OPTIMIZATION OF THE WATER DISTRIBUTION NETWORKS WITH SEARCH SPACE REDUCTION

Milan Čistý, Zbyněk Bajtek

Slovak University of Technology Bratislava, Faculty of the Civil Engineering

Abstract

A water distribution network is a system containing pipes, reservoirs, pumps, and valves of different types, which are connected to each other to provide water to consumers. In the case of the design of a pipe network the optimization problem can be stated as follows: minimize the cost of the network components subject to the satisfactory performance of the water distribution system (mainly, the satisfaction of the allowable pressures in demand nodes). This leads to difficult constrained combinatorial optimization problem. Various algorithms ranging from artificial intelligence to the optimization domain have been applied. In this paper modified, two step GA methodology was used to solve this problem.

1 Introduction

A water distribution network is a system containing pipes, reservoirs, pumps, and valves of different types, which are connected to each other to provide water to consumers. In the case of the design of a pipe network the optimization problem can be stated as follows: minimize the cost of the network components subject to the satisfactory performance of the water distribution system (mainly, the satisfaction of the allowable pressures). Various algorithms ranging from artificial intelligence to the optimization domain have been applied. Alperovits and Shamir [1] presented a linear programming gradient for optimizing a water distribution network. Later, Fujiwara and Khang [7] extended the Alperovits and Shamir method to non-linear modeling. Also, Eiger, et al. [6] used the same formulation as Kessler and Shamir [10], which leads to a determination of the lengths of one or more segments in each link with discrete diameters. Researchers have focused on stochastic or so-called heuristic optimization methods since the early 1990s. Simpson, et al. [13] used a simple genetic algorithm in which each individual solution from the population of solutions is represented by a string of bits with identical lengths. Other heuristic techniques have also been applied to the optimization of a water distribution system, such as simulated annealing [3], an ant colony optimization algorithm [12], a shuffled frog leaping algorithm [5] and a harmony search [9], to name a few.

The reason for this work is that significant differences from the known global optimums are referred to even for single objective tasks and simple benchmark networks, while existing algorithms are applied. Reca, et al [11] evaluated the performance of several meta-heuristic techniques - genetic algorithms, simulated annealing, tabu search, and iterated local search. He compared (among the other testing accomplished) these techniques by applying them to medium-sized benchmark networks. The results which he obtained for the Hanoi network, which is well-known and often used in the optimization community (after ten different runs with five heuristic search techniques), varied from 6,173,421 to 6,352,526. These results differ by 1.5 – 4.5 % from the known global optimum for this task, which is a relatively large deviation for such a small network. Similar results were presented by Zecchin [15] in a comparative study of ant colony optimization algorithms in which other heuristic algorithms were also tested. It could be expected that this difference would be even greater for larger networks (e.g., more than 6% for the network tested in [2]. The main concern of this paper is to propose a method which is more dependable and converges more closely to a global optimum than existing algorithms.

It is known that the GA search is influenced by various parameters such as population size, coding scheme, penalty method, fitness function, selection and crossover operators, probability of crossover, probability of mutation , and hydraulic simulation technique. Parameters such as the size of a network i.e., the number of links and the number of commercially available pipes mainly influence the size of the GA search space. This parameter is maybe most important from the mentioned and that is why in present work we focused on its reduction by proposed iteration methodology.

2 Metodology

In order to overcome deficiencies of the mathematical programming techniques, heuristic optimization techniques have been introduced. First of them was applied genetic algorithms methodology, which is used also in this study. It is search procedure inspired by the mechanics of natural genetics and natural selection. This methodology is finding increased application in solving difficult problems of engineering, science, and commerce. Its basic ideas are briefly summarized below; a good introduction to the subject is given by Goldberg [8].

The first step is to represent a legal solution of the problem by a string of genes (seeking parameters) that can take on some value from a specified finite range. This string of genes, which represents the solution, is known as a chromosome. Then an initial population of legal chromosomes is constructed at random. Genetic algorithms are implemented as a computer simulation in which a population of chromosomes evolves toward better solutions by means of genetic operators such as inheritance, mutation, selection, or crossover. With each generation, the fitness of each chromosome in the population is measured. The fitter chromosomes are then selected to produce offspring for the next generation, which inherit the best characteristics of both parents. This process is repeated until some form of convergence in fitness is achieved. The goal of the optimization process is to minimize or maximize the fitness.

