Applied Soft Computing 10 (2010) 1141–1150

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Computer aided optimization of natural gas pipe networks using genetic algorithm

Omar Fayez Mohamed El-Mahdy*, Mohamed Ezz Hassan Ahmed, Sayed Metwalli

Mechanical Design & Production Dept., Faculty of Engineering, Cairo University, Egypt

ARTICLE INFO

Article history:
Received 26 September 2008
Received in revised form 19 April 2010
Accepted 11 May 2010
Available online 17 June 2010

Keywords:
Optimization
Evolutionary computation
Genetic algorithm
Natural gas
Pipe networks
Soft constraints
Hard constraint
Computer aided optimization
Penalty function
Integer pipe diameter sizes

ABSTRACT

This study is concerned to determine the optimum pipe size for networks used in natural gas applications. The genetic algorithm has been used in optimizing network parameters. The topology of the network is predefined. The study deals with the discrete nature of decision variables, namely, pipe diameters, as they are usually available in market in standard sizes. Hard constraints and soft constraints are considered. An imposed penalty factor is introduced to allow solutions that violate soft constraints to remain in the population during the solution progress guiding the algorithm convergence to a minimum network cost.

In a case study, engineers with average experience of 6 years in the design office of a gas company performed the design of a gas network problem using their experience and judgment. The adopted method by engineers depends on a trial and error, time consuming, procedure. Their results are compared with the results obtained from the developed genetic algorithm optimization technique.

The developed optimization technique has provided a distinctive reduction in the total cost of pipe networks over the existing heuristic approach which is based on human experience and judgment. A saving up to 12.1% has been achieved using the present analysis, in the special case studied.

1. Introduction

The objective of this work is to find the optimum sizing of pipe diameters for a given natural gas network topology. The objective is to minimize the cost of realizing the network without any violation of the network constraints. There are two types of constraints, namely, hard constraints and soft constraints. Hard constraints are formulated within the mathematical model simulating the gas network and should not be violated for any degree. The global optimum solution is expected to exist on the border of the pressure constraint. Allowing, some infeasible solutions to be retained in the population data-structure may help to guide the search towards an optimum solution. An optimum solution is likely to occur due to application of genetic operators such as crossover and mutation over an infeasible solution.

Gas is extracted from production platforms at pressure of over 100 bar. Then, it is transported from shore to regions using the high national transmission system (NTS) at high pressure ranging from 70 bar to 35 bar. In the populated areas, gas pressure is then further reduced in the medium, intermediate and low-pressure systems to enable gas delivery for domestic use.

Most of domestic customers are supplied from the low-pressure system. The low-pressure system has as source district governors with an inlet pressure of 75 mbar. A given city is divided into districts. Each district is subdivided into sectors. A sector has at its source a number of district governors which are fed from the medium pressure system and has the domestic customers as loads.

Due to the load variation over the day, the estimation of the load is a statistical criterion. "Diversity" is the statistical peaks average per house within a group of houses. Diversity is the load that is taken into consideration when making the design of a network.

Pressure at internal nodes (soft type constraint) should be maintained above a minimum given acceptable pressure. For low pressure networks, the pressure at any building service should never be below 50 mbar while the minimum acceptable pressure at burner tip is about 20 mbar. For higher-pressure regimes, the minimum acceptable pressure varies from one application to another, depending on the designer point of view and pressure reduction station configuration.

In the present study, the topology of the network is predefined and the optimization search is oriented to find pipe diameters for minimum cost of the gas network. This situation is typical for gas network applications especially in distribution networks. As

* Corresponding author at: Mechanical Design & Production Dept., Faculty of Engineering, Cairo University, (200) Saker Koreish Buildings, New Maadi, Cairo 11435, Egypt. Tel.: +20 101184606.
E-mail address: omarfayez@hotmail.com (O.F.M. El-Mahdy).

1568-4946/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2010.05.010

Nomenclature

GA	genetic algorithm
GIS	Geographic Information System
PRS	pressure reduction station
NTS	National Transportation System
LNG	liquefied natural gas
HPSV	high pressure storage vessel
LPS	low-pressure system
LPG	liquefied petroleum gas
EP	evolutionary programming
M	number of demand nodes
R	number of source nodes
N	number of pipe in the gas network
l_j	length of pipe (j)
d_j	diameter of pipe (j) in mm
p_i	pressure at node (i) in bar
D_s	standard sizes available at market in mm
c_s	cost of size D_s currency/unit length
$\bar{A}$	node-arc incidence matrix
A	demand nodes incidence matrix
A_R	source nodes incidence matrix
a_{ij}	unit forming incidence matrix
Q_i	load vector at demand nodes in m^3/h
h_R	vector of gas pressure feeding the network at source nodes in bar
q_j	gas flow in pipe (j) in m^3/h
$\overline{p_i}$	pressure of gas at source node (i) in bar
p_i^{min}	permissible pressure at node (i) in bar
Cost(d)	cost of gas network as function of selected pipes diameter
P	gas pressure in bar
V	gas volume in m^3
T	temperature in K
R	universal gas constant
R'	specific gas constant
Z	compressibility factor
F	viscous force
μ	coefficient of dynamic viscosity
U	average gas velocity
ΔP	pressure drop in bar
Re	Reynolds number
P	gas density
N	kinematics viscosity
Q_s	equivalent flow rate at standard condition
Hf	head loss
f	general friction factor
f_{sp}	friction pipe for smooth pip flow
f_{rp}	friction pipe for rough flow
S	specific gravity
K	pipe resistance in a flow equation
T_s	temperature at standard condition
P_s	pressure at standard condition
$F(X)$	objective function
$f(x)$	fitness function
$P(X)$	fitness probability of chromosome X
$P_c(X)$	cumulative fitness probability of chromosome X
N	population size
NR	Newton Raphson method
LTM	linear theory method
S	number of available standard pipe diameter sizes
K	iteration number
w	number of nodes violation permissible pressure constraint

each gas branch is feeding a customer or a group of customers, the elimination of any branch is not an alternative in this search case. For each network: pipe length, type of material, related roughness and efficiency factor are given as inputs of the problem. Loads at demand nodes are given at the maximum loading condition in case of high-pressure systems. In low-pressure gas networks feeding domestic customers, the loads are based on a diversity factor rather than maximum connected loads. The flow equation adopted depends on the pressure regime [1,2]. The network analysis method differs from one simulation software to another [3]. Pipe diameters are selected from a set of discrete standard sizes that are available in the market. The adopted algorithm deals directly with this discrete nature, which is an advantage that was hardly achieved by other techniques.

