practicable solution. . Todini (2000) presented a heuristic method considering cost function and resilience index, a reliability measure, as objectives. Although this method is a step forward in considering multiple objectives in pipe network optimization, the Pareto-optimal front obtained can at maximum be an approximation of true Pareto-optimal front because of the heuristic method used. Also, resilience index does not avoid loops made of widely different diameters. This paper presents a multi-objective genetic algorithm approach to the design of a water distribution network. The objectives considered are minimization of network cost and maximization of a reliability measure. In this study, a new reliability measure, called network resilience, is introduced. This measure mimics a designer's desire of providing excess power at nodes and designing reliable loops with practicable pipe diameters. The proposed method produces a set of Pareto-optimal solutions in the search space of cost and network resilience. Genetic algorithms are observed to be poor in handling constraints. To handle constraints in a better way, a constraint handling technique that does not require a penalty coefficient and applicable to water distribution systems is presented.

FORMULATION OF THE MODEL

The following is the proposed two-objective optimization model for a water distribution network design. The two objective functions are: (i) minimization of network cost, and (ii) maximization of a reliability measure.

Minimize $$f_1(D_i)=\sum_{i=1}^{np}C(D_i,L_i) \quad (1)$$

Maximize $$f_2=I_n \quad (2)$$

where $C(D_i,L_i)$ = cost of the pipe i with diameter D_i and length L_i np = number of pipes in the system; and I_n = network resilience. The above optimization model is subjected to the following constraints:

$$g_j(H,D)=0 \qquad j = 1,2,,nn \quad (3)$$

$$H_j \geq H_j^l \qquad j = 1,2,,nn \quad (4)$$

$$D_i \in \{A\} \qquad i = 1,2,,np \quad (5)$$

where nn = number of junction nodes; $g(H,D)$ = nodal flow continuity equations: H_j = head at any node j, which must be greater than a minimum specified value H_j^l; and all D_i's are discrete pipe sizes selected from a set of commercially available sizes.

The above-formulated model is a multi-objective mixed integer nonlinear optimization model. It can be solved using a multi-objective genetic algorithm. In this study, network hydraulic analysis is performed using the method developed by Gupta and Prasad (2000), which require a less number of iterations with any initial guess of pipe flows. For calculating pipe head losses, Hazen-Williams equation in the following form is used.

$$h_f=\frac{\omega L q^{1.852}}{C_H^{1.852} D^{4.87}} \quad (6)$$

where = a numerical constant, which depends on the units used; and C_H = Hazen-Williams coefficient. In this study =10.5088 (SI units) is used.

MULTI-OBJECTIVE GENETIC ALGORITHMS

In dealing with multicriterion optimization problems, classical search and optimization methods are not efficient. simply because (i) most of them cannot find multiple solutions in a single run, thereby requiring them to be applied as many times as the number of desired Paretooptimal solutions, (ii) multiple application of these methods do not guarantee finding widely different Paretooptimal solutions, and (iii) most of them cannot efficiently handle problems with discrete variables and problems having multiple optimal solutions. On the contrary, the studies on evolutionary search algorithms, over the past few years, have shown that these methods can be efficiently used to eliminate most of the above difficulties of classical methods (Deb 2001). Since they use a population of solutions in their search, multiple Paretooptimal solutions can, in principle, be found in one single run. The use of diversity preserving mechanisms can be added to the evolutionary search algorithms to find widely different Paretooptimal solutions. In this study, a multi-objective genetic algorithm, called nondominated sorting genetic algorithm, is used.

Nondominated Sorting Genetic Algorithm (NSGA)

The idea behind NSGA is that a ranking method is used to emphasize current nondominated points and a niching method is used to maintain diversity in the population. Before the selection is performed, the population is first ranked on the basis of an individual's nondomination level, which is found by the following procedure, and then fitness is assigned to each population member.

For a problem having more than one objective function, any two solutions $x^{(1)}$ and $x^{(2)}$ can have one of two possibilities, one dominates the other or none dominates the other. A solution $x^{(1)}$ is said to dominate the other solution $x^{(2)}$, if both the following conditions are true:

1. The solution $x^{(1)}$ is no worse (say the operator denotes worse and denotes better) than $x^{(2)}$ in all objectives, or $f_j(x^{(1)}) \quad f_j(x^{(2)})$ for all j = 1,2,, M, objectives.
2. The solution $x^{(1)}$ is strictly better than $x^{(2)}$ in at least one objective, or $f_j(x^{(1)}) \quad f_j(x^{(2)})$ for at least one j {1,2,,M}.

