Spatial and Temporal Visualisation of Evolutionary Algorithm Decisions in Water Distribution Network Optimisation

Ed Keedwell
College of Engineering, Mathematics and Physical Sciences
University of Exeter
North Park Road
Exeter
EX4 4QF
+44 1392 724014
E.C.Keedwell@ex.ac.uk

Matthew Johns
College of Engineering, Mathematics and Physical Sciences
University of Exeter
North Park Road
Exeter
EX4 4QF
+44 1392 724014
mbj202@ex.ac.uk

Dragan Savic
College of Engineering, Mathematics and Physical Sciences
University of Exeter
North Park Road
Exeter
EX4 4QF
+44 1392 723637
D.Savic@ex.ac.uk

ABSTRACT

Much research has been conducted into the visualisation of objective space and decision space landscapes. This work moves away from this and investigates a 3D interactive method for linking EA decisions through time with the design of engineering systems. The proposed system shows through an intuitive interface, the design space being explored by the algorithm including decision variable choices, locations that are fixed early on in the optimisation and those problem areas that are difficult for the algorithm to solve. The paper presents a case study in water distribution network design, although the methods described are, in principle, generalisable to other design domains.

Categories and Subject Descriptors

G.1.6 [Optimization]: *Global optimization*

Keywords

Evolutionary algorithms; visualization; problem understanding; water distribution network optimisation

GECCO'15 Companion, July 11 - 15, 2015, Madrid, Spain

ACM 978-1-4503-3488-4/15/07…$15.00
DOI: http://dx.doi.org/10.1145/2739482.2768442

1. INTRODUCTION

1.1 EA Visualisation

A great deal of previous research has focussed on the visualisation of objective spaces, both single objective, multi- and many objectives and in the visualisation of decision variable spaces as characterised by the search process. This analysis has naturally tended to focus on the spaces in which an evolutionary algorithm operates in an effort to improve performance and understanding of the progress of an evolutionary algorithm (EA). These visualisations often consider the problem statically e.g. in the visualisation of the final solution or Pareto front of an algorithm, or through temporal snapshots where the execution of the algorithm can be traced and the dynamics of the optimisation can be observed.

Furthermore traditional visualisation can either be seen to be focussing on the developer of algorithms to aid in the algorithm's development and tuning, or on the end user as a method to deliver the outcomes of an optimisation to its intended audience and to aid decision support in the real-world.

In this paper we propose a method that takes elements of all these approaches and combines them into a single visualisation that can simultaneously visualise aspects of the best solutions discovered and the areas of activity of an evolutionary algorithm within the problem class of water distribution network optimisation, with additional problem domains also considered from a theoretical perspective.

1.2 Water Distribution Network Optimisation

1.2.1 Problem Description

Water distribution network (WDN) optimisation is an important real-world application for optimisation techniques and evolutionary algorithms in particular. These networks deliver fresh drinking water from reservoirs, tanks and water treatment works to households and businesses via a network of pipes and making use of a variety of other assets such as pumps and valves to regulate pressures and water supply. Typically, the optimisation of these networks aims to design new networks or enhance existing ones, to deliver drinking water at an adequate pressure to all demand points for the minimum possible cost. Although this is the primary task for optimisation in this domain (and the focus of this paper), there are many other objectives that can be considered including the minimisation of water age (improving water quality), adherence to velocity and pressure constraints (reducing the prospect of leakage) and increasing the robustness of the network to reduce the potential for supply outages. In this particular problem set, only the simplest WDN optimisation problem is considered where the decision variables are a set of diameters for each pipe within the network and the objectives are to meet the required pressure (head) throughout the network and minimise the overall cost of constructing the network. Though simplified, this problem is still one of high real-world importance and optimality in the solutions developed can have large scale financial, social and environmental impacts when applied to large-scale real-world examples.

1.2.2 *Computational Formulation*

The WDN optimisation problem is characterised as an NP-Hard combinatorial optimisation problem with large-scale multi-modal search landscapes. The algorithm must select from a list of discrete diameter options for each pipe within the network which constitutes the set of decision variables for the algorithm. A full set of decision variables describes a new network that is simulated by a hydraulic simulator, in this case Epanet 2 [1], which provides the information necessary to calculate the hydraulic values and to determine to what extent the network meets the hydraulic constraints. In this formulation, the two objectives are:

$$cost = \sum_{i=1}^{k}(c_i \times l_i) \tag{1}$$

Where i represents one of the total number of pipes k in the WDN, and c represents the cost of the selected diameter of pipe i and l represents its length (in feet or metres), and:

$$headDeficit = \sum_{n=1}^{m}((h_t - h_n) > 0) \tag{2}$$

Where n represents one of the total number of demand nodes m in the WDN and h represents the hydraulic head (in feet or metres) at that node, h_t represents the target head for each node which is usually, but not necessarily, set as a uniform value for all nodes within the network. Only those nodes for which a deficit is recorded are considered to remove the possibility of nodes with head excess compensating for those with deficit.

