European Water 58: 375-381, 2017.

Efficient parallel evolutionary optimization algorithm applied to a water distribution system

T.T. Tanyimboh[1,2*] and A.G. Seyoum[1,3]
[1] *Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, UK*
[2] *School of Civil and Environmental Engineering, University of the Witwatersrand, Johannesburg, South Africa*
[3] *SWS Consultancy, Addis Ababa, Ethiopia*
[*] *e-mail: tiku.tanyimboh@wits.ac.za*

Abstract: Water distribution systems are key components of public infrastructure and it is essential to design, manage and rehabilitate them economically without compromising the required performance and regulatory standards. Evolutionary optimization algorithms such as genetic algorithms have become popular in providing optimal and near optimal solutions to various optimisation problems on water distribution systems. However, one of the main challenges associated with genetic algorithms in the optimization of water distribution systems is that they are time consuming when applied to problems on real-world water distribution networks with large numbers of pipes and multiple operating conditions. For example, in the optimisation of large water distribution systems, a single optimisation run may involve millions of hydraulic and water quality simulations that may take many days on modern powerful computers such as workstations. One way to address this difficulty is by utilising parallel computing methods. The aim of the research was to investigate the potential improvement in computational efficiency achievable through parallel computing by optimizing a design based on a real-world water distribution network with a large number of decision variables, large solution space and complex response surface. The results showed that the parallel algorithm developed was practical and found optimal and near-optimal solutions reliably and efficiently, and the solutions achieved were consistently competitive. The average speedup achieved on an eight-core workstation was 15, based on an optimization problem with hundreds of decision variables for a real-world water distribution network with extended period simulation.

Key words: Water distribution network, penalty-free multiobjective evolutionary optimization, pressure-driven analysis, discrete optimization, demand satisfaction ratio, controller-worker model

1. INTRODUCTION

Water distribution systems (WDSs) are key components of public infrastructure and it is essential to design and rehabilitate them in a cost effective manner without compromising the required performance and regulatory standards. Evolutionary algorithms (EAs) such as genetic algorithms (GAs) have become popular in providing optimal and near optimal solutions to WDS optimisation problems.

One of the main issues with genetic algorithms is that the algorithms are time consuming when applied to optimisation problems on real-world networks with large numbers of pipes and multiple operating conditions (Van Zyl *et al.* 2004). For example, in the optimisation of large water distribution systems, a single optimisation run may involve millions of hydraulic and water quality simulations (Rossman 2000, Ghebremichael *et al.* 2008, Spiliotis and Tsakiris 2011, Todini and Rossman 2013) that may take many days on modern computers such as workstations.

One way to improve the computational efficiency of evolutionary optimization algorithms is by adopting parallel computing. Indeed, parallel computing has the added benefit that it can be used together with other strategies. In the context of water distribution systems, some of the other strategies for improving efficiency include: (a) the use of supercomputers (Barlow and Tanyimboh 2014); (b) regression models to estimate the hydraulic properties rapidly (Vairavamoorthy and Ali 2000); (c) surrogate measures of system performance, such as flow entropy, that are relatively easy to calculate (Tanyimboh and Setiadi 2008, Moosavian and Lence 2016); (d) techniques for reducing the solution space of the optimization problem (Kadu *et al.* 2008, Kang and Lansey 2012); and (e) local search and hybrid methods (Haghighi *et al.* 2011).

Evolutionary algorithms for water distribution systems often use penalties to evaluate the merits of infeasible solutions when solving optimization problems that have constraints. By contrast, the penalty-free multi-objective evolutionary algorithm (PF-MOEA) proposed by Siew and Tanyimboh (2012b) uses pressure-dependent analysis that accounts for the pressure dependency of the nodal flows and thus avoids the need for penalties to address any violations of the minimum nodal pressure constraints. The hydraulic simulator used was a pressure-driven extension of EPANET 2 called EPANET-PDX (Seyoum and Tanyimboh 2014, 2016, Siew and Tanyimboh 2012a, Tsakiris and Spiliotis 2014, Sivakumar and Prasad 2015).