If any of the above conditions is violated, the solution $x^{(1)}$ does not dominate the solution $x^{(2)}$. If $x^{(1)}$ dominates the solution $x^{(2)}$, then $x^{(1)}$ is said to be nondominated solution. In this study, real coded NSGA with tournament selection, arithmetic crossover, and Gaussian mutation are used. For more information about NSGA the reader may refer Srinivas and Deb (1994)

Arithmetic crossover: If we assume $x^{(1)} = (x_1^1, x_2^1, \dots, x_{nd}^1)$ and $x^{(2)} = (x_1^2, x_2^2, \dots, x_{nd}^2)$ are two parents (solutions) selected for crossover, then two offspring are generated as follows:

$$y^{(k)} = (y_1^k, y_2^k, \dots, y_{nd}^k) \qquad k=1,2 \qquad (7)$$

where $y_i^1 = \lambda x_i^1 + (1-\lambda)x_i^2$; $y_i^2 = (1-\lambda)x_i^1 + \lambda x_i^2$; and is a constant (0 1). This crossover operator was found to give good results for water distribution networks optimization with = 0.75 (Vairavamoorthy and Ali 2000). The same is used in this study also.

Gaussian mutation: If y(k) is an offspring and y_i^k is a gene randomly selected for mutation, then the gene obtained after Gaussian mutation is as follows:

$$z_i^k = y_i^k + N(0,\sigma) \qquad (8)$$

where $N(0,\sigma)$ is a random Gaussian number with mean zero and standard deviation $\sigma = f(y_i^u)$, where y_i^u is the maximum value of the gene. Here the value of $\sigma = 0.1 \times y_i^u$ is used. With this scheme applied, if new gene values exceed their range at either end, the values are adjusted to take the limiting values.

Constraint handling: In the previous GA applications to water distribution network optimization many improvements were suggested for constraint handling (Dandy et.al. 1996; Savic and Walters 1997; and Vairavamoorthy and Ali 2000). However these methods are not elegant in the sense that they all require a penalty co-efficient. Identifying a penalty co-efficient is a difficult task and it may change from problem to problem. The penalty co-efficient must take a value that will not allow the best infeasible solution to be better than any feasible solution in the population (Simpson et al., 1994; Savic and Walters 1997). In this study a method of constraint handling, which does not require a penalty co-efficient to be specified and applicable to water distribution network is developed.

A solution $x^{(i)}$ is constraint-dominating a solution $x^{(j)}$, if any of the following are true:

1. solution $x^{(i)}$ is feasible and solution $x^{(j)}$ is infeasible.
2. solution $x^{(i)}$ and $x^{(j)}$ both are infeasible, but $x^{(i)}$ has a smaller constraint violation.
3. solution $x^{(i)}$ and $x^{(j)}$ are feasible and solution i dominates solution j

This way, feasible solutions are constraint-dominated to any infeasible solution and two infeasible solutions are compared based on their constraint violations only. However, when two feasible solutions are compared, they are checked on their domination level (fitness value). The constraint violation for any solution can be calculated using failure index as,

$$I_f = \frac{\sum_{j=1}^{nn} e_j}{\sum_{k=1}^{nr} Q_k H_k + \sum_{i=1}^{npu} (P_i/\gamma)} \quad \text{where,} \quad e_j = \begin{cases} 0 & \text{when } H_j \geq H_j^l \\ Q_j\left(H_j^l - H_j\right) & \text{otherwise} \end{cases} \tag{9}$$

The above constraint handling procedure does not require any penalty co-efficient and always a feasible solution has more priority than any infeasible solution.

Reliability Measures

A branched water distribution network will have severe consequences in terms of reliability under failure conditions. In order to reduce the risk of failure and improve the reliability of a water distribution network, often designers introduce redundancy into networks by adding pipes to close the loops. This causes the flow to reach demand points in alternative paths under failure conditions. Least cost design of pipe networks have resulted some of the pipes having minimum specified diameter and heads at some of the nodes barely satisfied (Savic and Walters 1997). Whenever there is a mechanical or hydraulic failure, the internal head losses will increase causing failure of the network. This increased head losses during failure conditions can be met, if sufficient excess power is available for internal dissipation. Based on this premise, the following reliability measures are defined.