Cost and head deficit can be treated as separate objectives in a multi-objective formulation or combined into a single objective in the standard fashion:

$$objective = cost \times \alpha(headDeficit) \tag{3}$$

Where α can be used to balance the optimisation between the cost and head deficit elements of the optimisation. This factor will be required as most WDN problems have costs in the millions and head deficits typically are in the range of small hundreds. α is usually set on a case-by-case basis for each problem and has been determined manually here to ensure balance between the objectives. A detailed analysis of this process is out of the scope of this paper.

1.2.3 *Computational Complexity*

The WDN optimisation problem necessitates a metaheuristic approach to optimisation (such as an EA) through the computational complexity of its search space. Each change in pipe diameter changes the hydraulic conditions in the network for surrounding elements and potentially for all elements in the network. Therefore the problem is combinatorial in nature and as such the number of possible combinations of diameters is:

$$possible\ networks = D^P \tag{4}$$

Where D is the number of diameters in the network and P is the number of pipes. Even small networks such as the New York Tunnels [2] problem that has 16 diameter choices and just 21 pipes to select from has 1.934×10^{25} possible network combinations. In addition, hydraulic simulations typically incur a non-trivial computational load and can require up to several seconds to run.

Clearly problems of this nature, particularly in real world scenarios necessitate a stochastic or metaheuristic based search to efficiently search these large spaces.

1.3 Water Distribution Network Optimisation with Evolutionary Algorithms

A large body of research exists within the literature relating to the optimisation of water distribution networks. This ranges from the early work with single-objective evolutionary algorithms [3], through multi- [4] and many- [5] objectives to recent work on multi-method search and hyperheuristics [6]. This work has concentrated on generating new algorithms and producing results on various benchmark problems and for large-scale real-world problems taken from the industry. Although there is some work on decision support within the industry, there has not as yet, been extensive development of tools for industrialists to interact with the algorithms and solutions created to better understand their networks.

1.4 Water Distribution Network Visualisation

Rudimentary visualisations of networks are available in Epanet 2 which show a plan view of the layout of the network along with colour coding to show a selected aspect of the network, e.g. pressures, velocities, heads, diameters etc.

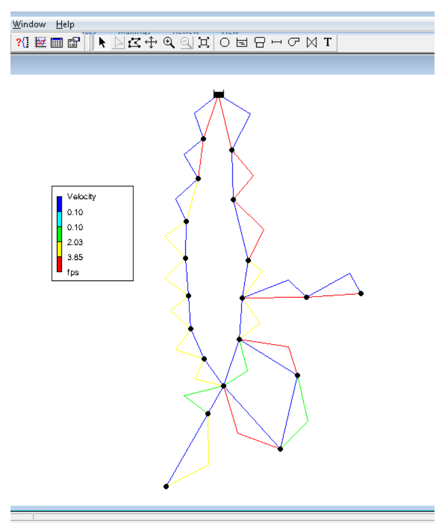


Figure 1: Epanet 2 visualisation of the New York Tunnels network.

Although more sophisticated visualisations exist within commercial packages in the industry, this plan view of the network is typical of the visualisations of WDNs.

There have also been efforts to visualise outputs in this area though. For example [7] has implemented various methods for the visualisation of the cloud of points in objective space generated by multi-objective methods through time. This work has focussed on supporting decision makers by allowing them to interact with the Pareto-surfaces generated by multi-objective algorithms. By allowing the manipulation of these surfaces, decision makers can be guided through N-dimensional clouds of points without imposing constraints and allowing 'what if' scenarios to be investigated. A further study, also in multi-criteria decision making, although not in the evolutionary optimisation area, uses the 'power index' to visualize the ranks of various key performance indicators in the water industry [8] and again shows the power of visualisation for complex datasets in this domain.

The above approaches show that there is interest in the industry for visualisations of networks and their potential to aid decision makers particularly in complex multi-dimensional objective and decision spaces.

