Environmental Modelling & Software xxx (2014) 1–11

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Embedding linear programming in multi objective genetic algorithms for reducing the size of the search space with application to leakage minimization in water distribution networks☆

E. Creaco [a,*], G. Pezzinga [b]

[a] College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, UK
[b] Dipartimento di Ingegneria Civile e Architettura, University of Catania, Viale A. Doria 6, 95125 Catania, Italy

A R T I C L E I N F O

Article history:
Received 12 March 2014
Received in revised form
25 September 2014
Accepted 6 October 2014
Available online xxx

Keywords:
Multi-objective optimization
Genetic algorithms
Linear programming
Water distribution
Leakage
Pressure
Control valve

A B S T R A C T

This paper shows how embedding a local search algorithm, such as the iterated linear programming (LP), in the multi-objective genetic algorithms (MOGAs) can lead to a reduction in the search space and then to the improvement of the computational efficiency of the MOGAs. In fact, when the optimization problem features both continuous real variables and discrete integer variables, the search space can be subdivided into two sub-spaces, related to the two kinds of variables respectively. The problem can then be structured in such a way that MOGAs can be used for the search within the sub-space of the discrete integer variables. For each solution proposed by the MOGAs, the iterated LP can be used for the search within the sub-space of the continuous real variables. An example of this hybrid algorithm is provided herein as far as water distribution networks are concerned. In particular, the problem of the optimal location of control valves for leakage attenuation is considered. In this framework, the MOGA NSGAII is used to search for the optimal valve locations and for the identification of the isolation valves which have to be closed in the network in order to improve the effectiveness of the control valves whereas the iterated linear programming is used to search for the optimal settings of the control valves. The application to two case studies clearly proves the reduction in the MOGA search space size to render the hybrid algorithm more efficient than the MOGA without iterated linear programming embedded.

1. Introduction

The reduction in the size of the search space is an important issue when applying evolutionary algorithms to real-world problems (Maier et al., 2014). In fact, the large size of the search space may sometimes lead population based algorithms to have difficulties finding solutions close to the global optimum. In this work, the reduction in size of the search space of a multi-objective genetic algorithm (MOGA) is obtained by embedding an iterated linear programming (LP) algorithm in it, and then by splitting the whole space into two parts, assigned to either of the algorithms. The resulting algorithm belongs to the category of the hybrid algorithms, which have been gaining popularity in the Hydroinformatics Community in the last two decades (Reis et al., 1997; Pezzinga and Gueli, 1999; Jourdan et al., 2004; Pezzinga and Pititto, 2005; Olsson et al., 2009; Raad et al., 2009; Haghighi et al., 2011; Creaco and Franchini, 2012, 2013; Wang et al., 2014; Creaco and Pezzinga, 2014), and whose structure is based on the combination of various sub-algorithms, which collaborate with each other in order to obtain improved computational efficiency. In particular, following the Talbi (2002) classification, the resulting algorithm can be considered as a Low Level Hybrid Algorithm (LLHA), since one of the component algorithms (iterated LP) is embedded in the other (the MOGA) as a functional part.

Among the LLHAs, there exist various examples in the water resources scientific literature featuring the use of a local search algorithm embedded in an evolutionary algorithm (see for example Reis et al., 1997; Pezzinga and Pititto, 2005; Haghighi et al., 2011; Creaco and Franchini, 2012, 2013; Creaco and Pezzinga, 2014). This approach turns out to be particularly fruitful from the viewpoint of computational efficiency when the search space features continuous real variables and discrete integer variables at the same time. In this case, the search space can be subdivided into two subspaces, related to the two kinds of variables respectively. In the context of the optimization, the evolutionary algorithm can be

☆ Thematic Issue on Evolutionary Algorithms.

* Corresponding author.
E-mail addresses: E.F.Creaco@exeter.ac.uk (E. Creaco), gpezzinga@dica.unict.it (G. Pezzinga).

http://dx.doi.org/10.1016/j.envsoft.2014.10.013
1364-8152/

applied to the sub-space of the discrete integer variables. The local search algorithm, instead, can be used for the sub-space of the continuous real variables, if this space is convex. A worthwhile example of the search space splitting was provided by Reis et al. (1997), who proposed an LLHA aimed at identifying the optimal location of control valves for leakage attenuation in water distribution networks. This algorithm was based on the coupling of a GA and of an embedded iterated LP algorithm for the search for optimal control valve locations (discrete integer variables) and control valve optimal settings (continuous real variables) respectively. In this LLHA, the problem was formulated as single objective and the objective function was the daily leakage volume to be minimized in the network. Later, Pezzinga and Pititto (2005) proposed an LLHA based on the same rationale, in order to deal with the problem of the combined optimization of pipes and control valves, always using the single objective formulation. To this end, they upgraded the LLHA of Reis et al. (1997) in order to render the GA capable of searching for the suitable diameters for the various network pipes besides the optimal location of the control valves. The objective function was the total installation cost, considered as the sum of the cost of the network pipes and of the present worth cost of the leakage volume distributed over the whole network life cycle.

In the context of the water distribution network design, other authors showed that LLHAs (made up of the combination of a GA and of a local search algorithm) can sometimes be constructed by assigning a sub-space of discrete integer variables (i.e. commercial pipe diameters) to the local search algorithm. In this case, an expedient consists in either using a local search algorithm for integer variables (such as the Integer Linear Programming ILP, see Haghighi et al., 2011) or in keeping on using a local search algorithm for continuous real variables (such as the LP, see Creaco and Franchini, 2012, 2013) and then rounding off the results after each algorithm application. Creaco and Franchini (2012, 2013) also showed that, in the context of the LLHAs, the sub-space assigned to the GA can also comprise continuous real variables.