Resilience Index (I_r): Todini (2000), proposed the resilience index based on the concept that the power input into the network is equal to the power lost internally to over come the friction plus the power that is delivered at demand points. The resilience index for the entire network is defined as

$$I_r = 1 - \left(\frac{P_{\text{int}}}{P_{\text{int}}^{\max}}\right) = \frac{\sum_{j=1}^{nn} Q_j\left(H_j - H_j^l\right)}{\left(\sum_{k=1}^{nr} Q_k H_k + \sum_{i=1}^{npu} P_i/\gamma\right) - \sum_{j=1}^{nn} Q_j H_j^l} \tag{10}$$

where P_{int}, is the amount of power dissipated in the network; and $P_{\text{int}}^{\max}$, is the maximum power that would be dissipated internally in order to satisfy design demand Q and design head H^l at junction nodes.

Maximization of resilience index improves the ability of a pipe network to counter the failure conditions.

Network Resilience (I_n): Maximization of resilience index may improve output power at junction nodes. However, it does not avoid loops having widely different diameters. Increase in the following reliability measure, called network resilience (I_n), not only tries to improve nodal surplus power but also uniformity in diameters of pipes connected to it. The surplus power at any node j is given by,

$$P_j = \gamma Q_j \left(H_j - H_j^l\right)$$

Reliable loops can be ensured, if the pipes connected to a node are not widely varying in diameter. If D1, D2 and D3 (where, D1 D2 D3) are the diameters of three pipes connected to node j then, uniformity of that node is given by,

$$C_j = \frac{(D_1 + D_2 + D_3)}{3D_1} \quad \text{or} \quad C_j = \frac{\sum_{i=1}^{npj} D_i}{np_j \times \max\{D_i\}} \tag{11}$$

where npj = number of pipes connected to node j. The value of C = 1, if all pipes connected to a node are having same diameter and C < 1, if the pipes connected to a node are having different diameters. For nodes connected with only one pipe, the value of C is taken to be one. The network resilience Inis then defined as the total resilience of junction nodes of a network per unit input power and is expressed as

$$I_n = \frac{\sum_{j=1}^{nn} C_j P_j}{P_{tot}} = \frac{\sum_{j=1}^{nn} C_j Q_j \left(H_j - H_j^l\right)}{\sum_{k=1}^{nr} Q_k H_k + \sum_{i=1}^{npu} (P_i/\gamma)} \tag{12}$$

The network resilience can also be viewed as equivalent to resilience index with each node j is given a weight of Cj based on the uniformity in diameter of pipes connected to it.

APPLICATION OF THE MODEL

Example 1: Comparison of the present method with that of Todini's is made on the basis of: (1) efficiency of optimization algorithm and (2) comparison of reliability measures. The example network shown in Fig.1, first used by Alperovits and Shamir (1977) and later on by many investigators, is chosen for this purpose. This network is a typical network as it contains many alternative solutions with the same network cost. The network is a two-loop network with 7 nodes and 8 pipes, each having a length of 1000m. Pipe cost data, and node data are given in Table 1 and Table 2.respectively. There are eight decision variables in this example. The model is first run with the objectives of minimizing cost and maximizingresilience index. The GA parameters used are: population size = 100; probability of crossover = 1.0; probability of mutation = 0.125 (approximately 1/P: P = number of variables); number of generations = 1000 and σ_{share} = 0.25. The designs, for which network cost is less than \$500,000, are presented in Table 3. Also, the designs reported by Todini (2000) are presented in this table. Comparison of the solutions obtained for a network cost of \$450,000, which has different network configurations, reveals the superiority of the proposed optimization algorithm as it gave a solution with higher resilience index value. Also, the solutions are uniformly distributed in this range. This can be easily understood as NSGA was observed to converge towards Pareto-optimal front in many test problems Deb (1999), compared to heuristic method used by Todini. Another advantage of NSGA is that it can give many solutions along the Pareto front (Fig.2). Although resilience index based approach has improved the surplus power at nodes, it could not eliminate impracticable loops. This aspect can be realized from the designs given in Table 3. In order to compare the reliability measures, two-loop network (Fig.1) is now solved with the objectives of minimizing the cost and maximizing network resilience. The GA parameters used are the same as before. The Pareto-optimal front is shown in Fig.3. Some of the designs obtained using this approach are given in table 4. These designs are not only having increased surplus power at nodes, but also having loops with practicable diameters.

