Environmental Modelling & Software 69 (2015) 414–424

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

An analysis of the interface between evolutionary algorithm operators and problem features for water resources problems. A case study in water distribution network design☆

K. McClymont, E. Keedwell*, D. Savic

College of Engineering, Mathematics and Physical Sciences, University of Exeter, Harrison Building, North Park Road, Exeter EX4 4QF, United Kingdom

ARTICLE INFO

Article history:
Received 12 September 2014
Received in revised form
24 November 2014
Accepted 1 December 2014
Available online 7 February 2015

Keywords:
Water distribution network design
Optimisation
Problem understanding
Problem analysis
Operator and problem linkage
Evolutionary algorithms

ABSTRACT

Evolutionary Algorithms (EAs) have been widely employed to solve water resources problems for nearly two decades with much success. However, recent research in hyperheuristics has raised the possibility of developing optimisers that adapt to the characteristics of the problem being solved. In order to select appropriate operators for such optimisers it is necessary to first understand the interaction between operator and problem. This paper explores the concept of EA operator behaviour in real world applications through the empirical study of performance using water distribution networks (WDN) as a case study. Artificial networks are created to embody specific WDN features which are then used to evaluate the impact of network features on operator performance. The method extracts key attributes of the problem which are encapsulated in the natural features of a WDN, such as topologies and assets, on which different EA operators can be tested. The method is demonstrated using small exemplar networks designed specifically so that they isolate individual features. A set of operators are tested on these artificial networks and their behaviour characterised. This process provides a systematic and quantitative approach to establishing detailed information about an algorithm's suitability to optimise certain types of problem. The experiment is then repeated on real-world inspired networks and the results are shown to fit with the expected results.

1. Introduction

Evolutionary algorithms (EAs) have been applied to a countless number of problems across a wide variety of disciplines. Their relative simplicity and ability to work well on new problems have led to them being adopted in fields as diverse as engineering, economics and robotics. As one would expect, EAs have also been applied to water resources problems with a large degree of success, for example in the fields of groundwater remediation (Piscopo et al., 2014), controlling channel bed morphology (Nicklow et al., 2003), determining the hydraulic characteristics of production wells (Jha et al., 2004) and in particular to the field of water distribution network optimisation (e.g. Savic and Walters, 1997, Bi et al., 2015). A key aspect of EAs is that they have a number of parameters to set when first considering a new problem (e.g. population sizes, mutation and crossover rates, selection pressure etc.) and this usually means that a period of parameter tuning is necessary to deliver acceptable performance. To alleviate this research has been carried out on algorithms to adapt these parameter settings automatically, removing the need for the parameter tuning period. A natural extension to this process then is to consider whether not just the parameters, but the operators themselves might be selected in an adaptive fashion, leading to the field known as selective hyperheuristics (Burke et al., 2013). In essence, this field explores the potential for algorithms to function beyond the strict application of selection, mutation and crossover phases and aims to develop a more dynamic approach with a greater number of operators to produce better results with less human intervention. However, to develop a suitable pool of operators from which a set might be chosen we must first understand the relationship between problem characteristics and operator function. The work herein introduces a process for exploring this interaction, and provides empirical results on a range of different artificial and real-world problems using water distribution networks as an example case study.

☆ Thematic Issue on the Evolutionary Algorithms.

* Corresponding author.
E-mail address: E.C.Keedwell@ex.ac.uk (E. Keedwell).

http://dx.doi.org/10.1016/j.envsoft.2014.12.023
1364-8152/

1.1. Water distribution network design problem

Water distribution networks (WDNs) represent one of the most complex and key infrastructures in use today and are responsible for the transportation of clean drinking water from reservoirs and storage tanks to industrial and residential consumers. Failure of these networks to adequately supply the demand can cause significant problems in the day-to-day running of businesses and homes.

A standard WDN is comprised of pipes, nodes (junctions and demand points), hydraulic devices (such as pumps) and sources (tanks and reservoirs) that constitute the entire infrastructure that delivers water from the source (e.g., reservoir) to various locations where it is drawn from the network for consumption (e.g., residential housing or industrial sites). With increasing demand and tighter regulation, water companies continue to search for more optimal operations and improvements in their networks and so have in the last two decades looked towards emerging optimisation methods to help solve their problems. Real-world WDNs are complicated structures that require constant operational management, maintenance and rehabilitation. In order to satisfy consumer demand, the networks must be constructed with a good layout that connects to all points of demand and should provide the best possible hydraulic conditions and operational requirements all whilst minimising network cost. This is known as the WDN design problem.