2. METHOD

2.1 WDNet3D

As described in [9] a 3D approach to the visualisation of WDNs has been implemented using the open source Panda 3D visualisation libraries[1]. Each network is comprised of the same set of elements as would be visible in the plan view but uses a 3D perspective (as shown in Figure 2) to provide important extra functionality.

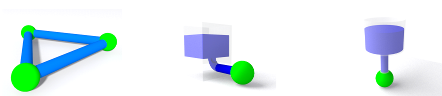


Figure 2: 3D visualisation of water distribution network components (visualized from left to right – pipes, a reservoir, a tank and a pump).

The notable extra functionality is that spatial elements such as distances, pipe diameters, lengths and crucially the elevation of elements, are visualized implicitly without recourse to colour coding or other artificial mechanisms. This effect is illustrated in Figure 3, a 3D visualisation of the Hanoi benchmark network. Although the pipe network grid appears flat, the elevated position of the reservoir can be seen implicitly and is a common arrangement in this type of network known as a 'gravity fed' system. Also notable is that the diameters of the pipes can be seen in the visualisation. This implicit visualisation allows the user to see potential bottlenecks and oversizing of mains within a network. As can be seen from Figure 3, the colour of the pipes has not yet been used to convey any information and so additional functionality is gained from visualizing velocity, pressure, water quality and other variables within the network. Of course, this is possible whilst also visualizing the implicit aspects of the pipe assets within the system.

The system is fully interactive and for larger networks the user can pan around the network and zoom in on particular areas of interest.

[1] https://www.panda3d.org/

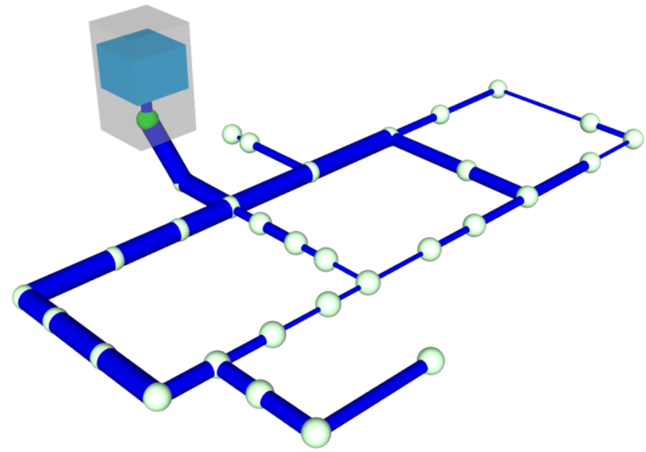


Figure 3: WDNet3D Visualisation of the Hanoi Benchmark WDN.

2.2 Incorporating EA Optimisation

By visualizing network construction implicitly, the colour of elements is available to visualize other aspects of the system and this work focuses on this to visualize evolutionary progress. In particular, the colour of pipes is used to represent the number of times that the EA has modified that variable in the best solution for the most recent N generations. This allows the end user to understand the interaction of the EA with the network on a spatial level and to better understand which pipes are requiring most effort from the EA to resolve.

2.3 Evolutionary Algorithm

A standard steady state, single objective evolutionary algorithm is used to produce solutions for visualisation. It uses a binary representation with each pipe represented by the requisite number of bits for the number of available pipe diameters, e.g. a pipe with 16 possible diameters requires 4 bits for each pipe. Simple random mutation and single point crossover are also used. This algorithm is popular in the application area and is sufficient to illustrate the proposed visualisation method. In practice, any EA formulation (including multi-objective formulations), or iterative optimisation algorithm for that matter could be used. The objective function was calculated as shown in equation 3.

2.4 Benchmark Networks

The system is applied to a number of publicly available benchmark networks[2]. The problems have the following characteristics:

Hanoi – A simple network with 32 nodes and 34 pipes with 6 possible diameters. Possible combinations: ~2.8×10^{23}.

Anytown – 35 existing pipes and 6 possible new pipes with 10 possible pipe diameters. 19 nodes with varying demands. It should be noted that the original version of this problem includes a number of operational aspects that have been fixed in this formulation where only pipe sizing is considered. Possible combinations: ~1.0×10^{41}.

