Water Resour Manage (2010) 24:1–24
DOI 10.1007/s11269-009-9434-1

Hybrid Genetic Algorithm and Linear Programming Method for Least-Cost Design of Water Distribution Systems

Milan Cisty

Received: 31 July 2008 / Accepted: 13 April 2009 /
Published online: 30 April 2009

Abstract The problems involved in the optimal design of water distribution networks belong to a class of large combinatorial optimization problems. Various heuristic and deterministic algorithms have been developed in the past two decades for solving optimization problems and applied to the design of water distribution systems. Nevertheless, there is still some uncertainty about finding a generally trustworthy method that can consistently find solutions which are really close to the global optimum of this problem. The paper proposes a combined genetic algorithm (GA) and linear programming (LP) method, named GALP for solving water distribution system design problems. It was investigated that the proposed method provides results that are more stable in terms of closeness to a global minimum. The main idea is that linear programming is more dependable than heuristic methods in finding the global optimum, but because it is suitable only for solving branched networks, the GA method is used in the proposed algorithm for decomposing a complex looped network into a group of branched networks. Linear programming is then applied for optimizing every branch network produced by GA from the original looped network. The proposed method was tested on three benchmark least-cost design problems and compared with other methods; the results suggest that the GALP consistently provides better solutions. The method is intended for use in the design and rehabilitation of drinking water systems and pressurized irrigation systems as well.

Keywords Genetic algorithms · Linear programming · Optimization · Water distribution system

M. Cisty (✉)
Department of Land and Water Resources Management, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovak Republic
e-mail: milan.cisty@stuba.sk

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. The problems involved in the optimal design of this system are huge; they belong to a class of problems known as NP-hard problems, where the problem is intractable, and it is not practical to perform a full enumeration using any rigorous algorithm. For this reason, various algorithms ranging from artificial intelligence to the optimization domain were applied. Alperovits and Shamir (Alperovits and Shamir 1977) presented a linear programming gradient for optimizing a water distribution network. This is an iterative method where, at each iteration, a fixed set of flows is tried, and every pipe is divided into segments, each with a different diameter. The decision variables are the lengths of each segment, and the problem is reduced to a linear one. In the successive iterations, the flow variables are heuristically adjusted according to the gradient of the objective function. Kessler and Shamir (Kessler and Shamir 1989) used the linear programming gradient method as an extension of this method. It consists of two stages: an LP problem is solved for a given flow distribution, and then a search is conducted in the space of the flow variables. Later, Fujiwara and Khang (Fujiwara and Khang 1990) used a two-phase decomposition method extending the method of Alperovits and Shamir to non-linear modeling. Also, Eiger et al. (1994) used the same formulation as Kessler and Shamir, which leads to a determination of the lengths of one or more segments in each link with discrete diameters. Nevertheless, these methods fail when solving problems of large looped systems.

The researchers have focused on stochastic or so-called heuristic optimization methods from the early 1990s. Simpson et al. (1994) 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. The simple GA was then improved by Dandy et al. (1996) using the concept of the variable power scaling of the fitness function, an adjacency mutation operator, and gray codes. Savic and Walters (Savic and Walters 1997) also used a simple GA in conjunction with an EPANET network solver.

Other heuristic techniques have also been applied to the optimization of a water distribution system, such as simulated annealing (Loganathan et al. 1995; Cunha and Sousa 2001); an ant colony optimization algorithm (Maier et al. 2001); a shuffled frog leaping algorithm (Eusuff and Lansey 2003) and a harmony search (Geem et al. 2002), to name a few.

Nowadays, the main concern in developing methods for the design of water distribution networks consists of solving various types of multi-objective tasks (for instance, a combination of an economic design with the reliability of a water distribution system, leakage reduction, location of the water quality sensors, etc.). These tasks could be implemented quite effectively in existing multi-objective optimizers such as NSGA-II, SPEA or others. The question is how different are pareto fronts obtained from these computations from the true optimal pareto fronts of these tasks (e.g., from a global optimum). Although research exists which deals with the task of performance assessment of multi-objective optimizers (e.g., Zitzler et al. 2003) mainly methods for relative comparisons of existing algorithms are available. That is why this paper deals with a least-cost design where a measurement of an approximation to a global optimum could be more easily evaluated. The author

expects that when an optimization model with a better approximation to the global optimum for this task evolves; a better basis for more complicated multi-objective tasks will also be obtained.

Significant differences from known global optimums are referred to even for single objective tasks and simple benchmark networks, while existing algorithms are applied. Reca et al. (2008) evaluated the performance of several meta-heuristic techniques—genetic algorithms, simulated annealing, tabu search, and iterated local search. He compared (among other testing accomplished) these techniques by applying them to medium-sized benchmark networks. The results which he obtained for the Hanoi network (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 problem. Similar results were presented by Zecchin (Zecchin et al. 2007) in a comparative study of ant colony optimization algorithms in which other heuristic algorithms were also tested. It could be expected for larger networks and networks for which the global optimum is not known (i.e., what should be obtained from computations), that this difference would be even greater (e.g., more than 6% for the network tested in Cisty et al. (1999). It should be noted that real life problems never get the privileges afforded the Hanoi or other benchmark network, which were computed in various optimizer hours or maybe years of CPU time. This runtime reduction for real designs also increases the possibility of greater errors in the solutions obtained in comparison to the global minimum. Similarly this difference between the local optimum obtained and the global optimum would logically be even greater for more complicated tasks (multi-objective optimization).

The difference from the global optimum mentioned is not only a problem of a bit higher cost of the network design, but designs which are closer to the global optimum are also usually more logical designs from an engineering point of view. The common experience of users who apply heuristic techniques (which offer suboptimal solutions) is, for instance, the observation that in the results of the computations, it often occurs that a larger diameter of a pipe is proposed by these methods after a smaller one in the flow direction (in the direction in which the flows are smaller because of the demands from the network), which is not acceptable in practical design.

Research on the optimal design of water distribution systems continues to this day. Many published papers have been concerned with improving the effectiveness of traditional heuristic algorithms, while others have been concerned with their efficiency or developing new search methods. For example, a recent development in the field of evolutionary algorithms is to use probabilistic methods to identify key building blocks (short, highly fit groups of genes). These techniques are replacing the traditional crossover and mutation methods in genetic algorithms for the generation of offspring with the building and sampling of a probabilistic model describing the genomes in a set of promising solutions (Olsson et al. 2008). New algorithms are being tested—e.g., Raad and Vuuren (2008) applied a modern metaheuristic framework known as AMALGAM, which was developed in 2007 by Vrugt and Robinson (2007). This framework allows for the simultaneous incorporation of multiple heuristics within a generic evolutionary framework in the hope of improving performance and efficiency by adopting the philosophy of power in diversity. Moreover, completely different methods are being proposed: Ekinci and Konak (2009)

proposed a method by which they built on the idea that head losses through pipelines should be minimal and as nearly equal as they can possibly be. In their study, the initial flows and pipe diameters are determined by a weighting optimization process to get a reliable solution satisfying those conditions in a concrete manner. A datum optimization procedure is realized to get a least-cost design alternative giving the optimum conditions for the system.

Nevertheless, it must be pointed out that there still exists uncertainty as to how close any heuristic method could approximate solving the task of a global optimum. The main concern of this paper is to propose a method which is more dependable in converging more closely to a global optimum than existing algorithms.

A heuristic search is a method that might not always find the best solution, but is usually capable of finding a near optimum solution in a reasonable time. The advantage of applying these methods in a water distribution design is that very complicated tasks can be solved by them, because heuristic methods have the potential to take everything that is possible to compute by a simulation model of a water distribution system as an optimization objective. On the other hand, the issue of precision is problematic.