The present optimization problem was tackled before using other optimization techniques, such as, linear programming, non-linear programming and partial enumeration methods as well as genetic algorithm [4]. None of these techniques had shown enough flexibility and strength to deal with real life problems. Gas companies adopted a heuristic technique which is based on human judgment and experience rather than known optimization techniques. The heuristic technique is based on a trial and error procedure using simulation softwares such as that given in Ref. [5]. Most of the studies in the literature were oriented to water supply networks [6] or limited to pipeline and tree gas networks [7]. Thus, the present study fills the gap of applying genetic algorithm on optimization of natural gas looped networks.

The present study does not focus on the choice of best equation that suits a certain regime. The present study does not focus neither on the comparison between different simulation methods nor on the accuracy and convergence of each one.

The only optimization variable in the optimization analysis is the diameter of each pipe branch in the network. Pipe diameters are selected from a set of discreet standard sizes that are available in the market. The adopted algorithm deals directly with this discrete nature, which is an advantage that was hardly achieved by other techniques as it will be described in the following sections.

The importance of pipe networks optimization is arising from the great number of applications it covers. Among these applications are applications associated with natural gas networks, oil transportation networks, water supply networks, water irrigation meshes and liquefied petroleum gas (LPG) networks. The principles governing different applications are the same. However, the choice of the best optimization technique and the best parameters setting should be carefully selected to suit each application. Few studies were found to deal directly with natural gas applications. Hansen [8] presented his work with application to a Danish gas network. The used optimization technique was applied on a case study representing a low-pressure natural gas network. The successive linear programming approach is used in an optimization method of the trust region type. Boyd [9] used genetic algorithm to optimize a real gas tree network problem that was constructed by British Gas. Wolf [10] solved a similar problem by application of the Bundle method for non-smooth optimization.

Zhang [7] applied his optimization work on two case studies representing high pressure networks where the inlet pressure at source node is a maximum pressure of 30 bar. The main feature of the proposed method is to treat the problem as a bilevel programming problem, then to simplify the lower level problem by conjugate duality theory and to handle the upper level problem by its piecewise linear and convex nature using a trust region algorithm. The lower level, concerned with the pressure and the flow in the network, used a sophisticated arithmetic method to transform the constrained problem to an unconstrained one then to solve it by Newton's method with some modifications.

Lin [11] applied an enumeration technique on the optimization of water pipe network. He introduced the enumeration technique to be a powerful technique to solve pipe network optimal design problem especially when reliability consideration are highlighted.

2. Genetic algorithm (GA)

Application of genetic algorithm optimization to real world problems, such as water supply networks, has emphasized two main features of great importance. The first, is the ability of genetic algorithm to deal directly with the discrete nature of decision variables (diameters of pipes) as they are available in the market only in a predefined standard sizes. Rounding-off a solution found by a traditional continuous variable-based technique towards the closer standard pipe size had shown to be of great difficulty [12]. The rounding procedure is an optimization problem in itself and the feasibility is not guaranteed. To guarantee the feasibility of the solution after the rounding process the only solution is to always round-up towards the higher standard diameter step. Sometimes, in doing so, the solution gets far away from the optimal solution and looses all the optimization effort acquired during the minimum cost search procedure [13].

Some studies were concerned with the comparison of different pipe networks optimization techniques. Dandy [14] revised four optimization techniques covering partial enumeration, nonlinear programming, linear programming and genetic algorithm. The risk of rejecting the optimal solution with the partial enumeration technique exists. In both linear and nonlinear programming, Dandy claimed that the rounding-off on used continuous pipe sizes is a difficult procedure with a risk to be stuck in a local optimum. Application of these techniques on a water networks showed that GA was the most capable technique to produce a minimum cost solution. Nonlinear method was able to achieve a cost value with is 4% greater than GA solution. Linear technique did not show any notable cost reduction when compared to GA.

Simpson and Dandy [4] compared three optimization techniques on a bench mark problem. The problem was solved using a pruned enumeration technique and a nonlinear optimization technique and genetic algorithm. GA was able to identify the global optimum 8 out of 10 times. Simpson and Dandy [17] compared results of various techniques on a case study that was solved and constructed based on Linear Programming Analysis. Results were compared in many studies using different techniques such as linear programming, gradient search, partial enumeration, linear programming with heuristics, and finally genetic algorithm as applied by Dandy. According to the author it was found that the power of GA is, not only achieving a cost reduction alternative, but also, handling the problem directly with discrete pipe sizes while keeping track on the most promising solutions.

Castillo [6] mathematically modeled the pipe network problem based on directed acyclic graphs. The problem was then presented as an optimization problem solved using genetic algorithm. The researcher did not study looped networks. Real coded chromosomes were used and a new set of problems specific genetic operators were introduced. “Individual Crossover” and “Linked Node Crossover” were introduced as two new crossover techniques.

Savic and Walters [15] described the method for coding genetic operators in order to preserve feasibility of pipe network solutions generated in any evolutionary programming (EP) run. A penalty term was embedded in the fitness function in order to reduce its strength relative to other feasible solutions.

Morley et al. [16] described architecture for an integrated optimization application, GAnet, which comprises a GA application, a GIS and a hydraulic network solver. The author made use of the advantage of multiprocessor machines together with the multiprocessing capabilities of Microsoft Windows NT. The software is tested in a case study from the regional Municipality of York where 300 new pipes are included in the model as decision variables. The GA based software achieves a solution, which makes a reduction saving of 35% when compared to the manually proposed heuristic solution.

Van Vuuren [13] applied his study on a segmented distribution pipeline. Genetic algorithm was adopted as a powerful technique. Single point crossover and uniform mutation were applied on a bit string representing chromosomes. The developed program keeps record of the best 20 found solutions for the choice of the decision maker.

A great advantage that characterized genetic algorithm is its ability to deal directly with the discrete nature of decision variables. Traditional optimization techniques did not have this ability and were obliged to use pipe diameters as continuous variable. Wolf [10] is an example of many works that go with this idea. The rounding step, that is to come after having an optimal continuous solution, has proved to be a hard optimization problem itself which is based on trial an error. Many research works have tried to find a way to come closer to the discrete handling of decision variables. Zhang [7] allows each pipe element to be subdivided into two smaller pipes. By this way, each pipe element is composed of two discreet pipe sizes and the new decision variable is the length percentage of each size.

Hansen [8] used the pipe diameter as a continuous variable in an iterative algorithm. In each iteration, the developed solution is approximated to the nearest discrete pipe sizes while maintaining the feasibility of the solution. Simpson and Dandy [14] in their comparison to other optimization techniques used pipe diameter as a continuous one. The solution of each technique was then rounded to the nearest discrete diameter size to allow the comparison with the genetic algorithm discrete solution.