The aim of this paper is to show how an LLHA can be constructed in the first case illustrated above (sub-space of discrete integer variables assigned to the GA and sub-space of continuous real variables assigned to the local search algorithm). The explicative example considered concerns the optimal location of control valves which have to be installed in order to attenuate leakage in water distribution networks. Compared to the LLHA of Reis et al. (1997) and Pezzinga and Pititto (2005), the new LLHA presented in this paper has the novelty of being developed following the multi-objective formulation, which is more widely preferred by decision makers since it enables results to be presented as Pareto front of optimal trade-off solutions between multiple objectives. The further advantage in comparison with the previous works in the same topic (Reis et al., 1997; Vairavamoorthy and Lumbers, 1998; Pezzinga and Gueli, 1999; Pezzinga and Pititto, 2005; Araujo et al., 2006; Liberatore and Sechi, 2009; Nicolini and Zovatto, 2009; Nicolini et al., 2011; Ali, 2014) lies in the fact that, thanks to a suitable and novel encoding of individual genes, the algorithm also enables identification of the isolation valves which have to be closed in order to contribute to leakage attenuation as well as to improve the effectiveness of the control valves installed by avoiding the formation of water paths around the control valves.

The problem of the optimal location of control valves and of the closure of suitable isolation valves can be framed within the more general paradigm of "divide and conquer", that is the partitioning and sectorization of a water network, recently addressed by some authors (Di Nardo et al., 2013a,b; 2014; Alvisi and Franchini, 2013). In particular, the algorithm proposed in this paper can be applied after a network has been subdivided into segments, in order to assess the suitable locations of the control valves to be installed and of the isolation valves to be closed.

In the following sections, a description of the structure of the new LLHA is given. Then, the application to a real case study follows, along with the comparison of the LLHA with a fully genetic algorithm, with no iterated LP embedded, in terms of computational efficiency.

2. Methodology

The proposed algorithm enables optimal location of the control valves as well as identification of the isolation valves which have to be closed, with the objective to simultaneously minimize the daily leakage volumes and installation costs. In this framework, the space of decisional variables comprises:

1. positions of the control valves to be installed and of the isolation valves to be closed (discrete integer variables);
2. settings of the control valves in the various time steps which characterize the network daily operation (continuous real variables).

The search space of decisional variables is subdivided between the two component algorithms. In particular, sub-space 1 is assigned to the multi-objective GA whereas sub-space 2 is assigned to the iterated LP. In the algorithm, the GA plays the role of main component algorithm. The other component algorithm, the iterated LP, is embedded in the GA as a functional part.

In the following sub-sections, first the iterated LP algorithm is described, followed by the definition of the steps to follow for the iterated LP solution; the description of the GA then follows.

2.1. Linear programming algorithm for the optimization of control valve settings

Within the multi-objective GA, which will be described in Section 2.3, an inner optimization algorithm based on the iterated LP is used in order to search for the optimal settings of the control valves in the jth of the $n_{\Delta t}$ time steps which characterize the network operation (in terms of hourly variations in source head and demand coefficient). In order to apply the iterated LP, the objective functions and constraints of the problem have to be linearized. In the following paragraphs, first the objective functions and constraints will be presented and then their linearized form will be derived. The linearization process differs from that originally proposed by Jowitt and Xu (1990), in that it was developed in matrix form, which is independent from the formula chosen to express pipe head losses.

The objective of the optimization is to minimize the leakage volume $W_{L,j}$ at each time step, calculated by means of the following relationship:

$$W_{L,j} = \sum_{i=1}^{n_p} Q_{Li}\Delta t, \tag{1}$$

where Q_{Li} is the leakage outflow from the ith of the n_p network pipes during the jth time step and Δt is the length of the time step.

Within the optimization problem, control valve settings can be defined as coefficients V, which take on real positive values lower than 1, by means of which the resistance of the generic pipe fitted with control valve is modified in such a way as to take account of the presence of the valve. The vector $\mathbf{V} = (V_1, V_2, \ldots, V_k, \ldots, V_{n_v})^T$ of the control valve settings can then be defined, where n_v is the number of pipes fitted with control valve and symbol "T" indicates the transpose vector.

Within the objective function in Eq. (1), the leakage outflow from the generic pipe can be assessed through the following relationship (Jowitt and Xu, 1990):

$$Q_{Li} = C_{L,i} L_i h_i^{n_{leak}}, \tag{2}$$

where for the ith pipe, h_i and L_i are the mean pressure head and the length respectively. $C_{L,i}$ and n_{leak} are empirical coefficients. The mean pressure head h_i can be calculated as:

$$h_i = \frac{H_{i,1} + H_{i,2} - z_{i,1} - z_{i,2}}{2}, \tag{3}$$

where $H_{i,1}$, $H_{i,2}$ and $z_{i,1}$, $z_{i,2}$ are the heads and elevations, respectively, for the end nodes of the pipe.

Leakage outflows Q_{Li} from pipes can be regarded as the elements of a vector $\mathbf{Q_L}$ ($n_p \times 1$), evaluated through the following vector equation:

$$\mathbf{Q_L} = \mathrm{diag}(\mathbf{CLL}) \left(\frac{|\mathbf{A_{10}}|(\mathbf{H_0} - \mathbf{z_0}) + |\mathbf{A_{12}}|(\mathbf{H} - \mathbf{z})}{2} \right)^{n_{leak}}, \tag{4}$$

where **CLL** is a vector with size $n_p \times 1$, whose elements are $C_{L,i}L_i$ and diag(**CLL**) represents the diagonal matrix associated with the vector. In Eq. (4) matrices $\mathbf{A_{12}}$ and $\mathbf{A_{10}}$ come from the incidence topological matrix **A**, with size $n_p \times$ nn (nn is the total number of network nodes). In the generic row of **A**, associated with the generic