How closely any solution from the computation of the optimization in a particular case could achieve a global optimum depends on the complexity of the problem. Generally, a more complex problem is a problem for which a larger or more complicated search space must be explored. The reduction of the search space is the main approach of this paper on how to design a method which has a greater potential to find results which are really close to a global optimum. A new hybrid GA-LP approach using a genetic algorithm and linear programming is proposed in this study for determining the least-cost design of a water distribution system. It is built on the advantages of both the deterministic and heuristic methods. The GA method is used in the outer loop of the proposed algorithm, which is intended for decomposing a complex looped network in a group of equivalent branched networks. LP is then used in an inner loop to solve each branched network and provide a minimum-cost design. After evaluating a high number of possible branched networks (by LP, which is nested in a GA fitness function), an optimal solution is found for the original looped network.

The advantage of using this hybrid method consists in the fact that a GA in this case has a much smaller searching space than in a case when GA methodology is used alone, which has a great impact when trying to achieve better results. How this is achieved will be explained later after introducing the details of the methodology.

2 Methodology

2.1 Deterministic Part of the Proposed Method—Linear Programming

The linear programming method has long been accepted as an approach for the optimal selection of diameters for pipes in branched networks, e.g., in the design of irrigation systems. The mathematical formulation of this problem is as follows:

$$\begin{aligned} X_{11} + X_{12} + \ldots + X_{1n} &= B_1 \\ X_{21} + X_{22} + \ldots + X_{2n} &= B_2 \end{aligned} \tag{1}$$

etc.

$$\begin{aligned} X_{m1} + X_{m2} + \ldots + X_{mn} &= B_m \\ A_{11}X_{11} + A_{12}X_{12} + \ldots + A_{xy}X_{xy} &\leq C_1 \end{aligned} \tag{2}$$

etc.

$$D_1X_{11} + D_2X_{11} + \ldots + D_jX_{mn} = \min \tag{3}$$

The solution has to comply with inequalities:

$$X_{11} > 0;\ X_{12} > 0 \text{ etc. up to } X_{mn} > 0 \tag{4}$$

Where

X_{ij} is the unknown length of the selected diameter j on section i. The allowable diameters must satisfy the velocity conditions on the section.
B_i the total length of the section
C_i allowable total loss for the constraint—described below
D_i the unit price of a pipeline with the diameter number i

The choice of a split-pipe design was made for the present work (more than one diameter can be proposed for a section; a section is the pipe connection between two nodes). The possible problem, that the algorithm will propose small lengths of particular diameters on a section could be easily overcome by including an additional system of expressions which would fix the minimal lengths by "greater than…" conditions (greater than the fixed length or a fraction of the total length of the section). Moreover, a discrete diameter design is possible to fix on some short sections, when a formulation of the linear programming method with binary variables is used. And finally, a split pipe design offers a better cost on the same network than are those obtained from a discrete diameter design, so there is some reserve for correction of the final solution.

When linear programming is applied in order to solve the optimal design of pipeline networks, the unknown will be the lengths of the individual pipeline diameters on the section. In conditions (Eq. 1) the requirement that the sum of the unknown lengths of the individual diameters in each section has to be equal to its total length is enforced. The second type of equations—constraints (Eq. 2)—represents the condition that the total pressure losses in a hydraulic path between a pump station or tank and every critical node (the end of the pipe network; the extreme elevation inside the network) should be equal to or less than the known value. This constraint is based on the minimum network pressure requirements needed for the operation of the system. There should be the same number of these constraints as there are critical nodes in the network. Given the minimization requirement for the investment costs, the objective function (Eq. 3) is the sum of the products of the individual pipeline unit prices and their required lengths. As will be described later, this system of the equation should be compounded automatically in the fitness function of the GA.

When multi-demand conditions in the LP model are incorporated, there will be a system of constraint (Eq. 2) for every demand pattern. When pumps are also included in a model, the main input parameter for a pump is its pump curve. The right sides of the constraints (Eq. 2) vary according to the pump's operating conditions. The head of the pump may be treated as an unknown variable.

2.2 Heuristic Part of the Proposed Method—Genetic Algorithms

In order to overcome the above-mentioned deficiencies of the linear programming techniques, heuristic optimization techniques have been introduced for solving the optimization of water distribution systems. Firstly, a genetic algorithms methodology, which is also used in this study was applied. This is a 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 concepts are briefly summarized below; a good introduction to the subject is given by Goldberg (1989).

The first step is to represent a solution to the problem by a string of genes 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 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. At each generation, the fitness of each chromosome in the population is measured. The fitter chromosomes are more often selected probabilistically to produce offspring for the next generation. This process is repeated until some form of convergence in the fitness is achieved. The goal of the optimization process is to maximize the fitness.

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 simplify the problem to designing only new pipes, the chromosome can be a string of the possible diameters of its corresponding section. An efficient and effective search for the optimum design solution of a water distribution network using genetic algorithms is governed by 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.

2.3 Hybrid GA—LP Approach to the Optimal Design of Water Distribution Systems

2.3.1 Basic Description of the Algorithm

The proposed method is based on a combination of linear programming methodology and a genetic algorithms approach. The main reason is that linear programming always finds the global optimum if it exists. But because LP is suitable only for solving branched networks, the GA method is used for decomposing a complex looped network into a group of branched networks. These branched networks are characterized by hydraulic behavior identical to a looped network on the condition of having identical diameters on the corresponding sections of the network. Identical hydraulic behavior means that there are identical flows in the corresponding pipes and identical pressures in the corresponding nodes in the original looped network and branched networks investigated. The decomposition of a looped network means that before the optimization, every loop is split in some demand node (or node in which a branch is connected) which is part of this loop. There are many possibilities for accomplishing such a decomposition in a water distribution network. Linear

programming is then applied for optimizing every branch network produced by GA from the original looped network, and GA is simultaneously applied for the evolution of the best splitting option.

The main aspect of the motivation to propose the new optimization algorithm presented is that the definition of a chromosome is in an obvious GA approach as long (in the sense of the number of genes) as many pipe sections are in the water distribution network. The definition of a chromosome proposed herein produces significantly shorter strings (which mean a smaller search space) for most networks because the number of genes will be the same as the number of loops in the network and not equal to the number of pipes. This means easier searching with better results.

An example of possibilities for splitting a loop is shown in Fig. 1. Let us suppose that all the parameters of the network in Fig. 1a) are already known (diameters, lengths, demands, flows, etc.), so we are not in the design stage for this network, but we can suppose that it already exists. The loop can be transformed to a branch layout without affecting the hydraulic behavior of the network (the flows and pressures remain the same) if the loop is split in the demand node in which the flows are entering from the two pipes connected to this node. There is one such node for every loop in the network. This node could be a hydrant or branch connection. In Fig. 1a) the original loop with a node in which there is a demand $3\,\mathrm{l\,s^{-1}}$ is shown. As can be seen, flow $2\,\mathrm{l\,s^{-1}}$ is entering this node from the left side, and flow $1\,\mathrm{l\,s^{-1}}$ is entering this node from the right side. So this is the node in which it is possible to break the loop without affecting the hydraulic behavior of the network in the current conditions. A loop could be split in this node in such a way that a "twin" node is introduced to the network; the identical elevation on it is assumed as in the original node (and de facto identical position). The original demand has to be split between these two nodes: the left node will have the demand $2\,\mathrm{l\,s^{-1}}$ and the right node $1\,\mathrm{l\,s^{-1}}$ (Fig. 1b). This is what could be (theoretically) done on the existing network. The branch network obtained is hydraulically identical to the original looped network.