Given that fitness evaluation is often the most computationally expensive step in evolutionary optimization algorithms (Alba and Tomassini 2002, Schutte *et al.* 2004), the aim of this research was to investigate the improvement in the computational efficiency achievable through parallel computing. A design problem on a real-world water distribution network with a large number of decision variables and large solution space was investigated using extended period simulation.

2. OPTIMIZATION PROBLEM FORMULATION AND SOLUTION

Evolutionary algorithms (EAs) by nature start with a randomly generated set of candidate solutions that may include both feasible and infeasible solutions. To address the node pressure constraints, penalty methods have been applied widely (Broad *et al.* 2005, Ostfeld and Tubaltzev 2008). The major drawback of the penalty-based approach is that additional case-specific parameters are required whose calibration is generally challenging (Prasad and Park 2004).

In an attempt to alleviate the difficulties associated with the minimum node pressure constraints, Siew and Tanyimboh (2012b) proposed a penalty-free multi-objective evolutionary algorithm (PF-MOEA) based on NSGA II (Deb *et al.* 2002). The PF-MOEA algorithm uses pressure-driven analysis to assess each individual in the population of solutions. In this way, the pressure-driven analysis addresses the node pressure constraints as an integral part of the hydraulic simulations. PF-MOEA employs the pressure-driven extension of EPANET 2 that is known as EPANET-PDX (Siew and Tanyimboh 2012a) and carries out pressure-driven analysis seamlessly. Hence, when comparing infeasible solutions, PF-MOEA does not employ the constraint dominance approach used in NSGA II.

The PF-MOEA formulation has been applied previously to various aspects of WDS optimization that included design, operation and long-term rehabilitation and upgrading (Siew and Tanyimboh 2012b, Siew *et al.* 2014, 2016). Overall, PF-MOEA generated superior results for the optimization problems solved in terms of the cost, hydraulic performance and computational efficiency, compared to all other results in the literature.

The PF-MOEA approach is practical and straightforward to implement as it utilises the most basic operators such as single-point crossover and single-bit mutation. The algorithm was selected in this research to investigate the relative merits of parallel computing due to the above-mentioned advantages. The decision variables in PF-MOEA are represented using binary coding. Single-point crossover, single-bit mutation and binary tournament selection for crossover are the genetic operators used in the algorithm.

Minimising the total network cost (capital and operation) and maximising the network hydraulic performance were the two conflicting objectives adopted, subject to the constraints of conservation of mass and energy and the minimum nodal pressures, as follows.

$$\text{Minimise: } f_1 = (CR)^2 \tag{1}$$

$$\text{Maximise: } f_2 = (DSR)^4 \tag{2}$$

where f_1 and f_2 represent the first and second objective functions, respectively. *CR* is the cost ratio i.e. the ratio of the cost of a particular solution to the cost of the most expensive solution in the whole population within a single generation. *DSR* is the demand satisfaction ratio i.e. the ratio of the available flow to the required flow and measures the feasibility of a solution. Feasible solutions have a *DSR* value of unity; solutions with *DSR* values below unity are infeasible.

Large problems can be divided often into smaller ones that can be solved simultaneously on parallel processors in a shorter time (Trobec *et al.* 2009). A controller-worker approach was applied to parallelise the penalty-free optimization algorithm employed. The controller-worker approach, with a potential to improve the computational performances of algorithms significantly, is relatively straightforward (Cantú-Paz and Goldberg 2000).