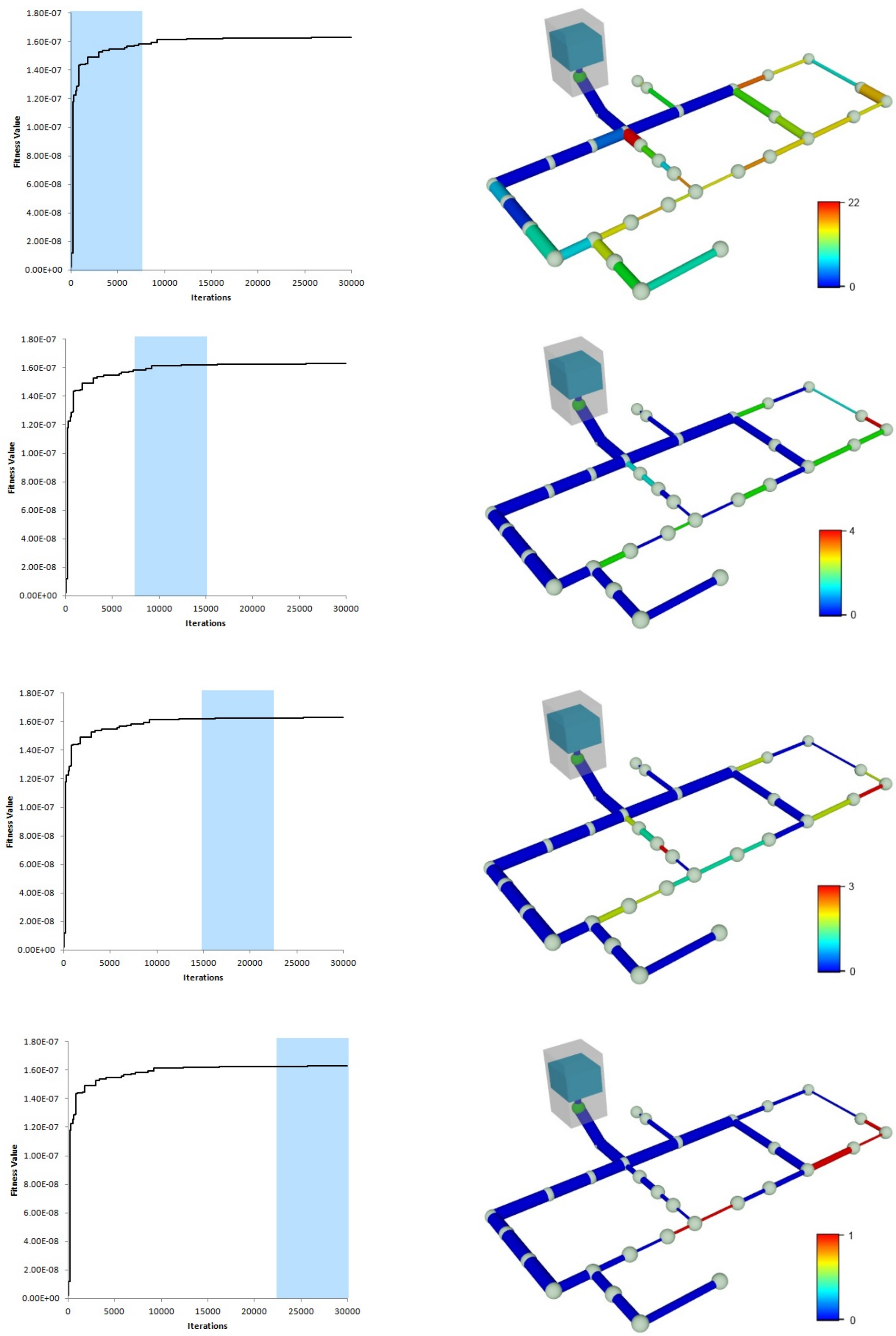
Epanet Example Network 3 – Larger real-world inspired network with 117 pipes and 92 nodes. 16 pipe diameters are used in this formulation. Possible combinations: ~ 7.6×10^{140}.