The situation is slightly different on a network which is just in the design stage. Because the parameters of the network such as flows, etc., are not known before the design of the network is ready, a loop split in the computation of the optimization should be tried and evaluated in every demand node which is part of the loop. Moreover, the original demand in every split node selected has to be divided between the original node and its "twin node" according to some rule with a rational number of alternatives (e.g., Fig. 1b, c). The number of alternatives depends on the demand rate—a high demand in a node requires more alternatives. So the number of demand nodes multiplied by the number of demand-splitting alternatives gives the number

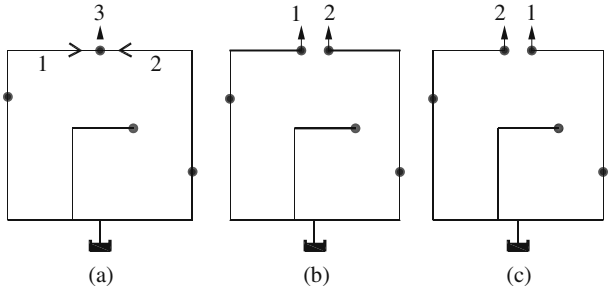

Fig. 1 Alternatives for splitting the loop in one node—**a** original loop network, **b** original network transformed to a branch network, **c** another possibility of dividing the demand in the design stage

of possible branch networks in which it is possible to divide one loop network. For networks with more loops, a combination of alternatives for every loop should be considered. Because we are now talking about designing the network, it is necessary to propose diameters for every alternative of the branch network which we get by this procedure. This is accomplished by LP. The cheapest one is an optimum which an algorithm is searching for. The search for the best possibility of splitting the loops is guided by the GA, which proposes one possible branch network for every chromosome in which both the splitting nodes and parameters for splitting the demand in these nodes are coded. The branch network (or this chromosome) is evaluated by accomplishing its design by LP, which is nested in the fitness function of the GA. This is one iteration of the algorithm. The population of chromosomes evolves toward better solutions by means of genetic operators such as inheritance, mutation, selection, or crossover as is usual in the application of GA. With each generation, the fitness of each chromosome in the population is calculated, and better splitting options are evolving.

2.3.2 Details of the Proposed Algorithm

In this part of the paper some details of the proposed algorithm are discussed. These include formulating the definition of the chromosome and establishing the so-called Loop Links matrix (**LL**) and Loop Splitting Options matrix (**LSO**), both of which help to organize searching for the best loop-splitting locations on the network, and finally the R and R_s parameters for splitting the demands in the nodes where the loops are divided. All these concepts are explained hereinafter with the help of Fig. 2.

Because the basic principles of GA are known and were briefly described in Part 2.2, there is no whole algorithm scheme but only one demonstrative iteration of the proposed algorithm displayed in Fig. 2. Therefore, the scheme displayed therein explains the structure of a chromosome, its decoding and an evaluation of its fitness. Other things such as initialization of the population, the evolving of the population through genetic operators to the next generations, etc., work as usual in GAs.

The first thing which should be determined in this type of algorithm is *the definition of a chromosome*. The chromosome of the illustrative three-looped network displayed in Fig. 2 can be seen on its left side. In the proposed method the chromosome consists of two parts. The first half of the genes holds the value of a parameter, which defines where a particular loop should be divided. There are the same number of these genes as the number of loops in a network. Genes are coded as integer numbers, which indicate a row number in an **LSO** matrix. This matrix defines all the splitting possibilities for every loop and will be described hereinafter. The second half of the genes in a chromosome is also coded as an integer value, and this value defines the ratio by which the demand in the original split node should be divided between the original node and its "twin node" after splitting the particular loop in its location.

For every loop the vector of links of which it consists should be determined. This will be done in successive steps for loop after loop, and when some link has already been assigned in preceding the algorithm run to any previously analyzed loop (it is already part of the vector of pipes for some loop), it will not be assigned to a vector of links of a loop subsequently analyzed. The result of this analysis is an **LL** matrix (e.g., Loop Links). It has a number of columns equal to the number of loops and a

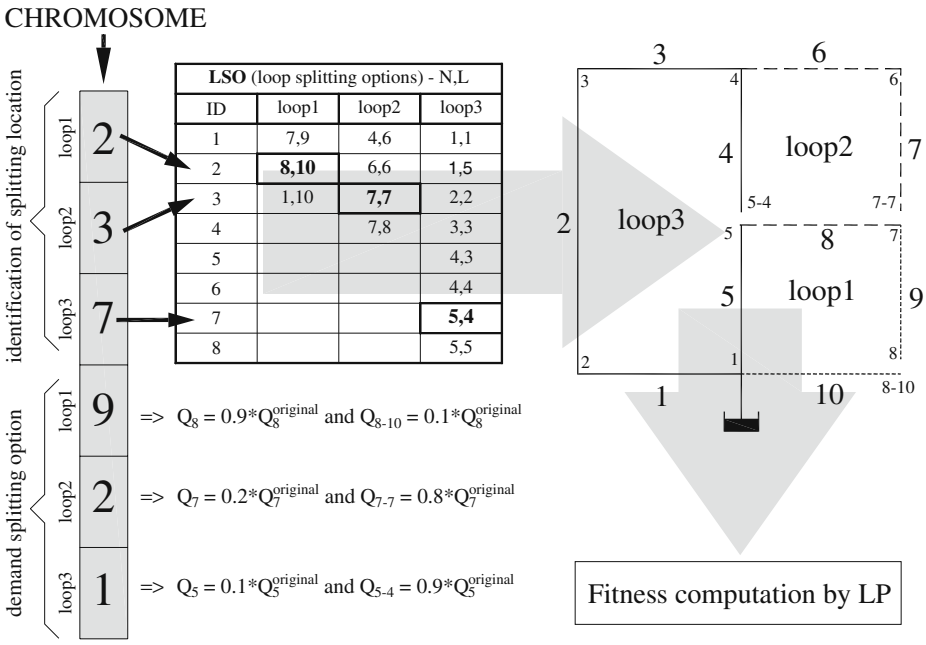

Fig. 2 Explanation of the basic concepts of an algorithm on one of its iteration

number of rows equal to the number of links in the largest loop of the network. This assignment of pipes to loops (or **LL** columns) is shown in the network scheme in Fig. 2 (loop 1—links 9, 10, loop 2–6, 7, 8 and loop 3–1, 2, 3, 4, 5). It is visualized here by the line types in Fig. 2.

An additional matrix is established in which the locations for splitting loops are defined. It is the **LSO** (Loop Splitting Options) matrix in Fig. 2. It has a column for every loop and a row for every splitting possibility for the loops. The main value in the cells of this matrix is the identification (number) of the splitting node N. It is a node in which a loop can be divided. There is also a second parameter, which has not yet been discussed, because it belongs to the details of the method. Together with the splitting node split link L should also be specified. The split link is chosen from the column of the **LL** matrix for a particular loop, and it is the one which is connected to N. There can be one or two possible links L for every node N. The original node N will be cloned to its twin node N–L for the sake of splitting the loop. This twin node N–L replaces the original node N on link L, but on the other pipes which were in the original looped network connected to N, this original node N remains. It is the mechanism through which the idea of splitting the loops is realized. It is possible to explain the purpose of the second parameter L for a definition of the splitting locations (and why only the definition of the splitting node is not enough) with this example: as could be imagined with the help of Fig. 2, it produces different configurations of branched networks derived from the original looped network if the split combination N, L for loop 3 is 5, 4 or 5, 5 (in both cases the loop is split in the node $N = 5$, but L is different).

The second half of the genes in a chromosome have the purpose of defining the splitting of the demand in the node N selected between the original node and its "twin" node $N–L$ for every loop. For this purpose parameters R and R_s are defined for every loop in the network. Parameter R is a real number between 0 and 1. The demand in the split node is equal to $Q_N = RQ_N^{\text{original}}$ and $Q_{N-L} = (1 - R)\, Q_N^{\text{original}}$ in its "twin" node. We can define the set of allowable R by specifying R_s—a step for R which is a number between 0 and 0.5. If R_s is set to 0.2, then possible values of R could be {0.2; 0.4; 0.6; 0.8}. It is usually not less than 0.01, and it is recommended that R/R_s be an integer number. Parameter R for a particular loop is selected by a genetic algorithm as the value of a corresponding gene from the second half of the chromosome genes (Fig. 2). In the example in Fig. 2 R_s is equal to 0.1; for the sake of the integer coding is R multiplied by 10. In general the set of allowable R is every multiple of R_s between R_s and $R–R_s$. If there should also be a possibility incorporated in an optimization algorithm to eliminate some pipes from the network originally included in the input data, R could be every multiple of R_s between 0 and 1 (inclusive). In this case if R were finally set by computation to 0 or 1, then in some pipe connected to the node there will be zero flow, so there is no need for this pipe in the network.