A great advantage of GA, is that the routine of the search adopted by genetic algorithm makes it possible to keep track of the most promising solutions that were found during the search progress [12]. In many cases, some solutions, with a cost higher than that of the optimum solution, may have an advantage from the reliability point of view. The formulation of some considerations such as reliability, maintainability and ability to monitor the network performance are hard to be mathematically modeled. Keeping track of say the best 20 or 30 solutions give the designer flexibility to choose one of these solutions that may suit a specific consideration [17]. The fact that genetic algorithm search is dealing directly with the function cost of the problem rather than to implement different interrelated variables in the objective function gives the algorithm a great flexibility. The choice of applying some penalty factors to account for constraint violation has some advantages. Conventional optimization techniques do implement the constraints with the objective function in a mathematical sophisticated procedure.

The developed algorithm is composed of two modules. One module is responsible for simulating the proposed pipe combination, calculating the total network cost and allows for penalties to be imposed. The second module is used for applying the genetic search solely. Handling the problem this way makes it possible to solve real world problem. A variety of difficult modeling features could be introduced when adopting the penalty function. Variable loading condition, diameter constraints on some of the network branches, the choice of best location for source nodes, and variable source pressure are all examples of these features.

It should be mentioned that unless reliability and maintenance considerations were successfully mathematically modeled, GA simply automates the trial and error procedure that was carried by human experience. GA technique guides the computer to evaluate only potential solutions where a global optimum is expected in an automated structured search process.

3. Problem description

In the present study, the topology of the network is predefined. The problem is formed from $(m+r)$ nodes and (n) pipe branches. Nodes from $(1, \ldots, m)$ are demand nodes while nodes $(m+1, \ldots, m+r)$ are supply nodes (sources). Each pipe connects two nodes. For every pipe branch, l_j is the pipe length $(l_j > 0)$, q_j is the flow rate associated with pipe j (for $j = 1, \ldots n$).

A sign convention should be maintained for the definition of flow direction. In the present study, the flow is considered to be positive if the flow is in the direction from n_2 to n_1, where $n_2 > n_1$. For simplicity, it is recommended to start the node numbering from nodes away from the source, i.e. node (1) is the most distant node from the source.

A node-arc incidence matrix $(\bar{A})$ is constructed defining pipes connecting nodes. Where $\bar{A} \in M^{(m+r)xn}$.

$\bar{A}$ could be split into two matrices A and A_R. Where $A \in M^{mxn}$ & $A_R \in M^{rxn}$

a_{ij} is defining columns and rows of $\bar{A}$ as follows:

$$a_{ij} = \begin{cases} -1 & \text{if gas is entering the node} \\ +1 & \text{if gas is leaving the node} \\ 0 & \text{when no connecting pipe between node } i \text{ and } j, \end{cases}$$

For each row in matrix $\bar{A}$ only two entries (1 and −1) can be defined, while the remaining elements are zeros.

At node i $(i = 1, \ldots m+r)$, let p_i be the gas pressure $(p_i > 0)$, and Q_i be the demand flow rate $(Q_i > 0)$ if $(i \le m)$, or supply if $(i > m)$. Also, it is assumed that the capacity of a source is infinity.

The load vector defining loads at internal nodes (Q_i) constitutes the gas demand that should be supplied to costumer at node (i) and can be expressed as follows:

$$Q_i = [Q_1\, Q_2 \ldots Q_m]^T \quad i = 1, \ldots, m. \tag{1}$$

The source pressure vector (h_R) defining the pressure of gas when supplied to the network constitutes the maximum gas pressure that can take place at any point in the network:

$$h_R = [h_{m+1} \ldots h_{m+r}]^T \tag{2}$$

The pressure at source nodes is constant:

$$p_i = \bar{p}_i, \quad i = m+1, \ldots, m+r \tag{3}$$

The pipe diameter d_j selected for pipe (j) can be selected from a set of (s) discrete available pipe diameter sizes available at standard sizes in the market $(D_1 < D_2 < \ldots < D_s)$. Each pipe diameter has its own associated construction cost price $c_1 < c_2 < \ldots < c_s$ (unit in currency/meter):

$$d_j \in \left\{D_1, \ldots, D_s\right\} \tag{4}$$

$$c(d_j) \in \{c_1, \ldots, c_s\}. \tag{5}$$

4. Optimization problem formulation

The objective function is to minimize the total construction cost of this network. The cost as presented in the problem is an average cost where material, construction, administrative overhead and maintenance costs are all aggregated in a total cost per unit length. The cost values are calculated by networks construction companies by evaluating the total expenses of procuring pipe materials, fittings, valves, governors, constructing and maintaining the gas network over contract life time. The study did not evaluate the best method to asses these cost values and used the data available in the company historical data:

$$\text{minimize } f(x) = \text{Cost}(d)$$

where

$$\text{Cost}(d) = \sum_{j=1}^{n} l_j c(d_j) \tag{6}$$

The design variables of the problem are the pipes diameters for all the network branches:

$$d_j = [d_1\, d_2 \ldots d_n] \tag{7}$$

Two types of constraints are distinguished. Hard type constraints which must be satisfied and soft type constraints which could be violated but with a penalty cost added to the network construction cost.

Gas flow equation, Kirchoff first and second law and the discrete nature of pipe diameter are the hard constraints of the problem.

All gas flow equations will have the general form of Eq. (8) where the flow in a circular pipe is related to the associated pressure drop:

$$K l_j d_j^{\beta} |q_j|^{\alpha-1} q_j = \Delta y \quad j = 1, \ldots n \tag{8}$$

where $\Delta y = p_2^2 - p_1^2$, for medium and high pressure networks; $\Delta y = p_2 - p_1$, for low pressure networks; α, β and K are parameters that differ depending on the used flow equation.

Typically [1]:

$$-5.3 \le \beta \le -4.8$$
$$1.8 \le \alpha \le 2$$

K is the resistance of the pipe and depends on many factors including units to be used in the equation.

Depending on the pressure regime, Panhandel 'A' Equation was used as recommended by Osiadacz [1] for high pressure networks operating at above 7 bar gauge:

$$\Delta p = p_1^2 - p_2^2 = K Q^{1.854} \quad (\text{Panhandel '}A\text{' equation}) \tag{9}$$

where

$$K = 19.43 \left(\frac{L}{D^{4.854} E^2} \right)$$

E is the efficiency factor (=0.9); L is length in meters; D is the diameter in millimeters; p is the gas pressure in bar; Q is the gas flow rate in m^3/h.