[2] http://emps.exeter.ac.uk/engineering/research/cws/resources/benchmarks/

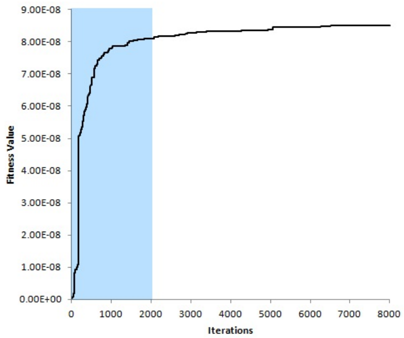
3. RESULTS

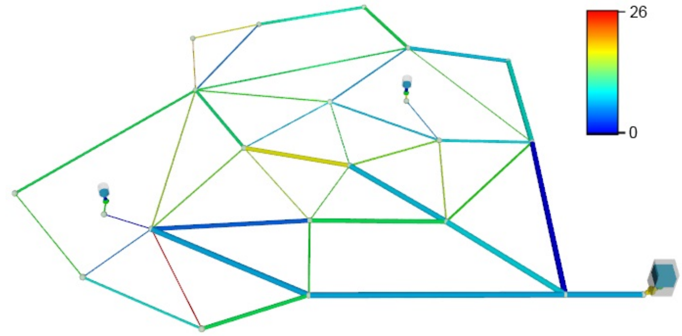
In each of the following, a network has been visualised with the WDNet3D system at points throughout the optimisation process. In each case, the period of the optimisation can be seen in the left panel highlighted in blue and network, which EA changes to pipe diameters shown in the right panel. The legend (which changes for each 'snapshot') shows the relationship between colour and the number of changes made to the best solution in the selected period of optimisation. This could also be visualised in video.

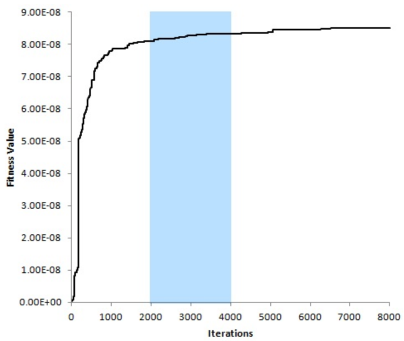
3.1 Hanoi

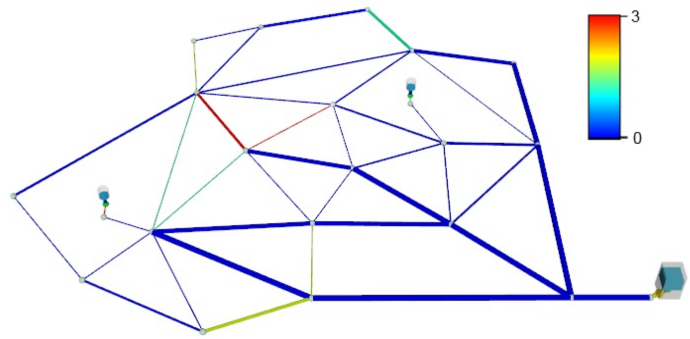


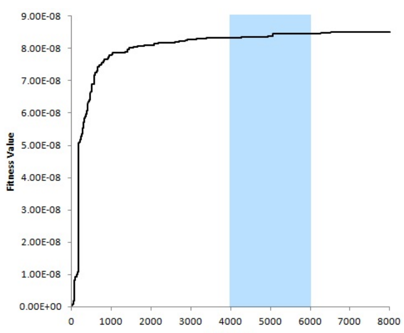
3.2 Anytown

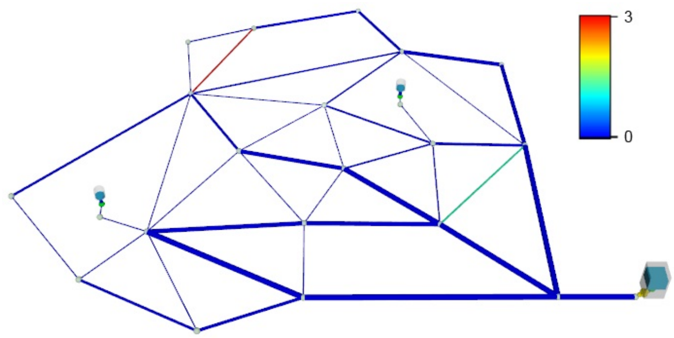

9.00E-08
8.00E-08
7.00E-08
6.00E-08
5.00E-08
4.00E-08
3.00E-08
2.00E-08
1.00E-08
0.00E+00
Fitness Value
0
1000
2000
3000
4000
5000
6000
7000
8000
Iterations