The proposed algorithm in its basic form can be summarized as follows:

Read input data (Epanet INP file, R_s for all loops, GA parameters, etc.)
*Matrix **LL** is determined*
*Matrix **LSO** containing loop splitting possibilities is determined*
Initial population generation
While termination condition is not satisfied, do:
{
*Decoding chromosomes with the help of **LSO** and R_s (which means the creation of as many branch networks as there are chromosomes in the population)*
Evaluation of the fitness of each individual in the population by LP
Select the best-ranking (cheapest) individuals to reproduce
Breed a new generation through crossover and/or mutation (genetic operations) and give birth to offspring
Evaluate the individual fitness of the offspring
Replace the worst ranked part of the population with offspring
}
Best solution output

In the last part of this section three remarks on the details of the proposed algorithm follow.

Firstly, one more condition should be added to the LP model in the case of using it in the context of solving looped networks.

The principle of the conservation of energy dictates that the difference in energy between two points must be the same regardless of the path that is taken. Thus, the difference in energy at any two points connected in a network is equal to the energy gains from the pumps and the energy losses in the pipes and fittings that occur in the path between them. This equation can be written for any open path between any two points. Of particular interest are paths around loops because the changes in energy must sum to zero. A linear programming model must also pay attention to this principle, so its formulation, which is expressed by formulas 1–4 for the branched network only, have to be expanded by the loop condition in the

proposed methodology; however each one of them was split into two branches by the described method (Figs. 1 and 2). There should be the same pressure in the node in which the loop was breaking and in its twin node after proposing the diameters in the network (Fig. 1). Therefore, while the LP model is being built, the following conditions should be added, which ensure that the same pressure must be in the original and corresponding dummy twin nodes:

$$E_i = A_{11}X_{11} + A_{12}X_{12} + \ldots + A_{xy}X_{xy} \tag{5}$$

$$E_{ti} = A_{11}X_{11} + A_{12}X_{12} + \ldots + A_{xy}X_{xy} \tag{6}$$

$$E_i - E_{ti} = 0 \tag{7}$$

Where

E_i and E_{ti}	are energy losses from the source (pump, reservoir) to the split node on loop i and to its twin node
A_{mn}	hydraulic loss in section m and diameter n, which belongs to the specified path
X_{ij}	is unknown length of the selected diameter j on section i

The second remark deals with the possible refinement of the proposed method, which could be applied especially when the larger network is being solved by it. By applying it, even a smaller search space (relatively) is to be searched for the sake of finding the optimal solution. This refinement is derived from the detection of the fact that for finding the optimal solution by the proposed method, it is much more important to find locations where the loops should be divided, than the determination of the R factors for dividing the demands in these locations between the two nodes mentioned. The following refinement is especially useful when one is solving networks with a large magnitude of demands.

In the proposed model a preprocessing phase could be implemented. In this phase the model is run with only half of the genes in the chromosome, because for all the split nodes, ratio R is fixed at 0.5. So there is only one possibility for dividing the demands, which means there is no need for the second half of the genes (Fig. 2), and they are left out from the definition of the chromosome in this phase. This produces an even smaller search space than in the basic GALP algorithm, which can in many cases be investigated with Steady State GA (although it is not necessary). A Steady State GA is a type of genetic algorithm in which only the worst member of the population gets replaced with each iteration. This method of progression tends to arrive at a good solution much faster than the Generational GA, which is used in the main part of the algorithm (a Generational GA is a type of genetic algorithm in which the entire population is replaced with each iteration). The preprocessing phase is based on the experience that in searching for an optimal design, it is more important to find the location of split nodes than to find the best ratio R for splitting the demands between the original and split nodes. The solution, which was found in the preprocessing phase, is seeded to the initial population of the basic part of the algorithm with a full chromosome which is described by a pseudocode hereinbefore. This means that one or two chromosomes from the initial population of the basic part

of the algorithm (which run with the complete chromosome) are changed according to the best chromosomes from the preprocessing phase. Then the main phase with a complete chromosome is evaluated by the Generational GA. When the optimal solution is found, the third part of the algorithm can be (eventually) run. This serves for the refinement of the solution which was found in the main (second) phase. This refinement consists in better splitting of the demands in the already defined splitting locations, which were found in the previous phases. Once again there is a reduced definition of the chromosome used in the post-processing phase, but in this case only the second half of the genes are present in its definition. In this stage only one splitting options from **LSO** which were found in the previous step are used (set of N, L for all the loops). So there is no need for the first half of the genes (Fig. 2). For the second half of the genes the same R_s value as in the previous phase or the smaller one could be used. Usually there is no need to run all three phases; running only the second phase is possible; running the first and second is another option; or running the first and third would be enough for finding the optimal solution.

The scheme of the algorithm with these refinements is shown in Fig. 3.

The third remark on the details of the proposed algorithm relates to its compatibility with EPANET. In an obvious simulation-based GA approach a hydraulic network solver handles the pressure and velocity constraints and simultaneously evaluates the hydraulic performance of each trial solution. The most commonly used simulation model to analyze the network in such a manner is EPANET. In the proposed method

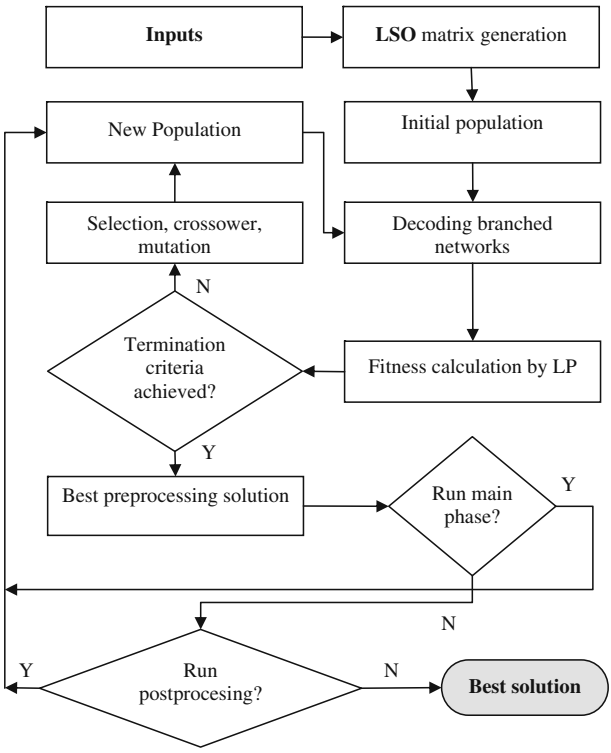

Fig. 3 The scheme of the GALP algorithm with preprocessing and postprocessing phases

simulation by EPANET (or another hydraulic engine) is not incorporated in the core of the objective function (or fitness function), but linear programming is used instead. But for the sake of compatibility with other optimization models, the computation of the friction headlosses A_{mn} (Eq. 2) is executed by EPANET by calling its functions through the EPANET Toolkit.

3 Results

The performance of the GALP model developed for the optimization of the least-cost design of a water distribution network problem is evaluated by the optimization of the well-known Hanoi network and the double and triple Hanoi water supply networks. The first problem is taken from the literature. The second and third problems were introduced by the author for the sake of evaluating the proposed method for greater problems than the known and thoroughly investigated benchmark models. There is also a second reason why the Hanoi network and its multiplications were chosen – the main goal of this paper is to find a method which offers results closer to the global optimum than existing methods do. The global optimum for the Hanoi network is known because the long usage of this benchmark in the optimization methods development community. The double and triple Hanoi networks are derived from it in such a way that their global optimums could also be evaluated. On this basis it is possible to compare the results obtained in testing runs with the known global optimums in all three cases.