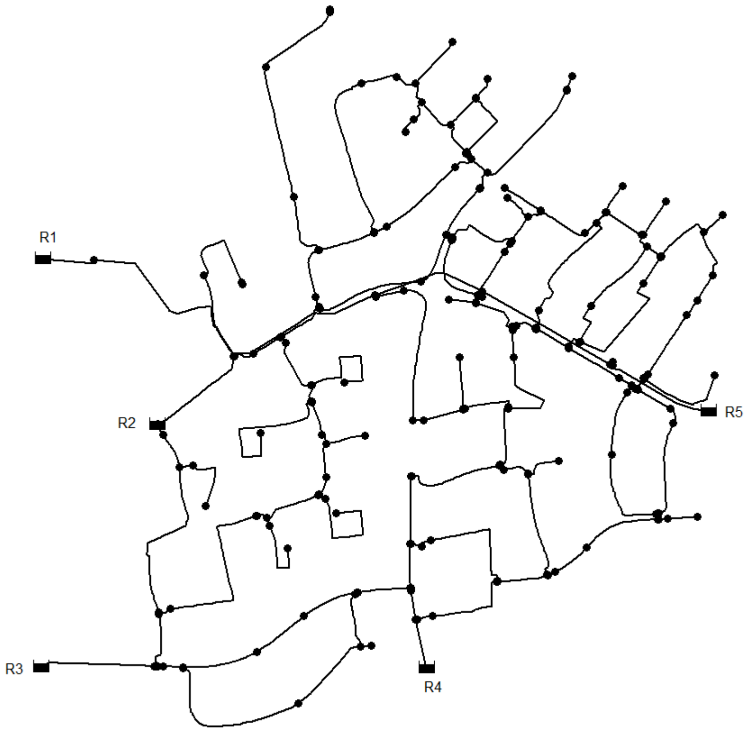
The fitness evaluation of the child population was parallelized, as fitness evaluation is often the most computationally expensive step in evolutionary algorithms (Alba and Tomassini 2002, Schutte *et al.* 2004). The fitness evaluations were divided equally among all the processors including the controller processor. The controller processor performed the rest of the operations of the evolutionary algorithm including selection, crossover and mutation, and all the other procedures of the optimization. The authors wrote the parallel algorithm program in C++ with Message Passing Interface (MPI) routines. Microsoft HPC pack 2008 was used to run the program in Microsoft Visual Studio (version 2010). A detailed description of the parallel evolutionary algorithm is available in Seyoum (2015).

3. RESULTS AND DISCUSSION

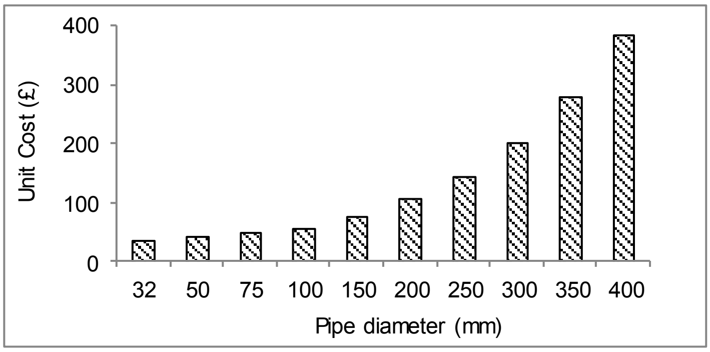
The example considered was one of the water supply zones of a network in the UK (Seyoum and Tanyimboh 2014, 2016). The network and pipe cost data were obtained from the water utility. The system comprised 251 pipes, 199 demand nodes, 29 fire hydrants, 5 variable-head supply nodes and three water demand categories including leakage (Kanakoudis and Gonelas 2016, Tabesh *et al.* 2009). The minimum residual pressure required at the demand nodes was 20 m. The minimum residual pressure required at the fire hydrants for a flow rate of 8 litres per second was 3 m. A fire demand was operational in each hour, except for the first and last hours. Extended period simulation for a period of operation of 31 hours with a hydraulic time step of one hour was used. The Darcy–Weisbach formula for the head loss due to friction in the pipes was used (Rossman 2000). Figure 1 shows the network topology and pipe diameter options.

The network was optimised as a new design. Ten commercially available pipe diameters were selected based on the existing network's pipe diameters that range from 32 mm to 400 mm. The 10 candidate pipe diameters correspond to 10^{251} feasible and infeasible solutions in total. The crossover and mutation probabilities were 1.0 and 0.005, respectively. The maximum number of function evaluations allowed per optimization run was 1,000,000. Ten optimization runs with different initial populations of 1000 were carried out. A workstation that comprised dual Intel Xeon 2.4 GHz CPU (four cores each) and 16 GB RAM running the Windows operating system was used. The workstation had eight cores in total and all were utilized for the parallel computing.