The constraint is formulated as follows:

$$h_1(dj) = K l_j d_j^{\beta} |q_j|^{\alpha-1} q_j - \Delta y = 0 \quad j = 1, \ldots n \tag{10}$$

At each node the sum of gas entering the node must be balanced with the gas leaving the node:

$$Q_j = \sum_{i=1}^{n} a_{ij} q_i, \quad j = 1, \ldots, m \tag{11}$$

Which is simplified in a matrix form as follows:

$$Aq = Q \tag{12}$$

$$h_2(d_j) = Aq - Q \quad j = 1, \ldots, n \tag{13}$$

Kirchoff's second law is responsible to maintain a zero pressure drop over a closed network loop:

$$A^T P = \Delta P \tag{14}$$

The constraint is formulates as follows:

$$h_3(d_j) = A^T P - \Delta P \quad j = 1, \ldots, n \tag{15}$$

The pipe diameter d_j selected for pipe (j) is selected form a set of (s) discrete available pipe diameter sizes. $d_j \in \left\{D_1, \ldots, D_s\right\}$

$$h_4(d_j) = d_j \quad j = 1, \ldots, n \tag{16}$$

Permissible minimum pressure at demand nodes is the soft constraint where the pressure at demand nodes must be higher than a predefined minimum pressure:

$$p_i \geq p_i^{\min}, \quad i = 1, \ldots, m \tag{17}$$

$$g_1(d_i) = p_i - p_i^{\min} \geq 0, \quad i = 1, \ldots, m \tag{18}$$

The constraints defined by Eqs. (10), (13) and (15) are hard constraints which are formulated within the used network analysis method. This study adopted the iterative equation (19) which can be used for network simulation analysis [3]:

$$h_{k+1} = h_k - \gamma \left[A^T D_k^{-1} A\right]^{-1} \left[Q + A^T D_k^{-1}(A h_k + A_R h_R)\right] \tag{19}$$

where $\gamma = 1$ for LTM linear theory method; $\gamma = \alpha$ for NR Newton Raphson method; A is the pipes incidence matrix; A_R is the pipe source incidence matrix; h is internal node pressure vector; h_R is source pressure vector; Q is the internal nodes load vector; D_k is the diagonal matrix flow resistance.

The constraint defined by Eq. (18) is a soft constraint formulated within the developed genetic algorithm procedure where a penalty cost is imposed over the total network cost each time a pressure constraint is violated. The penalty cost is calculated to ensure that the violating solution cost is higher than the maximum possible cost of the network:

$$\text{Penalty} = (C_s - C_1) \sum_{j=1}^{n} l_j \tag{20}$$

where C_s and C_1 are the cost (unit in currency/meter) associated with the largest and smallest pipe diameter, respectively as defined by Eq. (5).

This penalty cost is applied each time a node violates the minimum permissible pressure constraint. Thus, the total cost of the network which will account for the fitness function calculation is defined by Eq. (21) as follows:

$$\text{Cost}(d) = \sum_{j=1}^{n} l_j c(d_j) + w(C_s - C_1) \sum_{j=1}^{n} l_j \tag{21}$$

where w is the total number of nodes violating constraint 18.

5. Genetic algorithm solution procedure

The first step of performing an optimization search using genetic algorithm is the formation of the first generation. It is formed by randomly selecting a group of solutions or alternatives for the problem under study.

The formed population is fed to the simulation module. Pressure at internal nodes, flow in each pipe element, network cost, and any additional penalty cost are calculated. A power factor called (fitness) is assigned to each alternative. This fitness is used to judge how much a solution is near the optimum solution. In a random controlled way (based on the selection of the fittest), this population is allowed to reproduce and generate a new population. The regeneration is achieved using many techniques such as the roulette wheel technique [18]. The new generation mainly has its characteristics inherited from the previous one. It is expected that the average fitness of this population and hence its ability to survive is improved compared to its parent. To assure that this improvement from a generation to another takes place, genetic operators are applied. After the population is coded in binary bits, simple crossover is applied over all the population. Then mutation is applied if required. While applying these genetic operators, the best chromosome is always kept away from deformation and is transformed to the next step safely. This is called the elitism technique. The mutation and crossover implement new solutions into the population and explore the working space of the problem. This procedure is repeated for a predefined number of generations set by the user.

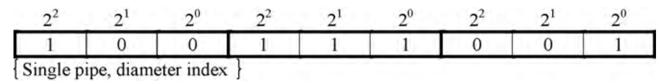

Fig. 1. Binary chromosome representation of a three pipe network.

5.1. Chromosome formation

The chromosome is chosen to deal directly with the pipe indices rather than to code the actual pipe diameter. This method enables to easily formulate the discrete nature of pipe diameters while having a small chromosome length. As shown in Fig. 1, the chromosome is a string of coded bits that represent a single solution of the network under study. The chromosome is divided into genes where each gene is the coded representation of the diameter of a single branch pipe of the network. The set of coded genes for all pipe diameters constitutes the chromosome which is a suggested pipe combination of the network under study.

Fig. 1 represents a chromosome for a three pipe branches network. The bits represent the binary code of the pipe index associated with each standard pipe diameter as presented in Table 1. As an example, using the coding system of Table 1, (d_1, d_2, d_3) of Fig. 1 will get a pipe diameter (180, 250HD, 63 mm) for pipes defined of the chromosome, respectively.

5.2. Fitness function

Fitness function is the representative of the objective function in the genetic algorithm. It is the way to rank chromosomes following their ability to achieve the objective function criteria. It is always recommended to deal with optimization problems as a maximization problem when using the genetic algorithm [18]. This implies that when the objective is to minimize a cost function the fitness function will be inversely proportional to the cost function. In many reviewed papers in literature, the fitness function was presented as the inverse of the objective function [4,14,15]. Dandy et al. [19] studied the use of the objective function in a more general form as follows:

$$f(X_k) = \left(\frac{1}{F(X_k)}\right)^n \tag{22}$$

All these forms of fitness function where tested during the preliminary phase of this work. However, it did not proof to have a significant effect on the algorithm effectiveness. The fitness function in the simplest form of Eq. (23) was applied on the case studied:

$$f(X_k) = \frac{1}{F(X_k)} \tag{23}$$

Table 1
Example of the binary code for commercially available standard pipe sizes.

Index	Binary code	Pipe diameter (mm)	Cost/unit length
1	000	32	100
2	001	63	140
3	010	90	170
4	011	125	250
5	100	180	375
6	101	250	580
7	110	180 (high density)	600
8	111	250 (high density)	700

5.3. Method of selecting parent chromosome

The selection operator is responsible for determining the parent chromosomes that will be allowed to mate and produce new solutions, hopefully with a better fitness function values. Weighted roulette Wheel selection method is used.

In this method, each chromosome X_k is assigned with a fitness value $f(X_k)$. Depending on the fitness, a probability value $P(X_k)$ for each chromosome is obtained as follows:

$$P(X_k) = \frac{f(X_k)}{\sum_{k=1}^{N} f(X_k)} \tag{24}$$

where N is the population size.

The whole population is sorted in an ascending order following the fitness probability. A cumulative probability is then calculated for each chromosome in this new sorted population. Now, for N times (population size) a random number (α) between [0 1] is generated. Each time, the chromosome with the nearest higher cumulative probability value is selected to jump to the next generation.