The water distribution trunk network in Hanoi, Vietnam, which was first introduced by Fujiwara and Khang (1990) is shown in Fig. 4. The network consists of 34 pipes, 32 nodes, and three loops. It is a gravity-fed system from a single fixed head

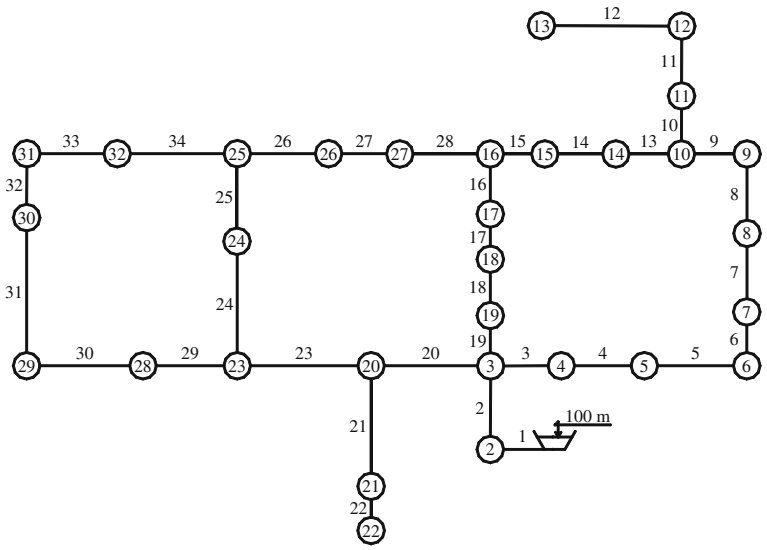

Fig. 4 Hanoi network

source and is designed to satisfy given demands at the required pressures. In this problem six sizes of commercial pipe diameters are available, and the cost of each pipe i with a diameter D_i and length L_i is calculated from $C_i = 1.1 \times D_i^{1.5} L_i$, where the cost C_i is in dollars, the diameter is in inches, and the length is in meters. The Hazen-Williams coefficient is fixed at 130 for all the pipes. The data necessary for the optimization can be found in the work of Fujiwara and Khang (1990).

The network consists of three loops, so three **LSO** matrices were determined by the algorithm. Roulette wheel selection was used to choose the parents for the next generation. A one-point crossover was used because of the relatively short chromosome, and the probability of the selected pair of strings being subjected to the crossover operator was taken as $p_c = 0.9$. The mutation rate was set to be $p_m = 0.1$. A very simple penalty function was used: if the LP (the final producer of every partial solution) does not find a solution (which could happen in some configurations), the algorithm gives this solution a significantly higher cost than the highest cost in the previous generation. The preprocessing phase of the algorithm runs as Steady State GA, and it found a solution of \$6,116,086.98 on every run in iteration 185 or very close to that (which is a few seconds). This means that this is a really simple problem for the algorithm. In the main part of the algorithm the ratio R dedicated to dividing the original demand between the twin nodes in which the network is disjoined was chosen as 0.1, and the population was seeded with local minima obtained in the previous step. In this part the algorithm found a solution of \$6,057,697 with 1655 iterations. The third part of the algorithm was not run for this network. The important thing here is that these computations were performed without any refinement of the GA settings; although the experiments were conducted by changing the settings of the main parameters, the impact of these changes on the results was found to be rather small, both in the terms of the optimal cost and the number of iterations. This means that for this network and for networks of a similar size, there is no need to perform very often tedious and time-consuming fine tuning of the GA parameters. The author chose the alternative described in Tables 1 and 2 in which the preprocessing and main phases of the algorithm were running as the reference (global) optimal solution. It is the best solution for the optimization of the Hanoi network referred in the literature which is feasible in terms of the allowable pressures (computed by EPANET). Because the proposed method has stochastic features, the result is not exactly the same on every run, but as was mentioned above, it is always very similar. That is why the run-time plots and other statistical evaluations of the computation will be described hereinafter only for the triple Hanoi network, which is larger (100 pipes).

The performance of the GALP model for the optimization of the least-cost design of a water distribution network problem is also evaluated by its application for the double and triple Hanoi water supply networks. Because these two networks are derived from the basic Hanoi network, their optimal cost is known (Figs. 5 and 6).

All the parameters for the reservoir, nodes and lines in the double Hanoi water distribution network are the same as in the original Hanoi network on both mirrored parts except for the first pipe (from the reservoir to node 2), which is shortened from the original 100 to 28.9 m. This change was made for the sake of obtaining the same head in node 2 (with a diameter of 1016 mm, which will certainly be proposed here by any optimization method) as in the original Hanoi network. In such conditions the optimal solution for the double Hanoi network should have the same diameters on

Table 1 Results for the Hanoi network

Pipe ID	Length (m)	Diameter (mm)
1	100	1,016
2	1,350	1,016
3	900	1,016
4	1,150	1,016
5	1,450	1,016
6	450	1,016
7	850	1,016
8	850	1,016
9	97.5	762
	702.5	1,016
10	950	762
11	1,200	609.6
12	3,500	609.6
13	299.5	406.4
	500.5	508
14	500	406.4
15	550	304.8
16	2,730	304.8
17	1,750	406.4
18	400.3	508
	399.7	609.6
19	400	609.6
20	2,200	1,016
21	489.1	406.4
	1,010.9	508
22	500	304.8
23	2,650	1,016
24	1,230	762
25	1,300	762
26	850	508
27	263.7	304.8
	36.3	406.4
28	750	304.8
29	1,500	406.4
30	1,629.1	304.8
	370.9	406.4
31	1,600	304.8
32	150	406.4
33	860	406.4
34	59.9	508
	890.1	609.6
Cost ($)	6,057,697	

the corresponding pipes as in the original Hanoi network (on both mirrored parts). In node 2 the same demand is as in the original Hanoi network; it is not doubled. Under these conditions the *reference* optimal solution (global) could be evaluated as follows:

$$C_{\mathrm{DH}} = 2C_{\mathrm{H}} - 2L_1C_1 + 28.9C_1 \tag{8}$$

Table 2 Optimal pressure heads for the Hanoi network

Node ID	Pressure (m)
1	0
2	97.14
3	61.67
4	56.98
5	51.17
6	45.05
7	43.62
8	41.91
9	40.56
10	39.2
11	37.64
12	34.21
13	30
14	33.37
15	31.88
16	30.24
17	32.92
18	49.95
19	58.93
20	50.5
21	35.16
22	30
23	44.33
24	38.66
25	35.01
26	31.13
27	30
28	38.42
29	30
30	30.13
31	30.4
32	32.72

Where

C_{DH} is the optimal cost of the double Hanoi network
C_H reference optimal cost of the Hanoi network ($6,057,697)
L_1 length of the first pipe on the original network (100 m)
C_1 unit price of diameter 1,016 mm ($278.28)

For our solution described in Tables 1 and 2 (cost $6,057,697), which is the best solution of the basic Hanoi network compared to the results known to the author, this Means (according to Eq. 8) that the reference optimal solution of the double Hanoi network should be $12,067,780.29. Only solutions which are feasible in terms of the allowable nodal pressures computed by the EPANET network solver are taken into consideration.