The least expensive solution achieved by the parallel algorithm had a cost of £418,685 within 975,000 function evaluations. This is the best solution obtained for the network to date. The solution is 0.2% cheaper than the best solution reported in Tanyimboh and Seyoum (2016), i.e. £419,514 within 985,000 function evaluations. Figure 2 shows the pipe diameters of the solutions from the parallel and serial algorithms and existing network. The solution for the serial algorithm was obtained from Tanyimboh and Seyoum (2016). All the solutions shown are fully feasible, i.e. the minimum residual pressure requirements for all the operating conditions were satisfied. Figure 3 illustrates the progress of the algorithm in terms of the least expensive feasible solution in successive generations.



(a) Topology



(b) Pipe diameters selected for the optimization

Figure 1. Network topology and pipe diameter options. R1-R5 represent the supply nodes.

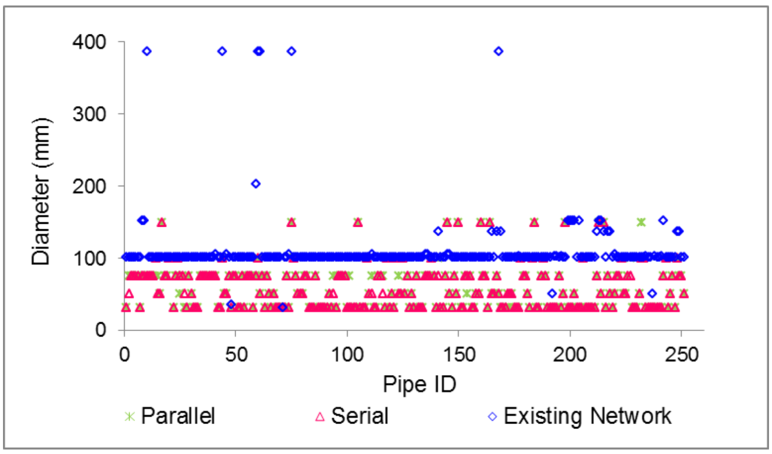


Figure 2. Comparison of the existing and optimised pipe diameters

A single optimization run with 1,000,000 function evaluations using the serial algorithm required an average CPU time of 30 days while the parallel algorithm required only two days on average. Thus, the parallel algorithm achieved a considerable improvement in the execution time. The speedup achieved by the parallel algorithm ranged from 10.96 to 17.15 as shown in Figure 4. On average, the parallel algorithm achieved a speedup of 15; in other words, the CPU time for the parallel algorithm was approximately 6.7 % of the CPU time for the serial algorithm.

4. CONCLUSIONS

The results showed that the parallel algorithm developed found optimal and near-optimal solutions reliably and efficiently; the solutions were consistently competitive and a new best solution was achieved. The number of solutions analysed before finding the Pareto-optimal solutions was extremely small in comparison to the total number of solutions available.

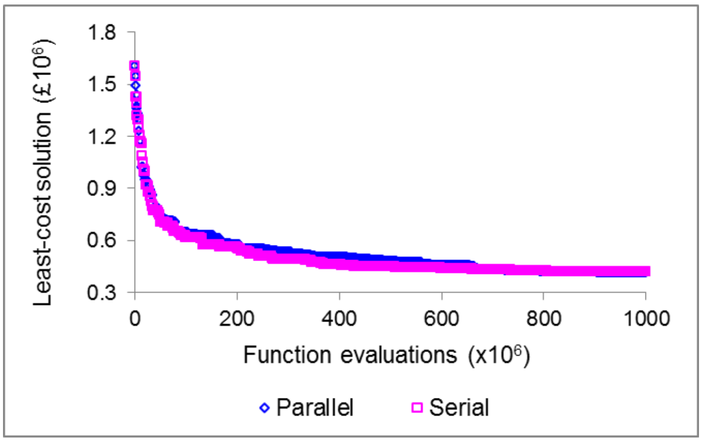


Figure 3. Illustration of the progress of the least expensive feasible solution in successive generations. The graphs show the least-cost feasible solutions for the serial and parallel algorithms, using the same initial population.