The NSGA was implemented by Srinivas and Deb as the method for solving the multi-objective problem. But the high computational complexity of nondominated sorting, lack of elitism and sharing parameter δ share have been criticized for years. Deb [4] presented NSGA-II as an improvement of NSGA by introducing a fast non dominated sorting procedure with less computational complexity, an elitist-preserving mechanism and a parameterless niching operator for diversity preservation. NSGA-II also performs well for solving the constrained multiobjective optimization problems. A non-dominated sorting genetic algorithm is a multi-objective genetic algorithm which is designed to solve optimization problems with multiple objectives and multiple constraints. In addition to a single population, it uses a merged population composed of parent population and offspring population. The non-dominated sorting is performed on the merged population and separates them into different non-dominated level sets. The next generation population is chosen based on ConstrainDominate selection operator and crowding distance from one solution to surrounding solutions. Additionally, the Hypervolume for the Pareto-optimal front is monitored to determine the convergence of the evolution.

In the case of the design of a pipe network the optimization problem can be stated as follows: minimize the cost of the network components subject to the satisfactory performance of the water distribution system. If we reduce the problem to designing only new pipes, the chromosome can be an integer string of numbers (genes) representing the possible diameters in each section. An efficient and effective search for the optimum design solution of a water distribution network using genetic algorithms is governed by several factors such as a representation scheme, population size, hydraulic simulation model, fitness function, penalty method, GA operators, number of generations and, more importantly, the size of the search space.

A well-known optimization problem from the literature has been chosen. The Hanoi WDS has attracted many researchers since it was first introduced by Fujiwara and Khang [7]. However, it has been mainly tackled as a single objective problem where the sole objective was to minimise the cost of the solution subject to constraint of a zero head deficit at all nodes. The optimization of the Hanoi WDS is a network design problem in which 34 pipes may take one of six pipe diameter options giving a search space size of is 634 = 2.86512E+26. A global minimum pressure constraint of 30 metres applies to the optimization. The set of commercially available diameters is 304.8, 406, 508, 609.6, 762 and 1016 millimeters. The topology of the network is shown in Figure 1.

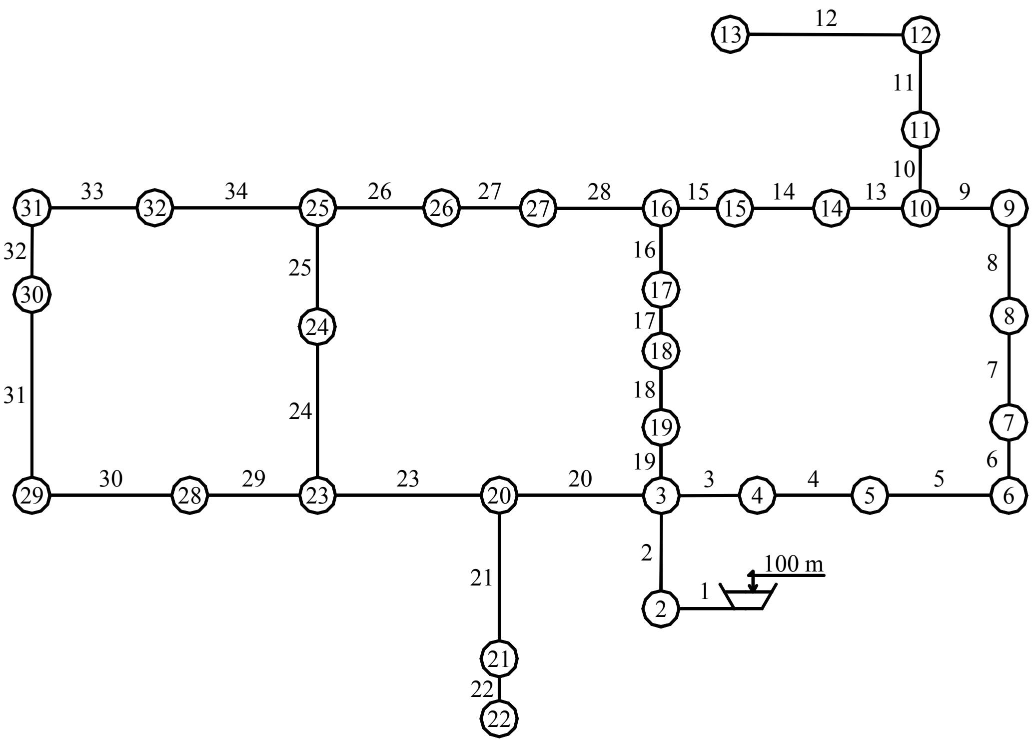

Figure 1: Hanoi network

In this case study, the problem has been reformulated as a MO optimization problem with two objectives defined as follows:

Minimise $$C = \sum_{i=1}^{Np} 1.1.D_i^{1.5}.L_i \qquad (1)$$

Minimise $$Hd = \sum_{j=1}^{N_N} \max(H_{\min} - H_j;0) \qquad (2)$$

where: C is cost in US dollars,

D_i is diameter of ith pipe in inches,

L_i is length of ith pipe in metres,

N_P is number of pipes in the network (i.e. 34 pipes),

H_d is total head deficit,

H_{min} is minimum required head at a node (i.e. 30m),

H_j is available head at jth node,

N_N is number of nodes in the network (i.e. 32 nodes) and subject to hydraulic model constraints (i.e. mass and energy conservation equations).

This definition (MO) was chosen because when it is not necessary to use penalties as by the definition of SO. SO problem can be formulated similarly to the MO version, however the sole objective was to minimise the cost of the network as defined in Equation 1. In order to obtain feasible results with a zero head deficit, a penalty function as defined below:

$$Penalty = \text{penaltya} * H_d \wedge 2 + \text{penaltyb} \qquad (3)$$

was applied to all solutions having a non-zero head deficit.

The calculation was carried out in two steps, calculation in first step been carry with entire gene (which means that the GA have available all 6 diameters) for each section (pipe -34). At the moment when it was reached approximate prize for this problem the best solution was chose as offspring for the reduction of search space. Based on the flow from this network and known minimum [0.1m/s] and maximum [3.0 m/s] required flow speed (These are dependent on the hydraulic network requirements) and a known flow rate was subsequently determined the minimum and maximum diameter of the pipe. These define the range of diameters i.e. size of a single gene. In this way we are able to achieved a significant reduction in search space(in same cases only one choice remain).

The source code is an extension of the GA toolbox implementation in C++ (Sastry,[4]) so that it can be used inside matlab with fitness functions written in matlab. The code is available from ftp://ftp-illigal.ge.uiuc.edu/pub/src/GA/Gatoolbox matlab.

To run the GA toolbox, at the command prompt, following command should be used:

GAtbxm('<input file name>')

For example to use **input_sga_maxSpec** as the input file type **GAtbxm('input_sga_maxSpec')**.

The toolbox solve both single- and multi-objective problems with or without constraints. The multiobjective genetic algorithm is the non-dominated sorting GA II (NSGA-II) [4]. The decision variables of the problem are encoded as real numbers or as integers within their specified ranges. This encoding procedure permits the decision variables to be binary, discrete, or a real number. The decision variable type needs to be specified in the input file and the valid options are **double** or **int.**

In the input file we also sets the parameters for:

- Selection
- Recombination
- Mutation
- Constraint handling
- Local search
- Elitist Replacement

The code for the objective function is the matlab function **sgaFitnessFunction.m** (for SO) and **mogaFitnessFunction.m** (for MO) and it is called within C++ library in the file **userDefinables.cpp.** These are the only files that has to be rewrite in order to try own fitness function. The function header which calls the matlab fitness function is as follows:

void globalEvaluate(double *x, double *objArray, double *constraintViolation, double *penalty, int *noOfViolations)

globalEvaluate takes as argument an array **x,** whose l elements contains the decision variables of a candidate solution whose fitness is being evaluated. Were, l is the problem length (# of genes).

The decision variables are then passed on to the matlab function **sgaFitnessFunction.m** (or **mogaFitnessFunction.m)** which evaluates the objective and constraint violation values of the given candidate solution. The function header is as follows:

function objConst = sgaFitnessFunction(decVars)