Let $P_C(X_k)$ be the cumulative probability assigned to chromosome (X_k).

When $P_C(X_k)\langle\alpha\langle P_C(X_{k+1})$

The chromosome (X_{k+1}) is selected as a parent in the temporary population that will be the basic pool for the formation of the next generation.

5.4. Crossover and mutation

The single point crossover and the uniform mutation techniques were adopted in the present study. Simpson et al. recommended the use of these techniques in problems associated with water pipe networks. The crossover probability rate between (0.6) and (1) was also recommended. A crossover probability of (0.7) was found to be very effective in this study.

In the single point crossover technique, a random break location is chosen in the chromosomes representing the two parents. At this break location, fragments of the parent chromosomes are swapped to create two new chromosomes.

The uniform mutation technique is a powerful source of diversity in the population. Each bit location through the chromosome length may be switched from (0) to (1) and vice versa on a probabilistic way.

Literally, a small mutation rate ranging between (0.001) and (0.01) was always adopted [4,14,15]. Haupt [20] investigated the use of a large mutation rate on an application of the genetic algorithm in the electromagnetic field problems. A suggested mutation rate for GA problems lied between (0.05) and (0.2). The effectiveness of using a large mutation rate was interpreted that, as the algorithm progresses, the probability to have children that are identical to the parent chromosomes magnifies. The use of a high mutation rate eliminates the stuck in a local minimum. This new trend of using high mutation rate was adopted in this study with a mutation rate of (0.3).

6. Developed software for genetic algorithm optimization

The process of selecting the proper values of genetic algorithm parameters and getting a human based solution to compare with requires many trials. For the purpose of facilitating this process, a software program is developed. The developed software named "GAGAS" compromises mainly two modules in a graphical user interface.

The first module is a simulation module. It is used to calculate the flow in each pipe element and the pressure at each internal node for a specific given pipe combinations. The second module is a genetic algorithm (GA) optimization module. A genetic algorithm library was constructed allowing the choice between some of the genetic operators for crossover, mutation and selection. Different genetic algorithm parameters are selected, namely, the population size, the number of generations, the crossover and mutation types and probabilities, fitness function equation and selection operators.

6.1. Program description

Fig. 2 shows a schematic representation of the program flow chart. The simulation module gets different pipe combination either manually by the designer or automatically when the GA module is activated. Following the embedded gas flow equations, the program gives the output flow rates and pressure relevant to the input pipe diameters combination. In the simulation mode, the output flow rate and pressure at internal nodes are displayed on the screen helping the designer to re-adjust his proposed pipe combination. If the GA module is activated the pressure is fed back to the GA module to check for pressure constraint violation. In case of pressure violation a penalty function is considered and is added to the network cost.

As the GA module is activated, it generates the first generation on a random basis. The population comprises a group of pipe combinations. This set of pipe combinations is delivered to the simulation module. Relevant pressure, network cost and penalties (if any) are calculated and fed back to the GA module. Then, the pipe combi-

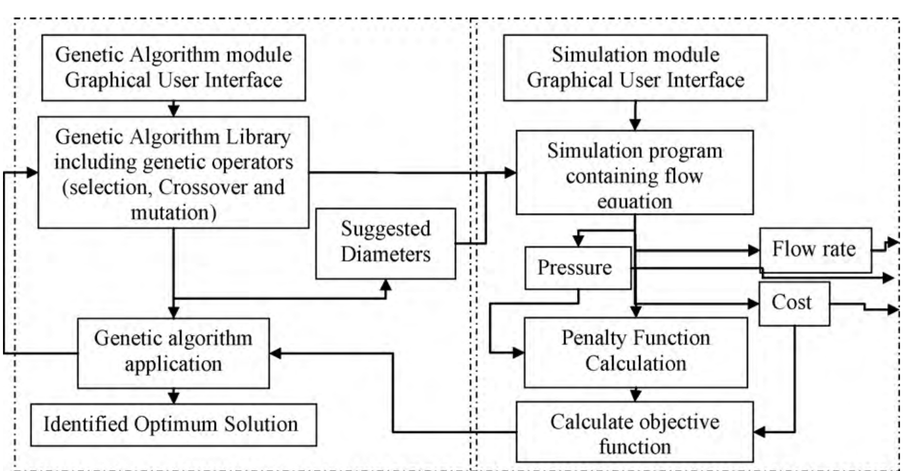

Fig. 2. Schematic for the developed software flow chart.

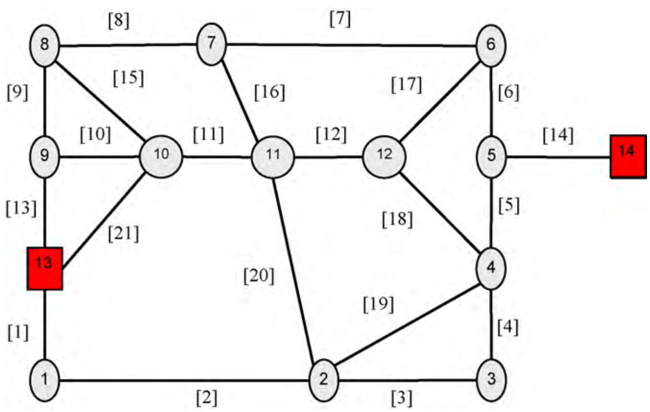

Fig. 3. Layout of the case study.

Table 2
Pipes length in (m) for the case study.

Pipe	Length
[1]	6300
[2]	8400
[3]	7700
[4]	9100
[5]	1400
[6]	6300
[7]	12,600
[8]	3850
[9]	4900
[10]	1400
[11]	4900
[12]	4200
[13]	8400
[14]	1750
[15]	7700
[16]	5600
[17]	14,700
[18]	17,500
[19]	10,500
[20]	15,400
[21]	9100

nations are binary coded. The fitness function is calculated based on the network cost and any additional penalties. Then the coded population is subject to the application of selection, crossover and mutation operators. The result of this process leads to an entirely new generation that is enhanced in its genetic characteristic. The cycle is repeated for a defined number of generations. In each generation, a copy of the best solution with a minimum cost is saved in a separate memory location to enlarge the final choice of the designer within a number of suggested good solutions.

7. Results

The results from the developed genetic algorithm application are compared to a heuristic approach which is based on human experience. In this heuristic method, the designer (with average 6 years experience) suggests preliminary pipe combinations. Using simulation software, the suggested pipe sizes are analyzed to determine the flow in each pipe and the pressure at each internal node. Based on the analysis output from the simulation software, the designer starts to modify his suggested design. The procedure is iterative and it never guarantees to achieve an optimum solution. Nevertheless, the success of each heuristic approach depends solely on the designer experience and on the first preliminary suggestion.