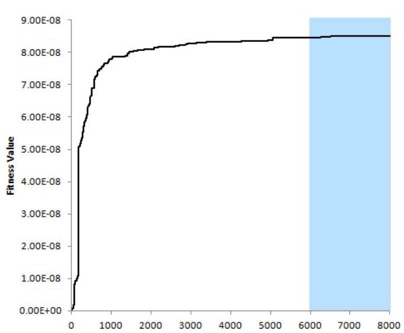

26
0

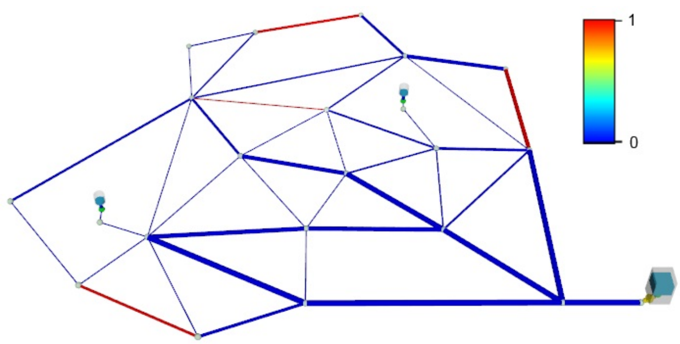

9.00E-08
8.00E-08
7.00E-08
6.00E-08
5.00E-08
4.00E-08
3.00E-08
2.00E-08
1.00E-08
0.00E+00
Fitness Value
0
1000
2000
3000
4000
5000
6000
7000
8000
Iterations


3
0

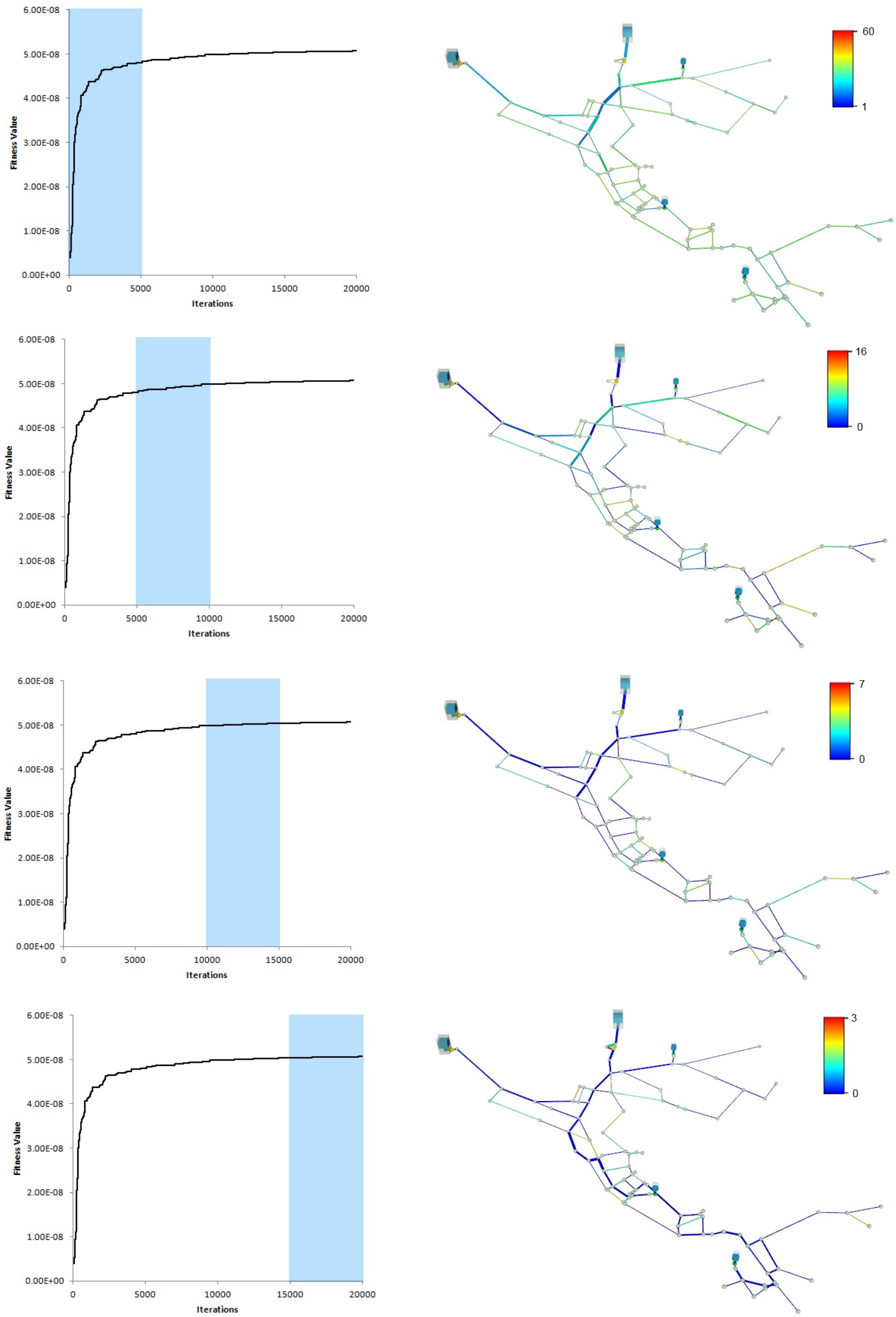

9.00E-08
8.00E-08
7.00E-08
6.00E-08
5.00E-08
4.00E-08
3.00E-08
2.00E-08
1.00E-08
0.00E+00
Fitness Value
0
1000
2000
3000
4000
5000
6000
7000
8000
Iterations