In the triple Hanoi water distribution network all the corresponding parameters for the nodes and lines are the same as in the original Hanoi network on all three (single Hanoi network) parts except for four pipes, the head in the reservoir and the demand in one node. These changes were made for the sake of obtaining the same

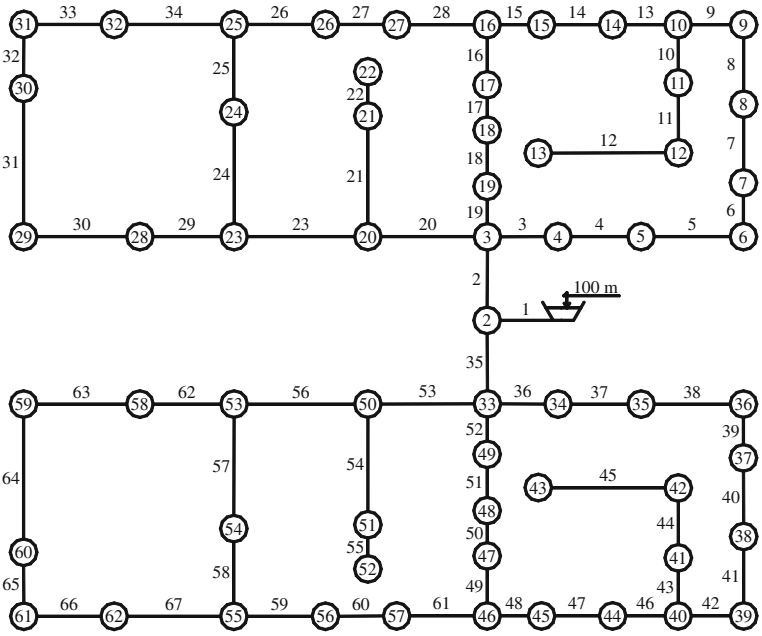

Fig. 5 Double Hanoi network

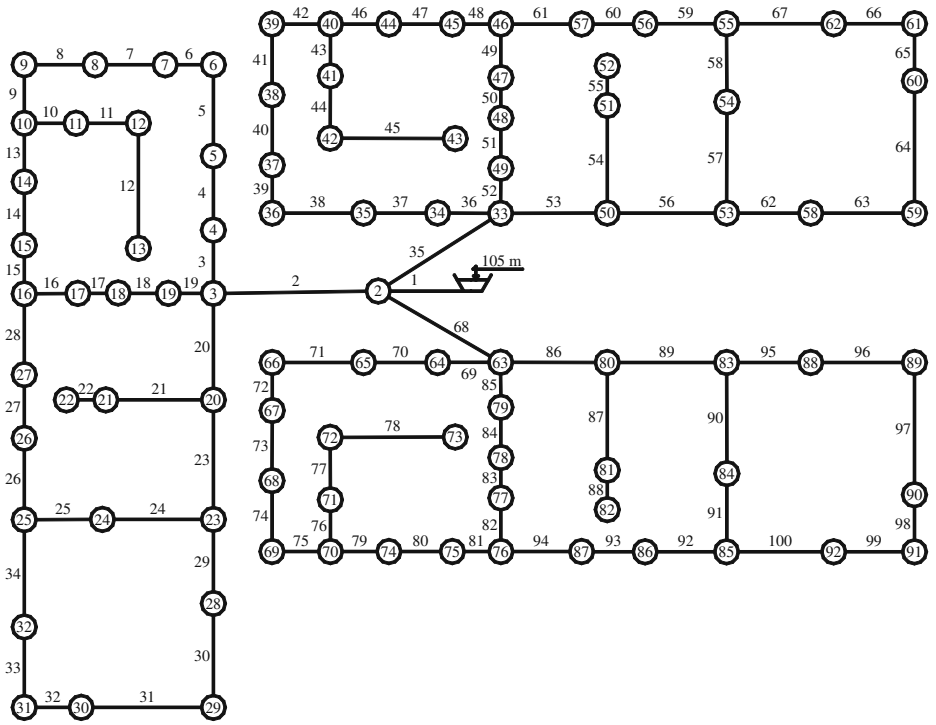

Fig. 6 Triple Hanoi network

pressure in nodes 3, 33 and 63 (with a diameter of 1,016 mm on pipes 1, 2, 35 and 68, which will certainly be proposed here by any optimization method because of the large flow in them) as in the original Hanoi network in node 3. In such conditions the same diameters should be the optimal solution for the corresponding pipes as in the original Hanoi network. These are the changes mentioned: the head in the reservoir is set to 105 m; the length of pipe 1 is 1 m; the length of pipe 2 is 1,786.50 m; and the lengths of pipes 35 and 68 are 1,641.69 m. In junction 3 the demand is equal to zero. Under these conditions the *reference* (global) optimal solution of the triple Hanoi network could be evaluated as follows:

$$C_{TH} = 3 * C_{\mathrm{H}} - 3\ L_1 C_1 - 3 L_2 C_1 + (1 + 1,786.5 + (2 \times 1,641.69))\ C_1 \qquad (9)$$

Where

C_{TH} is the optimal cost of the triple Hanoi network
C_{H} reference optimal cost of the Hanoi network ($6,057,744)
L_1 length of the first pipe on the original network (100 m)
L_2 length of the second pipe on the original network (100 m)
C_1 unit price of diameter 1,016 mm ($278.28)

For our solution, which is the best solution of the basic Hanoi network compared to the results known to the author, this means that the optimal solution of the triple Hanoi network should be $18,373,697.49.

A comparison of the results obtained by the proposed method (which can propose two diameters on some sections) with those published in the literature and obtained by a discrete diameter design (only one diameter is proposed for a section) is a bit problematic, and it is necessary to accomplish it comparatively to the corresponding global minimums for the discrete and split pipe designs. A discrete diameter design has somewhat limited possibilities in comparison with a split pipe design, which can propose two diameters for some sections of the network from the point of view of the resulting cost. Discrete diameter solutions for this reason always cost somewhat more than split pipe design solutions. That is why the *percentual differences* were computed from the reference global minimum—in the case of GALP it is the reference global minimum for a split pipe design, and in the case of a discrete diameter design, another reference global minimum will be taken as the basis. This will be evaluated by Eqs. 8 and 9, but by the global minimum for the original Hanoi network, which was obtained by the discrete diameter design methodology. So we will take as a reference (the best) discrete diameter design solution obtained by (Geem et al. 2002). This author used the harmony search method and found a feasible solution at a cost of $6,081,087, and it is the best feasible solution for a discrete diameter design which has been published in the literature. Only solutions which are feasible in terms of the allowable nodal pressures computed by the EPANET network solver are taken into consideration. When this is taken as the reference (optimal) discrete diameter design of the Hanoi network, the reference (global) optimal *discrete* diameter solution for the double Hanoi network should be $12,114,560.29 and $18,443,867.49 for the triple Hanoi network.

The same parameters were also used for GALP as in the original Hanoi network for the computations of the double Hanoi network. Through the application of the preprocessing and main phase of the proposed algorithm, a solution of $12,073,039

was found for the double Hanoi network. The table with the results (diameters and pressures) is not specified here because it is too lengthy—it is available from the author. Because the proposed method has stochastic features, the results are not always exactly the same, but they are always very similar. That is why the run-time plots and other statistical evaluations of the computation will be described hereinafter only for the triple Hanoi network, which is larger (100 pipes).

More extensive testing was conducted for the triple Hanoi network, because it is the largest of the three benchmarks used in this study, and its evaluation is the most important. The main idea of these computations was to obtain a comparison of GALP with the GA methodology which is incorporated as the heuristic part of the GALP algorithm. The GALP optimization model was run with various GA parameter settings. The size of the population was chosen as 100, 150, 200 or 250 chromosomes in various runs; the number of generations was chosen to be so big so that there were 50,000 iterations in every optimization run (500, 334, 250 and 200). The crossover operator was taken as $p_c = (0.8; 0.9; 0.95)$; the mutation rate was set to be $p_m = (0.05; 0.1; 0.15)$; and every combination between all of these parameter values served as input into the optimization run. This means that 36 runs, each with different parameter settings and 50,000 iterations, were performed. There was no need to run more computations, because the results were very consistent as will be shown.