A case study is given as a challenge to the design office (with average 6 years experience) in a gas company, in order to obtain a comparative heuristic solution. The network resembles a typical gas network feeding an industrial area supplied directly from the medium pressure network.

Table 4
Commercially available standard pipe sizes in the case study.

Index	Diameter (mm)	Cost ($/m)
1	100	1637
2	150	1796
3	200	2122
4	250	2908
5	300	2940
6	400	4139

7.1. Case study description

The case study topology is shown in Fig. 3. The network is formed from 21 pipes, 12 internal nodes and 2 sources of pressure. The problem consists of 8 internal loops. Pipe lengths are given in Table 2. The maximum pressure which is fed to the network is 17.5 bar and the minimum allowed pressure at any internal node is 2.5 bar. Gas demand at each internal node is as given in Table 3.

Pipe diameters are selected from 6 different pipe sizes available. Table 4 shows the details of each available pipe size with the diameter and cost per meter.

7.2. Results of heuristic approach

The problem is given to the designers in a gas company with no limited time frame. Pressure constraint violation, with any degree, is not tolerated neither for the engineers nor for the developed algorithm.

Table 5 shows submitted solutions by the engineers. Each solution is identified by the index of selected pipe combination, internal diameter, the total cost of the network and the minimum pressure at demand nodes. An average computation time found by engineers to submit one solution is from 3 to 4 h.

7.3. Results of genetic algorithm

Genetic algorithm is applied on the case study. A random population size of (250) is selected and allowed to reproduce for (250) generations. Crossover probability of (0.7) and a mutation rate of (0.3) are adopted. Elitism strategy was applied.

Table 6 gives the results for 10 different single algorithm computer runs. Each solution is identified by the index of selected pipe combination, the total cost of the network and the minimum pressure at demand nodes. Each solution cost is compared with three cost values, namely, minimum, maximum and the average of solutions submitted by engineers as shown in Table 5. The columns Red.(1), Red.(2) and Red.(3) in Table 6 represent the percent of cost reduction of each solution when, respectively, compared to the minimum, average and maximum of suggested solution by engineers.

Table 3
Demand flow rate at nodes of the case study.

Node	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Demand flow rate (m^3/h)	11,500	11,500	11,500	11,500	0	11,500	8750	8750	0	8750	8750	12,500

Table 5
Results of the heuristic approach of the case study.

	Engineer (A)		Engineer (B)		Engineer (C)	
	Index	*d*	Index	*d*	Index	*d*
Pipe						
1	3	200	2	150	3	200
2	2	150	4	200	1	100
3	2	150	3	200	2	150
4	2	150	3	200	3	200
5	2	150	3	200	3	200
6	2	150	3	200	3	200
7	2	150	3	200	1	100
8	2	150	2	150	1	100
9	2	150	2	150	3	200
10	3	200	2	150	1	100
11	2	150	1	100	3	200
12	1	100	1	100	3	200
13	4	250	3	200	3	200
14	4	250	3	200	3	200
15	1	100	1	100	1	100
16	1	100	2	150	1	100
17	2	150	3	200	1	100
18	2	150	3	200	3	200
19	1	100	3	200	1	100
20	1	100	3	200	1	100
21	3	200	2	150	3	200
Min. pressure	2.8497		2.6		4.8	
Cost	300,276,200		324,824,500		301,744,450	

Figs. 4–6 show the progress of the genetic algorithm solution for three different single computer runs as a function of the number of generations. In Fig. 4 the algorithm find a solution with a cost 3.52% less than the minimum cost obtained by engineers, while in Fig. 5 the algorithm was able to find a solution only 0.9% less than the minimum cost obtained by engineers. Fig. 6 shows a case where the algorithm found a solution higher than the minimum suggested by engineers with a slightly increased cost of 1.13% but still the cost is lower than both the average and the maximum solution obtained by engineers.

Fig. 7 represents a summary plot of 100 computer runs of the algorithm. Results from the developed genetic algorithm applications are plotted in comparison with minimum, maximum and average of suggested solution by engineers. Table 7 summarizes results of Fig. 7 where the number (out of 100) the algorithm reached a solution less than respectively the minimum, average and the maximum of engineers identified solution is presented. During 100 consecutive runs the algorithm identified 100 times a cost that is less than the maximum and average solution submitted by engineers. While an 85 times the cost achieved was less than the minimum solution proposed by engineers.

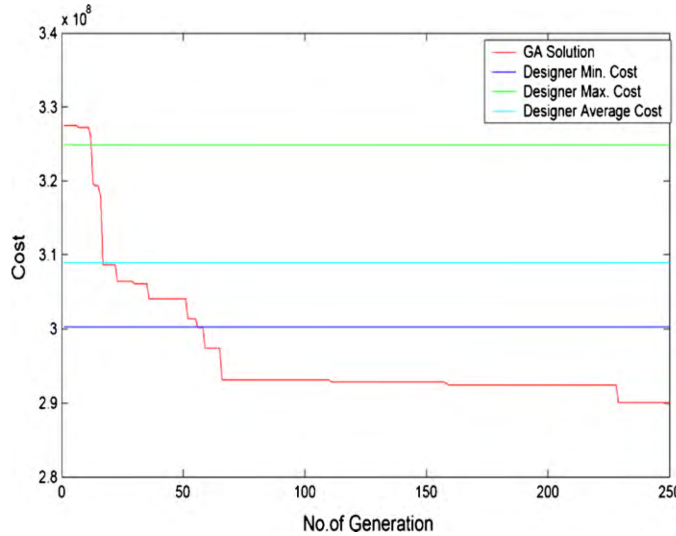

Fig. 4. Progress of the cost function as a function of the number of generations compared with engineers' solutions.

8. Discussion

Results given in Table 6 as well as results in Figs. 4–7 show that the developed algorithm is capable of identifying solutions with a cost saving up to 4% when compared with minimum network cost obtained by engineers. When comparing results with the average

Table 6
Results of 10 GA application for the case study.