3
0


9.00E-08
8.00E-08
7.00E-08
6.00E-08
5.00E-08
4.00E-08
3.00E-08
2.00E-08
1.00E-08
0.00E+00
Fitness Value
0
1000
2000
3000
4000
5000
6000
7000
8000
Iterations


1
0

3.3 Example Network 3



4. DISCUSSION

In the previous section, EA behaviour for three networks has been visualised for various windows over the optimisation run. The colours of the pipes indicate the number of times that these have been changed in the best solution. Cooler colours represent fewer modifications and warmer colours indicate larger numbers of changes by the algorithm over time. It should be noted that the colour range is reset and normalised for each timeslice to allow for easier interpretation.

4.1 Hanoi

For this simple network it is clear that the algorithm has decided the pipe sizes for the main transmission elements of the system (the 'trunk' mains) very early on in the optimisation as shown by the larger blue pipes emanating from the reservoir. The following snapshots show an increasingly fixed backbone of a network with only the extremities being modified later in the optimisation.

4.2 Anytown

From this visualisation it is clear that Anytown presents a more difficult problem to the EA, with very few pipes achieving 'fixed' status over the first 2000 iterations. From thereon, the story is somewhat similar with infrastructure emanating from the main reservoir being progressively fixed along with some close to the tanks. It is clear that the majority of effort is being expanded along the more 'looped' sections of the network where small improvements can be made.

4.3 Network 3

Network 3 is much closer to a real-world style network and has more variables than the other two examples. In this example, the EA first optimises the important infrastructure between the two reservoirs and tank towards the North of the network. It then proceeds to identify and 'fix' the central 'trunk' mains that link the reservoirs to the North with the rest of the network in the South. By the final snapshot, as with the two previous examples, the algorithm is concentrating on other sections and the majority of refinement appears to be taking place in the smaller mains parallel to the trunk.

4.4 Conclusion

The three example networks all show that the evolutionary algorithm tends to size and 'fix' the trunk main infrastructure (effectively the macro-level problem) early on in the optimisation. However, it would appear that the extent to which this occurs is dependent on the size of the problem but also its' interconnectedness. Despite its relatively modest size, Anytown shows a resistance to this fixing behaviour, that is likely to be due to its lack of a central set of trunk mains and highly looped structure. The larger 'real-world' Network 3 experiences more fixing despite its larger number of assets because it has a more traditional 'trunk main' style layout.

As these changes are made to the best solution in each generation, there may be a case for fixing these parameters during the EA optimisation to reduce running times and to allow the algorithm to focus on those areas of the network that require modification. The relatively small number of changes seen in the latter stages of the optimisation belie the fact that, through the uniform random mutation operator, many thousands of alternative 'trunk main' configurations will have been tried and discarded.

5. CONCLUSION

A system , WDNet3D has been shown to provide 3-dimensional visualisations of water distribution networks. It has been demonstrated that these visualisations can provide more information to the end user than the typical plan view of the network. Furthermore it has been shown that these visualisations can be used to provide information on the decisions being made by the evolutionary algorithm throughout the optimisation. This has revealed that the algorithm tends to concentrate on the central supply mains first and fine tunes this configuration in loops and more extraneous areas later on in the optimisation run. Key differences among the optimized networks show that the extent to which the algorithm 'fixes' the infrastructure is dependent on the degree to which the network possesses a trunk main style of system. This raises the possibility of potentially 'fixing' these decision variables and removing them from consideration for the EA, although this is not trialled here. Finally, by viewing the visualised networks throughout the optimisation, it is possible for the end user and EA developer to gain a greater understanding of which parts of the network are more difficult for the algorithm to optimize. This is likely to lead to greater understanding of the interface between the optimisation algorithm and the problem domain for this important class of real-world problems.

5.1 Application to Other Problem Domains

The approach applied here to water distribution network optimisation could be implemented in other problem domains to aid developer and end-user understanding of the optimisation process. A number of such application domains are discussed here.