A comparison of solving all the benchmark problems with the solutions obtained using other methods was made. The author's own GA model (it is the GA which was also used in the GALP), was also more intensively tested with the triple Hanoi network. This was done for the opportunity to compare the performance of a heuristic method with the proposed GALP method (which uses the same heuristic, but is strengthened by including the LP). Similar settings for the GA were used as were described for GALP in the previous paragraph. Because GALP has a little more computationally intensive iteration for its combination of LP and GA, the GA has the advantage of a longer run (not 50,000 iterations but 500,000 iterations in contrast with GALP in every optimization run). In addition to more iterations, 10 runs with every single combination of the setting of the parameters was also run by the GA methodology, and only the best run was taken for an evaluation (this means 360 runs every with 500,000 iterations were performed with GA).

This comparative system of testing GALP and GA is evaluated by the histograms in Figs. 7 and 8. There it can be clearly seen that the errors of the solutions obtained by the GA (the cost of the particular solution minus the global minimum in percents is on the X axis) with respect to the global minimum have a higher order in comparison with GALP. The GA results show that there was often an excess in relation to the global optimum greater than 10%. The histogram of the GALP results also shows that its runs are always very similar and with a significantly smaller excess in relation to the global optimum.

The best results obtained with GALP, the GA optimization model OptiDesigner and the HSNet model based on the Harmony Search methodology are summarized in Table 3. The results in Table 3 demonstrate that the proposed GALP methodology gives significantly better results in terms of closeness to the global minimum. This is a consequence of the reduction of search space for GA in a GALP context. There are three loops in the basic Hanoi network. This means that the chromosome for the Hanoi network consists of 6 genes. There are 14 possibilities for splitting the first

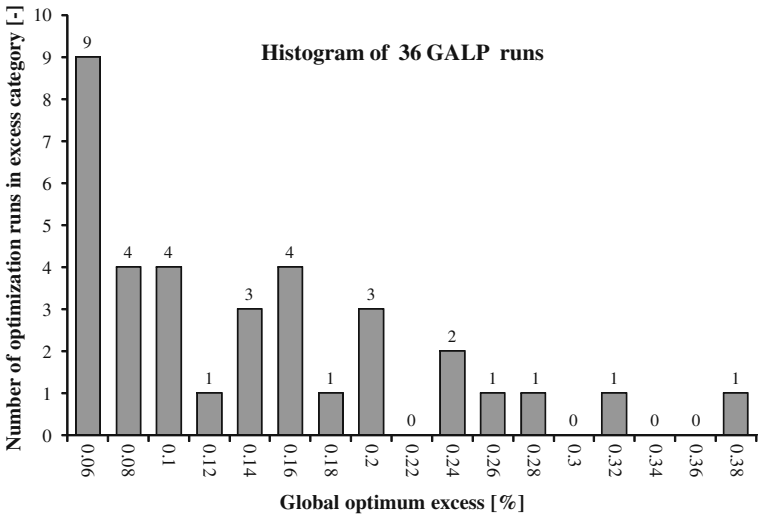

Fig. 7 Histogram of GALP runs of the triple Hanoi network with various settings. The difference between the particular solution obtained from the global optimum is displayed on the X axis, and the Y axis demonstrates how many tested runs have this difference (for example, there were nine runs which had a solution only 0–0.06% greater than the global minimum)

loop, eight for the second and seven for the third. The R_s factor used was set to be 0.1, so there are nine possibilities for splitting the demand in the split node in every loop (01; 0.2; …; 0.9). This means that there are $14 \times 8 \times 7 \times 9 \times 9 \times 9 = 571{,}536$ possibilities of which the search space of GALP for the basic Hanoi network consists. The search space of the GA for the same problem is $6^{34} = 2.86512\text{E}{+}26$, which is significantly greater. In the case of the triple Hanoi network there are $571.536^3 =$

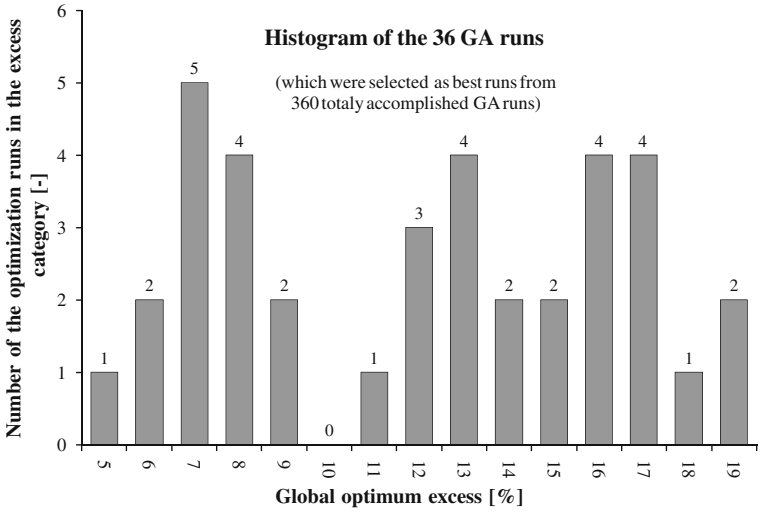

Fig. 8 Histogram of the GA runs of the triple Hanoi network with various settings reveals significantly greater global optimum excess than GALP

Table 3 Comparison of the best results when applying various methods

Method	Hanoi	Double Hanoi	Deviation from reference global optimum (%)	Triple Hanoi	Deviation from reference global optimum (%)
GALP	6,057,697[a]	12,073,039	0.04	18,394,255	0.04
GA	6,081,087	12,600,624	4.01	19,269,160	4.47
OptiDesigner	6,115,055	12,795,541	5.62	–[b]	–[b]
Harmony search	6,081,087[a]	12,404,680	2.39	18,839,302	2.67

[a]Reference global solution
[b]A feasible solution was not found

1.86694E+17 possibilities for GALP and $6^{100} = 6.53319E+77$ possibilities for GA. In this case it can be seen that a three times greater problem has a smaller search space when using GALP in comparison to when GA is used alone.

GALP is the method proposed by the author of this paper; references to other methods are GA (Cisty 2002), Harmony search (GEEM 2002) and OptiDesigner—http://www.optiwater.com/optidesigner.html. Other methods were also tested and compared to GALP, but they provide similar or worst results than those in the Table 3.

A typical run of the GALP optimization model can be seen in Fig. 9. It can be seen here that a good result was obtained by GALP already after 60 generations (in this particular run 100 chromosomes were in the population; crossover operator was taken as $p_c = 0.95$, and the mutation rate was set to be $p_m = 0.05$).

4 Summary and Conclusions

The design of an optimal water distribution network is a complex task. Various deterministic and heuristic algorithms have been proposed and attempted for solving this problem. Researchers have focused on stochastic or so-called heuristic optimization methods in the past two decades. Heuristic methods are used for solving a very general class of computational problems by combining various rules for finding a solution in a hopefully efficient way. Heuristic methods are generally applied to

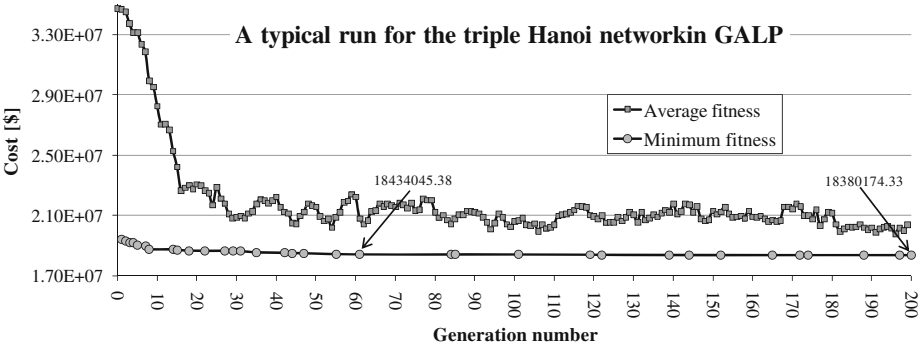

Fig. 9 A typical run of the triple Hanoi network in GALP

problems for which there is no satisfactory problem-specific algorithm or when it is not practical to implement such a method. The advantage of heuristic optimization techniques over the conventional optimization techniques is their robustness, speed, flexibility and ability to solve large and complex combinatorial problems. If, in the general optimization community, a heuristic algorithm already exists, its application for water distribution optimization usually involves only coding a chromosome and fitness function, and it is potentially applicable for the optimization of every parameter and type of network which could be run in a simulation program (e.g., EPANET). This and some other issues are the reason why they are so popular. Although the heuristic methods offer very good results from many points of view (and it is sure that they will be consecutively developed and used), they also have some disadvantages:

- They cannot guarantee the generation of a global or near global optimal solution, particularly for large-scale systems
- They require extensive fine-tuning of the algorithm's parameters, which are highly dependent on the individual problem
- They can produce solutions in which the diameters can be optimally found from an economic point of view, but the diameters can in some conditions be distributed randomly: for instance, for the branched part of networks, there could be larger diameters in some parts after the smaller ones in the direction of the flow, which is wrong.

For these reasons the author has proposed a method in which genetic algorithms (the heuristic method) are incorporated, but the final solution is produced by linear programming. This method is described in the paper and successfully tested on the benchmark networks. It was determined that the method gives results that are more trustworthy in terms of closeness to a global minimum. This is because of the fact that involving LP in an algorithm reduces the search space for heuristics very dramatically on most network configurations. Although one iteration of the proposed method is a little more computationally intensive than in the case of using GA only, this reduction is finally more important for the effectiveness of the algorithm.

This paper is a first effort to present a novel methodology, so only the basic task of optimizing water distribution systems is presented. Some suggestions for the scope of future research are therefore proposed in this paragraph. In this paper a method is applied, while a single demand loading is given as the input for optimization, but a solution for multi-demand loading is described in some remarks hereinabove. The effectiveness of the proposed concept should be tested, but the possibility of accomplishing this exists by adding the set of conditions described in Eq. 2 for every demand loading. In a practical design only two sets of demands are usually considered—peak demands and fire flow demands. Also, a solution for a split pipe design is preferred here, but a solution for designing only one pipe in a section is possible using a mixed integer programming definition of the LP problem, which is known. Split pipe design is often criticized in the water distribution systems optimization community because of the preconception that it can propose small sections with different diameters on a network—how to avoid this is described in Part 2.1 of this paper. Also, pump head sizing is possible (but should be tested) by the proposed GALP algorithm if the energy cost evaluation is known by adding the unknown variable for the head to Eq. 2 and the corresponding cost parameter to

the objective function (Eq. 3). The application of the method for the rehabilitation or enlargement of existing systems is just a matter of software implementation—for instance, when the rehabilitation of an existing network is to be solved, every first unknown X_{i1} in Eq. 1 should have a cost D_i in the objective function (Eq. 3) equal to zero (it is the cost of leaving the original pipe diameter in the section).

The method is proposed as an alternative to existing methods, mainly when a least-cost design is to be solved. Its extension to solving multi-objective tasks should be evaluated in future research. Nevertheless, the author supposes it should be tested too, if it is not better to use a more precise method for this single objective task and then refine the output from the least-cost design according to other criteria (or with other demand pattern) by applying a simulation model. The design is always multi objective, but that does not mean that this multi-objectivity must be covered in one computation with too many objectives (and less precision). That is why the author believes that this method can also be practically useful also in the stage described and tested in this paper.

In this work only the basic GA method was used in the stochastic part of the algorithm. It gives good results, but there is a possibility open to 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.

Acknowledgements This study was supported by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences, Grant No. 1/0496/08 and 1/0585/08 and by the Slovak Research and Development Agency, Grant No. APVV-0443-07.

References

Alperovits E, Shamir U (1977) Design of optimal water distribution systems. Water Resour Res AGU 13(6):885–900. doi:10.1029/WR013i006p00885

Cisty M (2002) Rehabilitation of irrigation pressurised pipe systems using optimisation techniques. J Land Water Dev, Pol Acad Sci 6(6):117–128

Cisty M, Savic DA, Walters GA (1999) Rehabilitation of pressurized pipe networks using genetic algorithms. In: Proc., water for agriculture in the next millennium. 17th congress on irrigation and drainage. International Commission on Irrigation and Drainage, Granada, pp 13–27

Cunha MC, Sousa J (2001) Hydraulic infrastructures design using simulated annealing. J Infrastruct Syst 7(1):32–39. doi:10.1061/(ASCE)1076-0342(2001)7:1(32)

Dandy GC, Simpson AR, Murphy LJ (1996) An improved genetic algorithm for pipe network optimization. Water Resour Res 32(2):449. doi:10.1029/95WR02917

Eiger G, Shamir U, Ben-Tal A (1994) Optimal design of water distribution networks. Water Resour Res 30(9):2637–2646. doi:10.1029/94WR00623

Ekinci O, Konak H (2009) An optimization strategy for water distribution networks. Water Resour Manage 23:169–185. doi:10.1007/s11269-008-9270-8

Eusuff MM, Lansey KE (2003) Optimization of water distribution network design using the shuffled frog leaping algorithm. J Water Resour Plan Manage 129(3):210–225. doi:10.1061/(ASCE)0733-9496(2003)129:3(210)

Fujiwara O, Khang DB (1990) A two-phase decomposition method for optimal design of looped water distribution networks. Water Resour Res 26(4):539–549

Geem ZW, Kim JH, Loganathan GV (2002) Harmony search optimization: application to pipe network design. Int J Model Simul 22(2):125–133

Goldberg DE (1989) Genetic algorithms in search, optimisation and machine learning. Addison-Wesley, New York

Kessler A, Shamir U (1989) Analysis of the linear programming gradient method for optimal design of water supply networks. Water Resour Res 25(7):1469–1480. doi:10.1029/WR025i007p01469

Loganathan GV, Greene JJ, Ahn TJ (1995) Design heuristic for globally minimum cost water-distribution systems. J Water Resour Plan Manage 121(2):182–192. doi:10.1061/(ASCE)0733-9496(1995)121:2(182)

Maier HR, Simpson AR, Foong WK, Phang KY, Seah HY, Tan CL (2001) Ant colony optimization for the design of water distribution systems. In: Proceedings of the world water and environmental resources congress, Orlando, FL

Olsson RJ, Kapelan Z, Savic DA (2008) Probabilistic building block identification for the multi-objective design and rehabilitation of water distribution systems. In: Proc., 10th annual water distribution systems analysis conference (WDSA 2008), Kruger National Park, South Africa

Raad D, Vuuren JV (2008) A multi-objective evolutionary meta-algorithmic approach towards water distribution system design. IFORS News, December 2008, pp 7–10

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

Savic DA, Walters GA (1997) Genetic Algorithms for least-cost design of water distribution networks. J Water Resour Plan Manage 123(2):67–77. doi:10.1061/(ASCE)0733-9496(1997)123:2(67)

Simpson AR, Dandy GC, Murphy LJ (1994) Genetic algorithms compared to other techniques for pipe optimization. J Water Resour Plan Manage 120(4):423–443. doi:10.1061/(ASCE)0733-9496(1994)120:4(423)

Vrugt JA, Robinson BA (2007) Improved evolutionary optimization from genetically adaptive multimethod search. Proc Natl Acad Sci U S A 104(3):708–711. doi:10.1073/pnas.0610471104

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 Water Resour Plan Manage 133(1):87–92. doi:10.1061/(ASCE)0733-9496(2007)133:1(87)

Zitzler E, Laumanns M, Bleuler S (2003) A tutorial on evolutionary multi-objective optimization. (Unpublished) Technical Report, Computer Engineering and Network Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Zurich