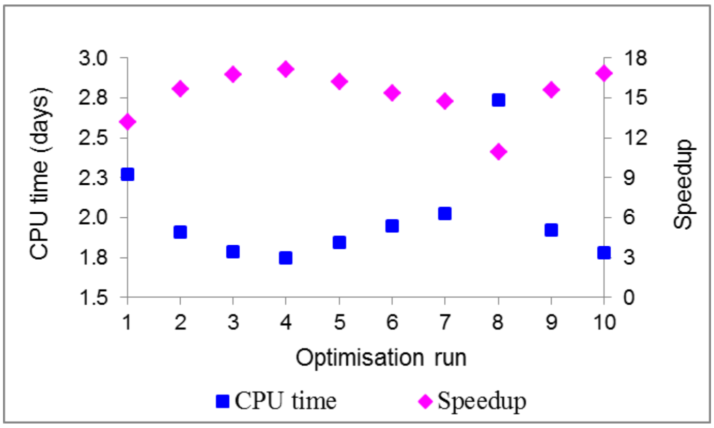


Figure 4. The CPU times and speedup achieved by the parallel algorithm

The average speedup achieved on an eight-core workstation was 15, based on an optimization problem with hundreds of decision variables for a real-world water distribution system. In other words, the average execution time for the parallel algorithm was approximately 6.7 % of the execution time for the serial algorithm. Though highly effective, the improvement achieved previously by search space reduction was smaller (Kadu *et al.* 2008, Siew *et al.* 2014). Thus, the results reinforce previous evidence in the literature that fitness evaluation is often the most computationally expensive step in evolutionary optimization algorithms (Alba and Tomassini 2002, Schutte *et al.* 2004). Hence, the results seem encouraging and further studies with more advanced approaches (e.g. the island model) may be worth considering as parallel computing has the added benefit that it can be used together with other strategies that can improve the computational efficiency further. The network considered comprised one water supply zone of a water distribution system in the UK. Therefore, additional investigations regarding the speed and other properties of the proposed parallel algorithm in the context of optimization problems with significantly more decision variables may be worth considering.

ACKNOWLEDGMENT

This project was funded in part by the UK Engineering and Physical Sciences Research Council (EPSRC Grant Reference EP/G055564/1), the British Government (Overseas Research Students Awards Scheme) and the University of Strathclyde. The authors wish to thank Veolia Water UK (now Affinity Water) for their invaluable contribution and support.

REFERENCES

Alba, E., Tomassini, M., 2002. Parallelism and evolutionary algorithms. IEEE Transactions on Evolutionary Computation; 6(5): 443-462.

Barlow, E., Tanyimboh, T. T., 2014. Multi-objective memetic algorithm applied to the optimisation of water distribution systems. Water Resources Management; 28(8): 2229-2242.

Broad, D. R., Dandy, G. C., Maier, H. R., 2005. Water distribution system optimization using metamodels. Journal of Water Resources Planning and Management; 131(3): 172-180.

Cantu-Paz, E., Goldberg, D. E., 2000. Efficient parallel genetic algorithms: theory and practice. Computer Methods in Applied Mechanics and Engineering; 186(2): 221-238.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multi-objective genetic algorithm: NSGA II. IEEE Transactions on Evolutionary Computation; 6(2): 182-197.

Ghebremichael, K., Gebremeskel, A., Trifunovic, N., Amy, G., 2008. Modelling disinfection by-products: coupling hydraulic and chemical models. Water Science and Technology: Water Supply; 8(3): 289-295.

Haghighi, A., Samani, H. M., Samani, Z. M., 2011. GA-ILP method for optimization of water distribution networks. Water Resources Management; 25(7): 1791-1808.

Kadu, M. S., Gupta, R., Bhave, P. R., 2008. Optimal design of water networks using a modified genetic algorithm with reduction in search space. Journal of Water Resources Planning and Management; 134(2): 147-160.

Kang, D., Lansey, K., 2012. Revisiting optimal water-distribution system design: issues and a heuristic hierarchical approach. J. Water Resour. Plann. Manage; 138(3): 208-217.

Kanakoudis, V., Gonelas, K., 2016. Analysis and calculation of the short and long run economic leakage level in a water distribution system. Water Utility Journal; 12:57-66.