Example 2: The present method is also applied to another benchmark problem, Hanoi water distribution network, reported by many investigators. The network layout is taken from Fujiwara and Khang (1990). This network consists of 32 nodes and 34 pipes, and is supplied by a fixed grade source at an elevation of 100m. The minimum required head at all junction nodes is specified to be 30m. The set of commercially available diameters (in inches) is D = [12, 16, 20, 24, 30, 40] and their corresponding cost per unit length is calculated using the equation $1.1D^{1.5}$. In this problem, the number of decision variables is 34 and the GA parameters used are: population size = 200; probability of crossover = 1.0; probability of mutation = 0.03; number of generations = 10000; and σ_{share} = 0.4. This method further substantiates the results obtained for two-loop network. The Pareto optimal front is as shown in Fig.4. The engineer can now use these various solutions for further analysis to select a suitable design.

CONCLUSIONS

Most of the water distribution network optimization models have considered network cost as a sole objective. This paper describes the application of a multi-objective genetic algorithm model to the design of a water distribution network. A better constraint handling technique that does not require a penalty coefficient and applicable to water distribution networks is presented. The objectives considered in this study are: minimization of network cost and maximization of a reliability measure. The reliability measure used is network resilience is a measure of both the nodal surplus power and the uniformity in diameters connected to that node. Increase in the value of network resilience improves the reliability of a network under failure conditions. Application of the model to the example problems revealed that the network resilience based approach gave better results in terms of pipe failure reliability than resilience index based approach.

REFERENCES

1. Alperovits, E., and Shamir, U. (1977). "Design of optimal water distribution system." Water Resour. Res., 13(6), 885-900.
2. Dandy, G.C., Simpson, A.R. and Murphy, L.J. (1996). "An improved genetic algorithm for pipe network optimization, Water Resour. Res., 32(2), 449-458.
3. Deb, K. (1999). "Multi-objective genetic algorithms: Problem difficulties and construction of test problems."J. Evolutionary Computation, 7(3), 205-230.
4. Deb, K. (2001). Multi-Objective Optimization using Evolutionary Algorithms. John Wiley & Sons Ltd. London.
5. Fujiwara, O., and Khang, D.B. (1990). "A two-phase decomposition method for optimal design of looped water distribution networks." Water Resour. Res., 23(6), 977-982.
6. Gupta, R., and Prasad, T.D. (2000). "Extended use of linear graph theory for analysis of pipe networks." J. Hydr. Engrg., ASCE, 126(1), 56-62.
7. Savic, D., and Walters, G. (1997). "Genetic algorithms for least cost design of water distribution network." J. Water Resour. Plng. and Mgmt., ASCE, 123(2), 67-77.
8. Simpson, A.R., Dandy, G.C., and Murphy, L.J. (1994). "Genetic algorithms compared to other

techniques for pipe optimization." J. Water Resour. Plng. and Mgmt., ASCE, 120(4), 423-443.

9. Srinivas, N., and Deb, K. (1994). Multi-objective function optimization using non-dominated sorting genetic algorithms." Evolutionary Computation, 2(3), 221-248.

10. Todini, E. (2000). "Looped water distribution networks design using a resilience index based heuristic approach." Urban Water, Elsevier Pubs., 2, 115-122.

11. Vairavamoorthy, K., and Ali, M. (2000). "Optimal design of water distribution systems using genetic algorithms." Comp. Aided Civ. And Infstr. Engrg., 15(2), 374-382.

12. Walski, T.M. (2001). Editorial note: The wrong paradigm-why water distribution optimization doesn't work. J. Water Resour. Plng. and Mgmt., ASCE, 127(4), 203-205.

Table 1: Pipe size and cost data for two-loop network

Dia (in)	1	2	3	4	6	8	10	12	14	16	18	20	22	24
Cost ($/m)	2	5	8	11	16	23	32	50	60	90	130	170	300	550

Table 2: Node data for two-loop network

Node	1	2	3	4	5	6	7
Head(m)	210	180	190	185	180	195	190
Demand (m3/h)	-1120	100	100	120	270	330	200