	Pipe																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
GA(1)	2	1	2	1	4	3	1	1	4	2	2	1	4	5	1	1	2	1	3	1	1
GA(2)	2	1	3	2	3	3	1	2	1	1	2	3	2	5	2	1	2	1	2	1	3
GA(3)	2	2	1	2	5	3	2	1	1	1	1	2	1	5	3	1	1	3	2	1	2
GA(4)	3	1	1	3	4	3	2	2	1	1	1	1	1	5	3	1	2	2	1	1	3
GA(5)	2	2	1	3	3	2	1	3	2	2	1	1	4	5	2	1	2	2	1	1	2
GA(6)	2	2	1	3	4	3	2	2	1	1	1	1	2	4	2	1	2	2	1	1	3
GA(7)	2	1	2	2	4	2	1	2	1	4	2	1	3	5	2	3	2	2	2	1	2
GA(8)	3	2	1	2	3	3	1	2	1	2	2	2	3	5	1	2	2	2	1	2	2
GA(9)	3	3	3	1	2	2	1	2	2	3	3	3	5	4	2	1	1	1	1	1	1
GA(10)	3	3	1	2	2	3	1	1	2	2	2	1	3	5	2	1	2	1	1	1	2

	Min. pressure (bar)	Network cost (HK$)	Percentage of cost reduction over heuristic approach		
			Red.(1) Min. cost	Red.(2) Avg. cost	Red.(3) Max. cost
GA(1)	6.2	299,379,850	0.3	3.2	8.5
GA(2)	3	291,309,200	3.1	6.1	11.5
GA(3)	3.9	293,656,650	2.3	5.2	10.6
GA(4)	4.9	295,170,400	1.7	4.7	10.0
GA(5)	2.9	296,751,000	1.2	4.1	9.5
GA(6)	3.2	293,221,600	2.4	5.4	10.8
GA(7)	4.5	292,858,300	2.5	5.5	10.9
GA(8)	4.2	292,817,000	2.5	5.5	10.9
GA(9)	6.8	297,668,700	0.9	3.8	9.1
GA(10)	2.8	289,700,950	3.7	6.6	12.1
Maximum percentage reduction			3.7	6.6	12.1

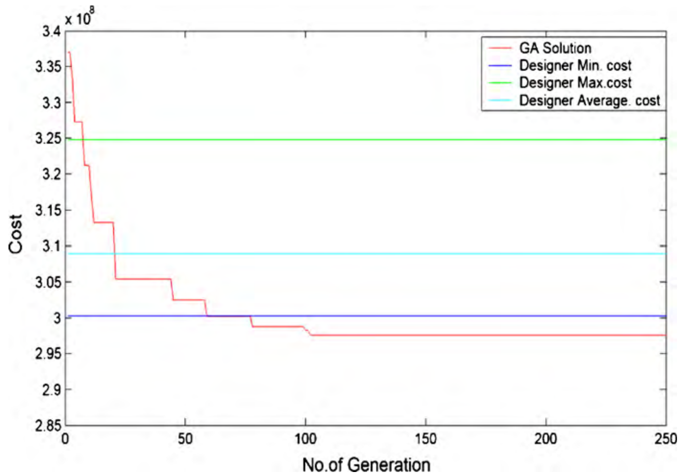

Fig. 5. Progress of the cost function as a function of the number of generations compared with engineers' solutions.

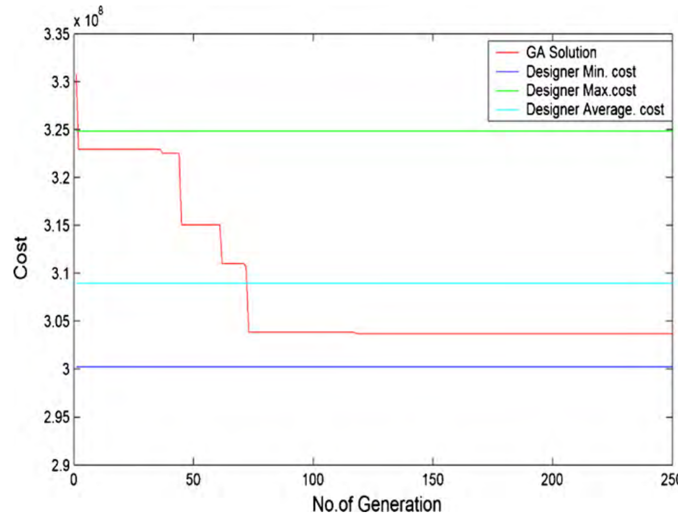

Fig. 6. Progress of the cost function as a function of the number of generations compared with engineers' solutions.

Table 7
Summary of the results given in Fig. 7.

	Results from Fig. 7
Number of Max. heuristic cost achieved out of 100 runs	100
Number of average heuristic cost achieved out of 100 runs	100
Number of Min. heuristic cost achieved out of 100 runs	84

and the maximum cost obtained by all engineers, the cost reduction is increased to 12.1%.

Achieving a global optimum using the algorithm is not guaranteed however a good solution with a considerable cost reduction is always obtained.

It could be seen that, as the complexity of the problem increases, the ability of the heuristic human based approach to identify a good solution is less effective. A single computer run of the developed algorithm takes around 8 min on an IBM ThinkPad Pentium IV processor, compared with a procedure that takes not less than 4 h by an average engineer.

9. Conclusions

1. The application of genetic algorithm to the optimization of looped natural gas pipe networks proved to be an efficient method for determining the optimum pipe sizing to minimize the network cost.
2. The nature of the developed algorithm makes it possible, not only to identify an optimum solution, but also to enumerate the most promising solutions that may exist spread in the workspace. These promising solutions may have a more expensive cost than the identified optimum solution but they may have another advantage from the reliability or maintainability point of view.
3. The comparison of results with the heuristic human based approach gives surprising pipe combinations that has a reduced

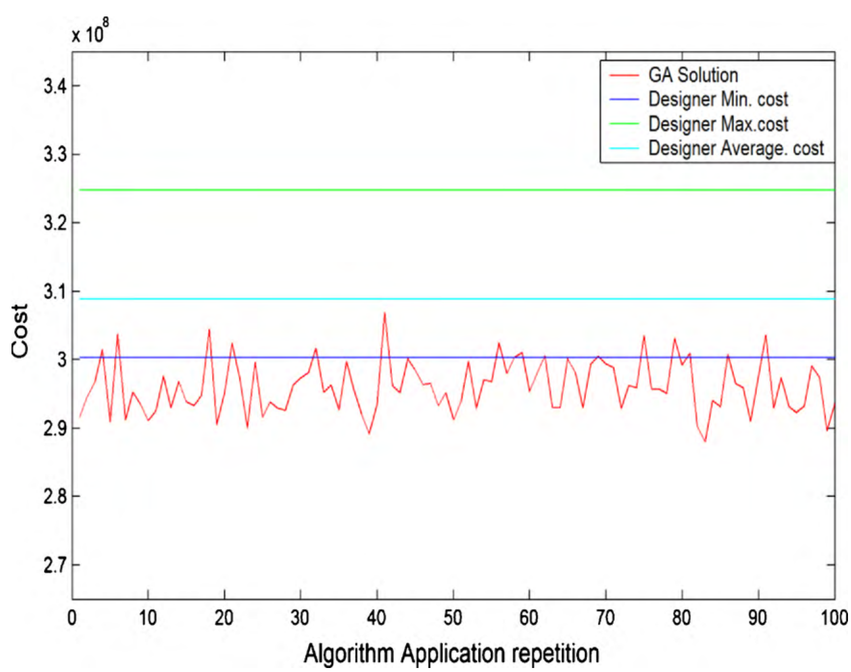

Fig. 7. Results of 100 consecutive computer runs of the algorithm application compared with engineers' solutions.

network cost price and were far away of designer suggestions. Application of the algorithm reveals the facts that sometimes using the highest available diameter at some branches makes it possible to use the minimum available diameters at other branches and that the optimum solution may exist within this new pipe combinations.
4. The application of the genetic algorithm is not suggested to replace the human judgment entirely, but its great importance is to replace the time consuming trial and error procedure previously adopted by designers and engineers. The developed algorithm provides the decision maker with a set of promising solutions that are to be judged and verified.
5. It is clear that as the complexity of the network increases, the ability of the algorithm to achieve a more competitive cost reduction compared to human based approach increases.