Moosavian, N., Lence, B. J., 2016. Nondominated sorting differential evolution algorithms for multiobjective optimization of water distribution systems. J. Water Resources Planning and Management; 10.1061/(ASCE)WR.1943-5452.0000741.

Ostfeld, A., Tubaltzev, A., 2008. Ant colony optimization for least-cost design and operation of pumping water distribution systems. J. Water Resour. Plann. Manage.; 134(2): 107-118.

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.

Rossman, L. A., 2000. EPANET 2 users manual. Water Supply and Water Resources Division. National Risk Management Research Laboratory, US EPA, Cincinnati.

Schutte, J. F., Reinbolt, J. A., Fregly, B. J., Haftka, R. T., George, A. D., 2004. Parallel global optimization with the particle swarm algorithm. International J. Numerical Methods in Engineering; 61(13): 2296-2315.

Seyoum, A. G., 2015. Head-dependent modelling and optimisation of water distribution systems. PhD thesis, University of Strathclyde, UK.

Seyoum, A. G., Tanyimboh, T. T., 2014. Pressure dependent network water quality modelling. Proceedings of ICE: Water Management; 167(6): 342-355.

Seyoum, A. G., Tanyimboh, T. T., 2016. Investigation into the pressure-driven extension of the EPANET hydraulic simulation model for water distribution systems. Water Resources Management; DOI:10.1007/s11269-016-1492-6.

Siew, C., Tanyimboh, T. T., 2012a. Pressure-dependent EPANET extension. Water Resour Manag; 26(6): 1477-1498.

Siew, C., Tanyimboh, T. T., 2012b. Penalty-free feasibility boundary convergent multi-objective evolutionary algorithm for the optimization of water distribution systems. Water Resour Manag; 26(15): 4485-4507.

Siew, C., Tanyimboh, T. T., Seyoum, A. G., 2014. Assessment of penalty-free multi-objective evolutionary optimization approach for the design and rehabilitation of water distribution systems. Water Resour Manag; 28(2): 373389.

Siew, C., Tanyimboh, T. T., Seyoum, A. G., 2016. Penalty-free multi-objective evolutionary approach to optimization of Anytown water distribution network. Water Resour Manag; 30(11): 3671-3688.

Sivakumar, P., Prasad, R. K., 2015. Extended period simulation of pressure-deficient networks using pressure reducing valves. Water

Resour Manag; 29(5): 1713-1730.
Spiliotis, M., Tsakiris, G., 2011. Water distribution system analysis: Newton-Raphson method revisited. J. Hydraul. Eng. ASCE; 137(8): 852-85.
Tabesh, M., Yekta, A. H. A., Burrows, R., 2009. An integrated model to evaluate losses in water distribution systems. Water Resour Manage; 23(3): 477-492.
Tanyimboh, T. T., Setiadi, Y., 2008. Sensitivity analysis of entropy-constrained designs of water distribution systems. Engineering Optimization; 40(5): 439-457.
Tanyimboh, T. T., Seyoum, A. G., 2016. Multiobjective evolutionary optimization of water distribution systems: Exploiting diversity with infeasible solutions. J. Environ Manag; 183:133-141, DOI:10.1016/j.jenvman.2016. 08.048.
Todini, E., Rossman, L., 2013. Unified framework for deriving simultaneous equation algorithms for water distribution networks. J. Hydraulic Engineering; 10.1061/(ASCE)HY.1943-7900.0000703, 139(5): 511-526.
Trobec, R., Vajtersic, M., Zinterhof, P., 2009. Parallel computing: numerics, applications, and trends. Springer Science & Business Media.
Tsakiris, G., Spiliotis, M., 2014. A Newton-Raphson analysis of urban water systems based on nodal head-driven outflow. European J. Environmental and Civil Eng.; 18(8): 882-896.
Vairavamoorthy, K., Ali, M., 2000. Optimal design of water distribution systems using genetic algorithms. Computer-Aided Civil and Infrastructure Engineering; 15(5): 374-382.
Van Zyl, J. E., Savic, D. A., Walters, G. A., 2004. Operational optimization of water distribution systems using a hybrid genetic algorithm. J. Water Resour. Plann. Manage; 130(2): 160-170.