5.1.1 Network-Based Problems

It is simple to envisage how this approach could be adapted to other network-based problems such as the travelling salesman problem or gene regulatory network optimisation. However, as these problems require the discovery of routes/subnetworks which represent a subset of all possible routes or the fully connected network, the approach would need to be adapted to this problem type. In this adaptation, edges in the final solution could be coloured according to the number of times they have featured in the best solution discovered by the algorithm over time in a similar fashion to the method presented here, although it would be possible to colour the entire network search space (e.g. every edge in the TSP) problems to provide more information on which edges have been visited most often.

5.1.2 Operations Research Problems

Operations research problems typically include problems such as resource allocation, scheduling and routing. In each case, the colouring of the resource being allocated, the event being scheduled or route fragment being selected (e.g. as in 5.1.1) could be adopted. An example of the former is the bin packing problem, an operations research problem that is not network-based. The application of the approach to this problem is illustrated here.

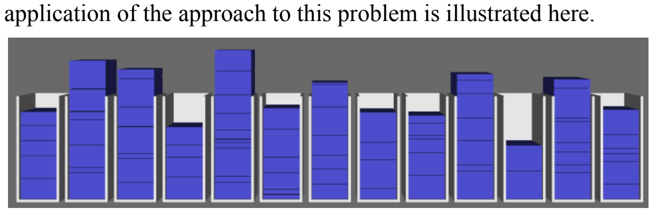


Figure 4: Bin packing illustration

Figure 4 illustrates a (non-optimised) solution to the bin packing problem where items of differing size are being packed into the various bins available. To adapt the proposed approach, each item would be coloured according to the number of times it has moved between bins in the best solution discovered by the algorithm. It is anticipated that this would yield similar results to the water distribution network optimisation problem in that certain items would be ‘fixed’ within the solution early in the optimisation and the algorithm would then seek to optimise around these fixed blocks, possibly highlighting those items that are more difficult for the algorithm to place.

5.1.3 Benchmark and Continuous Problems

Many EA benchmark problems contain decision variables that are not mapped to a physical construct in the same way that those in the above problems are. In this case, a potential adaptation could visualise the optimisation as a heatmap where rows represent decision variables and columns represent the ‘snapshots’ of the optimisation through time. The colour of each cell in the heatmap would represent the number of times that each decision variable has changed in the best solution discovered by the algorithm to that point in time.

One further adaptation to the method, regardless of optimisation domain, and a prospect for further work, would investigate not simply the number of changes to the network for each variable but also the magnitude and direction of those changes. This revised colouring would then provide more information for all the problem domains thus far described and would make the approach more amenable to real-valued optimisation problems where the magnitude and direction of change in the best solution variables is more important than simply whether the variable has changed.

6. REFERENCES

[1] Rossman, L. A., "EPANET 2 Users Manual", 2000

[2] Schaake J.C. & Lai D. (1969) Linear Programming and Dynamic Programming Applications to Water Distribution Network Design. Report 116 Hydrodyn, laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

[3] Dandy, G. C., Simpson, A. R., & Murphy, L. J. (1996). An improved genetic algorithm for pipe network optimization. Water Resources Research, 32(2), 449-458.

[4] Prasad, T. D., & Park, N. S. (2004). Multiobjective genetic algorithms for design of water distribution networks. Journal of Water Resources Planning and Management, 130(1), 73-82.

[5] Khu, S. T., & Keedwell, E. (2005). Introducing more choices (flexibility) in the upgrading of water distribution networks: the New York city tunnel network example. Engineering optimization, 37(3), 291-305.

[6] Raad, D., Sinske, A., & van Vuuren, J. (2010). Multiobjective optimization for water distribution system design using a hyperheuristic. Journal of Water Resources Planning and Management, 136(5), 592-596.

[7] Kollat, J. B., & Reed, P. (2007). A framework for visually interactive decision-making and design using evolutionary multi-objective optimization (VIDEO). Environmental Modelling & Software, 22(12), 1691-1704.

[8] McClymont, K., Walker, D. J., Keedwell, E., Everson, R. M., Fieldsend, J. E., Savic, D., & Randall-Smith, M. (2011) Novel Methods for Ranking District Metered Areas for Water Distribution Network Maintenance Scheduling. Computing and Control in the Water Industry, CCWI, Exeter, UK.

[9] Keedwell, E. C., Johns, M. B., & Savić, D. A. (2014). Interactive 3D Visualization Of Optimization For Water Distribution Systems 11th International Conference on Hydroinformatics, New York.