Table 3: Results of two-loop network using Resilience Index

Pipe No.	Design obtained using NSGA(Diameter in inches)					Design obtained using Todini's Method (Diameter in inches)				
	1	2	3	4	5	1	2	3	4	5
1	18	18	20	20	20	18	18	20	20	20
2	10	14	14	14	16	10	16	14	14	14
3	16	14	14	14	14	16	14	14	14	14
4	4	8	6	8	2	4	6	6	8	6
5	16	14	12	14	14	16	14	14	14	14
6	10	2	1	1	1	10	1	1	1	1
7	10	14	14	14	14	10	14	14	14	14
8	1	10	10	10	10	1	10	10	10	12
cost($)	419000	430000	450000	467000	479000	419000	450000	460000	467000	468000
I_r	0.2229	0.3612	0.4333	0.4796	0.5170	0.2229	0.4054	0.4681	0.4796	0.4905
Note:ω=10.5088 is used in Hazen-Williams equation for both the cases										

Table 4: Results of two-loop network using NSGA with Network Resilience

Pipe No.	Solution 1	Solution 2	Solution 3	Solution 4
1	18	18	18	18
2	14	14	14	14
3	16	16	16	16
4	6	10	10	10
5	14	14	14	14
6	8	8	6	8
7	10	10	12	14
8	10	10	10	10
Network cost($)	443,000	459,000	470,000	487,000
Network Resilience(I_n)	0.0291	0.0381	0.0393	0.0412
Resilience Index(I_r)	0.3227	0.3879	0.4239	0.4539
Total Surp. Head(m)	58.96	65.87	68.94	72.12
Min. Surp. Head(m)	0.0234	0.1006	1.29	1.37

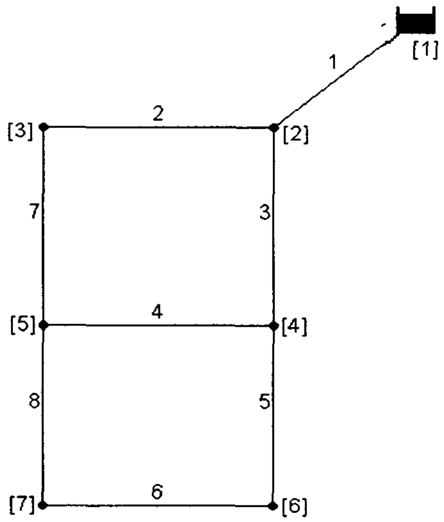

Fig. 1 Two-loop network

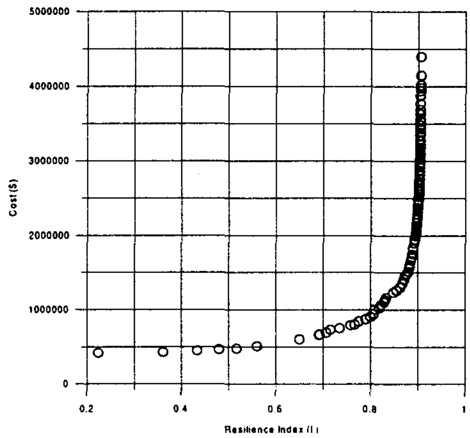

Fig.2 Pareto Front in Cost and Resilience Index Space for two-loop network

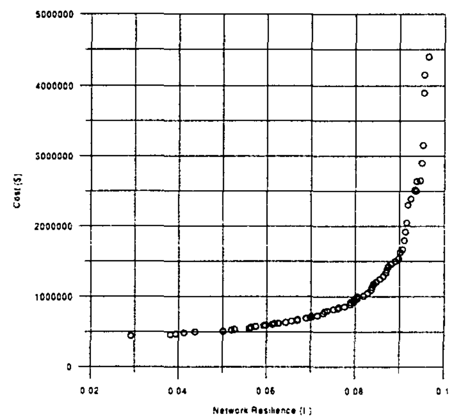

Fig.3 Pareto Front in Cost and Network Resilience space for two-loop network

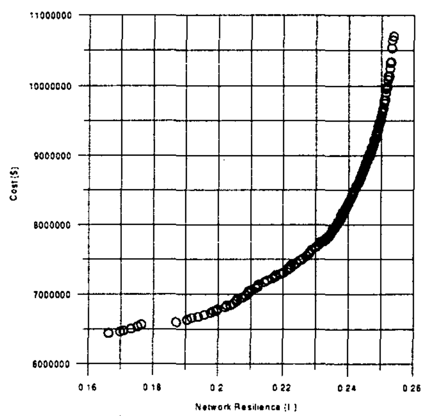

Fig.5 Pareto Front in Cost and Network Resilience space for Hanio Network