sgaFitnessFunction (alternatively **mogaFitnessFunction)** takes as argument an array, **decVars,** containing decision variables of the candidate solution. **sgaFitnessFunction** (alternatively **mogaFitnessFunction)** returns objective function value(s) and constraint violation value(s) in the array **objConst.**

```matlab

function objConst = mogaFitnessFunction(decVars)
x = decVars;  %v x je chromozom

global   minpresure nnodes npipes cenik minimum iteracia;
ele = -1;
iteracia = iteracia+1;
sumcost = 0;              %nulovanie celkovej ceny a
penaltyp = 0;             %nulovanie sumy nedostatku tlaku
for j=1:npipes
    h = x(j);
    err=calllib('epanet2', 'ENsetlinkvalue', j, 0, cenik(h,1));
    %nastavujem priemer potrubia
    err=calllib('epanet2', 'ENsetlinkvalue', j, 2, cenik(h,3));
    %nastavujem drsnosť potrubia
    [err, ele]=calllib('epanet2', 'ENgetlinkvalue', j, 1, ele);
    %citam dlzku useku do ele
    sumcost = sumcost + (ele * cenik(h,2));
    %nasobim dlzku jednotkovou cenou a pripocitavam k celkovej cene
end

err = calllib ('epanet2','ENinitH',1);
err = calllib ('epanet2','ENrunH',1);

for j= 1:nnodes
    [err, ele]=calllib('epanet2', 'ENgetnodetype', j, ele);
    if ele ==0
        [err, ele]=calllib('epanet2', 'ENgetnodevalue', j, 11, ele);      %TLAK v uzle j
        if ele < minpresure
            penaltyp = penaltyp + (minpresure-ele);   %penaltyp je suma nedostatkov tlaku
        end
    end
end

if penaltyp == 0 && minimum > sumcost
    err = calllib('epanet2','ENsaveinpfile',[num2str(sumcost) '.inp']);
    minimum = sumcost;
end

objConst(1)= sumcost;
objConst(2)= penaltyp;

```

Figure 2: Fitness function

TABLE 1: GA PARAMETERS USED IN THE CALCULATION

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Generation	1000	1000	1000	1000	1000	1000	1000	1000	1000	750	750	750	750	750	750
Population	150	150	150	150	150	150	150	150	150	200	200	200	200	200	200
Crossover	0.85	0.85	0.85	0.9	0.9	0.9	0.95	0.95	0.95	0.85	0.85	0.85	0.9	0.9	0.9
Mutation	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15

	16	17	18	19	20	21	22	23	24	25	26	27
Generation	750	750	750	600	600	600	600	600	600	600	600	600
Population	200	200	200	250	250	250	250	250	250	250	250	250
Crossover	0.95	0.95	0.95	0.85	0.85	0.85	0.9	0.9	0.9	0.95	0.95	0.95
Mutation	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15

3 Results

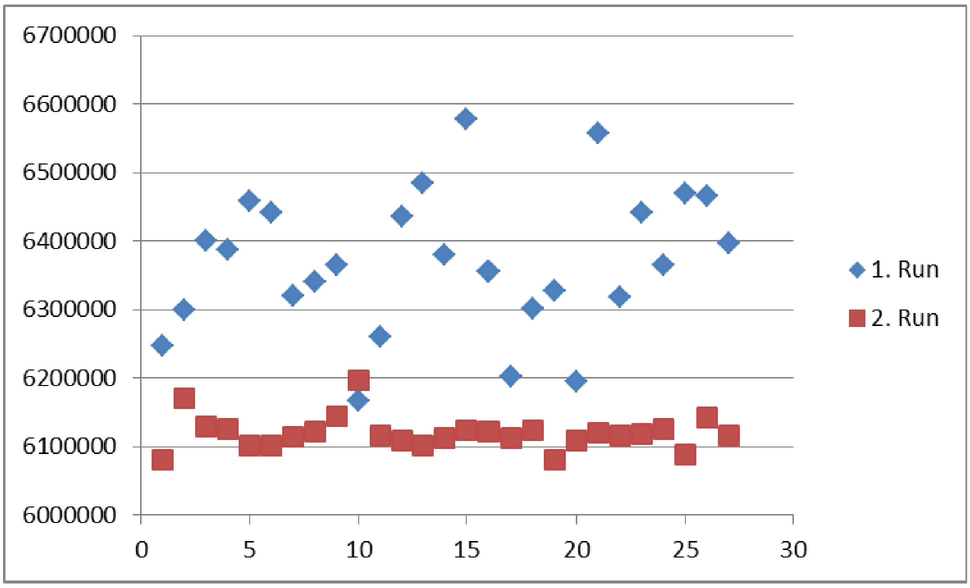

Figure 3: 1-th and 2 – nd Run

TABLE 2: RESULTS FROM 1-TH AND 2-ND RUN

1 Run	6248032	6298947	6400408	6386439	6457854	6440916	6320680	6339919	6365208
2. Run	6081119	6170111	6130362	6125579	6100900	6100900	6113817	6122220	6143952
1 Run	6167735	6260973	6435734	6484699	6577832	6354997	6201559	6301415	6326978
2. Run	6197071	6117148	6109539	6100900	6124159	6121212	6113237	6124283	6081119
1 Run	6194194	6556786	6317965	6441727	6364375	6469348	6466073	6395876	6395876
2. Run	6108686	6119440	6116572	6118621	6126557	6088521	6142475	6117150	6117150

Proposed methodology was tested on benchmark water distribution networks, used in water distribution system optimization community. Results clearly show better performance of the proposed methodology over traditional GA approach. As software tool, GA Toolbox from Illinois GA laboratory was used in this work, which works in MATLAB programming environment.