10. Recommendation

1. It is suggested to have more study consideration for the use of other advanced GA based models and to choose the most suitable model to this field of application [21,22].
2. It is suggested for future work to consider integrating the algorithm with simulation GIS based software and multiprocessor computer to extend the study to more complicated problems that involve higher number of pipe elements and to implement [26].
3. The use of multi-objective techniques as well as hybrid systems could be useful to implement some consideration related to the reliability and maintainability of the network [23–25].
4. The method used of applying a penalty function to characterize solutions violating the problem constraints gives the algorithm a great flexibility. It is possible, by fine-tuning of the penalty function cost factor, to give opportunity to solutions with slight pressure violation to be preserved as a promising solution for the decision maker consideration.

References

[1] A.J. Osiadacz, K. Pienkosz, Methods of steady-state simulation for gas networks, International Journal of Systems Science 19 (7) (1989) 1311–1321.
[2] N. El-Emam, A. Hassan, Most accurate equation and best reinforcement method in intermediate pressure natural gas network, in: Proceeding of the 1st International Gas Conference Al-Baath University-Homs, 2005, pp. 11–23.
[3] H.B. Nilsen, Methods for analyzing pipe networks, Journal of Hydraulic Engineering 115 (February (2)) (1989) 139–157.
[4] A. Simpson, G. Dandy, M. Laurey, Genetic algorithm compared to other techniques for pipe optimization, Water Resources Planning and Management, ASCE 120 (August (4)) (1994) 423–443.
[5] Stoner Associates, Inc., Technical Data Book SynerGEE Gas, Version 3.2, 2000.
[6] L. Castillo, A. Gonzalez, Distribution network optimization: finding the most economic solution by using genetic algorithms, European Journal of Operational Research 108 (3) (1998) 527–537.
[7] J. Zhang, D. Zhu, A bilevel programming method for pipe network optimization, SIAM Journal of Optimization 6 (August (3)) (1996) 838–857.
[8] C.T. Hansen, K. Madsen, H.B. Nilsen, Optimization of large networks, Mathematical Programming 52 (1991) 45–58.
[9] I.D. Boyd, P.D. Surry, N.J. Radcliffe, Constrained Gas Network Pipe Sizing with Genetic Algorithm, Technical Report EPCC-TR94, Edinburgh Parallel Computing Center, 1994.
[10] D.D. Wolf, Y. Smeers, Optimal dimensioning of pipe networks with application to gas transmission networks, Operation Research 44 (July–August (4)) (1996).
[11] B.L. Lin, H.M. Shau, W.C. Huang, R.S. Wu, S.L. Liaw, The enumeration algorithm for the practical optimal design of pipe network systems, Environmental Informatics Archives 2 (2004) 87–98.
[12] A.J. Abebe, D.P. Dolomatine, Application of global optimization to the design of pipe networks, in: Proceeding of 3rd International Conference on Hydro Informatics, Copenhagen, August, 1998.
[13] S.J. Van Vuuren, Application of genetic algorithms—determination of the optimal pipe diameters, Water SA 28 (April (2)) (2002) 217–226.
[14] G.C. Dandy, A.R. Simpson, L.J. Murphy, A Review of Pipe Network Optimisation Techniques, National Conference Publication Institution of Engineers, Australia, 1993, n 93 pt 2.
[15] D.A. Savic, G.A. Walters, Genetic operators and constraint handling for pipe network optimization, in: T.C. Fogarty (Ed.), Lecture Notes in Computer Science 993, Evolutionary Computing, Springer-Verlag, 1995, pp. 154–165.
[16] M.S. Morley, R.M. Atkinson, D.A. Savic, G.A. Walters, GAnet: genetic algorithm platform for pipe network optimization, Advances in Engineering Software 32 (June (6)) (2001) 467–475.
[17] J. Frey, A. Simpson, G. Dandy, M. Laurey, F. Terry, Genetic algorithm pipe network optimization: the next generation in distribution system analysis, Public-Works 127 (June (7)) (1996) 39–42.
[18] T. Back, D.B. Fogl, Z. Michalwicz, Handbook of Evolutionary Computation, IOP Publishing Ltd./Oxford University Press, 1997.
[19] A. Simpson, G. Dandy, M. Laurey, F. Terry, Pipe network optimization using genetic algorithms, in: Water Resources Planning and Management and Urban Water Resources, ASCE, New York, USA, August 1993, pp. 392–395.
[20] R.L. Haupt, Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors, in: Antennas and Propagation Society International Symposium, vol. 2, 2000, pp. 1034–1037.
[21] H.A. Hegazi, A.O. Nassef, S.M. Metwalli, Shape optimization of NURBS modeled 3D C-frames using hybrid genetic algorithm, in: Proceedings of the 2002 ASME Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Montreal, Canada, September 29–October 2, 2002, Paper # DETC2002/DAC-34107.
[22] A.O. Nassef, H.A. Hegazi, S.M. Metwalli, Design of C-frame using real-coded genetic optimization algorithm and Nurbs, in: Proceedings of the 1999 ASME Design Engineering Technical Conference, Las Vegas, Nevada, September 12–15, 1999, Paper No. DETC99/CIE-9138.
[23] A.V. Babayan, D.A. Savic, G.A. Walters, Z. Kapelan, Robust least-cost design of water distribution networks using redundancy and integrated based methodologies, Journal of Water Resources Planning and Management, ASCE 131 (1) (2007) 67–77.
[24] K. Behzadian, Z. Kapelan, D.A. Savic, A. Ardeshir, Stochastic sampling design using multiobjective genetic algorithm and adaptive neural networks, Environmental Modeling & Software 24 (4) (2009) 530–541.
[25] F. di Pierro, S.T. Khu, D.A. Savic, L. Berardi, Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms, Environmental Modeling & Software 24 (2) (2009) 202–213.
[26] R. Farmani, G.A. Walters, D.A. Savic, Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality, Journal of Hydroinformatics 8 (3) (2006) 165–179.