The WDN design problem is known to be an NP combinatorial problem (Yates et al., 1984). Even for relatively small networks, the number of possible combinations of pipes is very large which makes enumeration of all the possible designs impossible. If, for example, there were six potential sizes for each pipe in a network of just thirty pipes, there would be $6^{30} = 2.21 \times 10^{23}$ possible combinations – far more than is possible to fully enumerate within a reasonable time. This basic complexity is further compounded when advanced controls such as pump scheduling and valve operations are considered in combination with the much larger models (e.g. 1000s of pipes) that are likely to be found in the real world. Finally, when the potential for independency in looped structures and the non-linearity of the hydraulic equations is included, it becomes clear the WDN design problem is difficult non-linear, multi-modal problem. It is because of this that researchers and practitioners look to more advanced meta-heuristics to optimise their WDN designs.

1.2. Optimisation of WDNs

Since the first application of optimisation methods (Blum and Roli, 2003) to the problem of water distribution network (WDN) design researchers have collectively established a large body of literature on the subject (Marchi et al., 2014). The majority of these studies are focused on the application of novel optimisation methods to this problem, novel formulations of the problem (McClymont et al., 2013), or case studies of real-world instances. In addition, these studies have often employed or proposed new meta-heuristic methods; predominantly those from Evolutionary Computation (EC) (Coello et al., 2007) and variants of Evolutionary Algorithms (EAs) (Laumanns et al., 2000). More recently, there has been a shift in focus to hybrid (Keedwell and Khu, 2005) or more adaptive methods (Afshar, 2006) for the optimisation of these problems such as multi-method search (Vrugt and Robinson, 2007; Vrugt et al., 2009; Raad et al., 2010) and selective hyper-heuristics (McClymont et al., 2013).

The work presented here investigates search operators and their interaction with features in the fitness landscape for water distribution network optimisation, the first time this has been attempted, although attempts have been made to undertake studies of a similar nature in other domains (Franchini and Galeati, 1997). In a similar study, Zecchin et al. (2012) investigated Ant Colony Optimisation Algorithms in relation to Water Distribution Network problem characteristics and also highlighted the importance of these studies. Indeed, the optimisation of WDNs by Evolutionary Algorithms (EAs) in the early 1990s (Walters and Lohbeck, 1993; Simpson et al., 1994; Savic and Walters, 1995) was the start of a wider effort to find new, more efficient and effective optimisation techniques for this difficult real-world problem.

1.3. Performance analysis

In the search for better optimisation methods, papers frequently attempt to analyse the performance of meta-heuristics by applying them to a set of large, realistic WDNs (Walters et al., 1999; Cheung et al., 2003) or other water networks, such as in Fu et al. (2008). This experimental method provides vital information on the scalability of the proposed techniques. However, when considering the use of adaptive techniques to select operators it is important to understand the impact that individual search space features have on the behaviour and suitability of a method to that type of search space. What is required is to establish a fundamental understanding of the effect of different WDN features and landscape attributes on optimisation methodologies and lay the ground work for the new approaches described later. To make confident assertions about the true behaviour, such as the explorative or exploitative search of an algorithm, quantitative analysis of the algorithms is required (Deb and Jain, 2002; Merz, 2004).

Furthermore, while these works are important in developing better techniques for solving this class of difficult and constantly evolving real-world problems (McClymont et al., 2013), there has been relatively little work conducted on the analysis of the fundamental rules of how these optimisation methods behave under different conditions in the context of the WDN problem. Consider, for example, the significant differences in the hydraulic properties of a looped network versus a dendritic network (Walters and Lohbeck, 1993) or, similarly, a gravity fed network versus a network of pumps with tank storage. These variations in structural and hydraulic properties will result in very different optimisation search space landscapes for operators to traverse. Therefore, while one optimisation method might perform well on certain types of network, it is equally likely that it will perform less well on others. It is important to understand this relationship, between optimiser and problem, in order to make well-grounded claims about any one method's suitability for solving the WDN problem, or certain variants of it. This understanding will also help to guide algorithm and operator selection for this class of problems.

1.4. Problem and operator linkage

It is clearly shown by the No-Free-Lunch theorem (Wolpert and Macready, 1997) that not all optimisers are well suited to solving all problems. Similarly, it can be said that not all operators in an optimiser are well suited to solving all problems. This statement can be generalised somewhat to say that not all optimisation operators behave in the same way and therefore are not suited to all problems. The question therefore is: to what extent is it possible to ascertain a profile detailing the behaviour of an optimiser or its operator(s) and to determine how this profile relates to specific problems and problem features?

Malan and Engelbrecht (2013) provide some insight into the concept of characterising generalised fitness landscapes and the early work by Kauffman (1989) suggests adaptation based on these variances in the landscape is feasible. Furthermore, studies such as