In this work only two step reduction of search space was used according to the size of this problem. It gives good results, but there is a possibility open to improve replace it with some of the other and more effective heuristic methods which are available in the optimization community.

The author expects that this can even refine the method in the future. The effect of such a refinement will mainly be revealed when significantly larger networks than those tested here will be solved.

Acknowledgement

This study was also produced with the support of the OP Research and Development Center of Excellence project for flood protection area ITMS 26240120004 co-financed by the European Regional Development Fund and by the Slovak Research and Development Agency, Grant No. LPP-0319-09.

References

[1] Alperovits, E., and Shamir U., (1977), Design of optimal water distribution systems. Water Resource Research 13(6), 1977, 885-900.

[2] J. Bicik, M.S. Morley and D.A. Savić (2008), A rapid optimization prototyping tool for spreadsheet-based models. Proceedings of the 10th Annual Water Distribution Systems Analysis Conference WDSA2008, Van Zyl, J.E., Ilemobade, A.A., Jacobs, H.E. (eds.), August 17-20, 2008, Kruger National Park, South Africa.
[3] Cunha, M. C. and Sousa, J., (2001): Hydraulic infrastructures design using simulated annealing. Journal of Infrastructure Systems, ASCE, 7(1), 32-39.
[4] Deb, K., Ptatap, A.,Agarwal, S. Meyarivan, T. : A Fast and Elitist Multiobjective Genetic Algorithm NSGA-II. IEEE Transactions on Evolutionary Computation. Vol. 6. (2002) 182–197
[5] Eusuff MM and Lansey KE (2003) Optimization of water distribution network design using the shuffled frog leaping algorithm. J Wat Resour Plan and Manage ASCE 129(3): 210-225
[6] Eiger G, Shamir U and Ben-Tal A (1994) Optimal design of water distribution networks. Wat Resour Res 30(9): 2637-2646
[7] Fujiwara, O. and Khang, D.B. (1990): A Two-Phase Decomposition Method for Optimal Design of Looped Water Distribution Networks, Water Resources Research, 26(4), 539-549.
[8] Goldberg, D.E. (1989): Genetic Algorithms in Search, Optimisation and Machine Learning. New York, Addison-Wesley.
[9] Geem, Z. W., Kim, J. H. and Loganathan, G. V. (2002): Harmony search optimization: application to pipe network design. International Journal of Modelling and Simulation, 22(2), 125-133.
[10] Kessler, A. and Shamir, U. (1989): Analysis of the Linear Programming Gradient Method for Optimal Design of Water Supply Networks, Water Resources Research, 25(7), 1469-1480.
[11] Reca J, Martinez J, Gil C, Banos R (2008) Application of several meta-heuristic techniques to the optimization of real looped water distribution networks. Water Resour Manage 22(10):1367–1379, DOI 10.1007/s11269-007-9230-8
[12] Zecchin AC, Maier HR, Simpson AR, Leonard M & Nixon JB (2007) Ant colony optimization applied to water distribution system design: comparative study of five algorithms. J Wat Resour Plan and Manage, 133(1), pp. 87-92
[13] Simpson AR., Dandy GC and Murphy LJ (1994) Genetic algorithms compared to other techniques for pipe optimization. J Wat Resour Plan and Manage ASCE 120(4): 423-443

---

doc. Ing. Milan Čistý PhD.
milan.cisty@stuba.sk
Katedra vodného hospodárstva krajiny
Stavebná fakulta Slovenskej technickej univerzity v Bratislave
Radlinského 11
813 68 Bratislava

Ing. Zbyněk Bajtek PhD.
Zbynek.bajtek@stuba.sk
Katedra vodného hospodárstva krajiny
Stavebná fakulta Slovenskej technickej univerzity v Bratislave
Radlinského 11
813 68 Bratislava