13th Computer Control for Water Industry Conference, CCWI 2015

Sequence Analysis-based Hyper-heuristics for Water Distribution Network Optimisation

Ahmed Kheiri*, Edward Keedwell, Michael J Gibson, Dragan Savic

College of Engineering. Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter, EX4 4QF, UK

Abstract

Hyper-heuristics operate at the level above traditional (meta-)heuristics that 'optimise the optimiser'. These algorithms can combine low level heuristics to create bespoke algorithms for particular classes of problems. The low level heuristics can be mutation operators or hill climbing algorithms and can include industry expertise. This paper investigates the use of a new hyper-heuristic based on sequence analysis in the biosciences, to develop new optimisers that can outperform conventional evolutionary approaches. It demonstrates that the new algorithms develop high quality solutions on benchmark water distribution network optimisation problems efficiently, and can yield important information about the problem search space.

Keywords: Hyper-heuristic; Water Distribution Network; Hidden Markov Model

1. Introduction

A wide variety of meta-heuristic algorithms have been applied to the problem of water distribution network optimisation. Evolutionary algorithms [1,10] remain the most popular methods although ant colony optimisation [2], particle swarm optimisation [3] and shuffled leapfrog complex algorithms [4] have also been applied. These meta-heuristic methods have generally been successful in optimising many aspects of water distribution network design and operation due to their ability to be used as off-the-shelf techniques. Network design, rehabilitation and calibration have all been tackled along with pump scheduling as the main operational problem to be solved. As meta-heuristics, each of these methods has a fixed set of operations that are performed during the optimisation, (e.g.

---

* Corresponding author.
  E-mail address: A.Kheiri@ex.ac.uk

crossover and mutation in evolutionary algorithms). These fixed processes are often inspired by natural or other phenomena that have demonstrated success in the real-world and thus are used in computational optimisation. However, there has been a recent move towards higher level optimisation through multi-method search and hyper-heuristics. Multi-method search runs several algorithms in parallel and dynamically allocates computational resources to the most successful techniques during the optimisation. This method provides a diversity of approach and is well suited to modern multi-core machines where each of the methods can be executed in parallel. Hyper-heuristics similarly, operate above the level of (meta-)heuristics, but do so by *selecting* and *generating* low level heuristics (LLHs) which are similar to operators in meta-heuristics and perturb solutions in a variety of different ways. There are two primary classes of hyper-heuristics, selection hyper-heuristics and generation hyper-heuristics. Selection hyper-heuristics are provided with a set of LLHs to control and the task for the selection approach is to determine which LLH to apply at a given point during the optimisation. Generation hyper-heuristics build new heuristics from a set of components. Due to their ability to adapt to particular problem characteristics by selecting the most appropriate LLH at a given point in the optimisation, hyper-heuristics have been shown to improve on meta-heuristics in a number of different fields, but particularly in operational research (e.g. scheduling, timetabling and resource management) [5]. Improved performance is also possible by incorporating problem-specific and human (engineer)-derived heuristics into to the optimisation process. In this paper we investigate the use of a new sequence-based hyper-heuristic for the optimisation of the New York Tunnels water distribution network rehabilitation problem. The method is shown to find the best known solution within relatively few objective function calculations and a detailed analysis of the hyper-heuristic reveals information on how the method solves the problem. The work points the way towards the use of hyper-heuristics as the method of choice for this optimisation.

1.1. Water Distribution Network Design/Rehabilitation Problem

Water distribution network design/rehabilitation is an important real-world application for optimisation techniques. These networks deliver fresh drinking water from reservoirs, tanks and water treatment works to consumers via a network of pipes and make use of a pumps and valves to meet the demand of consumers. Typically, the optimisation of these networks aims to design new networks or rehabilitate 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, there are many other objectives that can be considered including the minimisation of water age, adherence to velocity and pressure constraints and increasing the robustness of the network to reduce the potential for supply outages. In this particular problem set, only the simplest 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. Problem 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 [6], 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:

$$cost = \sum_{i=1}^{k}(1.1d_i^{1.24} \times l_i) \tag{1}$$

Where *i* represents one of the total number of pipes *k* in the WDN, and *d* represents 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:

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

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.

1.3. Selection Hyper-heuristics

A large number of hyper-heuristics have been presented in the literature since their development in the early 2000s. The simplest possible selection hyper-heuristic is 'simple random' where LLHs are selected at random throughout the optimisation and does not incorporate any learning in the selection process. This algorithm is often used as a baseline from which to compare other more complex selection hyper-heuristics. Popular learning selection hyper-heuristics include the 'TSroulettewheel' which augments the random selection with a process for learning the utility of the LLH during the optimisation gained from the ability for that LLH to generate improving solutions. Other popular methods include the choice function and reinforcement learning and the reader is directed to [5] for a survey of hyper-heuristic approaches. The approach used in this paper is known as the sequence-based selection hyper-heuristic [7] and effectively uses a hidden Markov model to manage the process of determining the transition between heuristics and the selection of acceptance strategy. The approach has been shown to work well for a number of problems and is now applied to water distribution network design and rehabilitation.

1.4. Water Distribution Network Optimisation using Hyper-heuristics

There is very little research relating to the optimisation of water distribution networks using hyper-heuristics. McClymont et. al. [8] studied the use of a markov-chain based hyper-heuristic for multi-objective optimisation of water distribution networks with discolouration risk and Raad [9] used the AMALGAM multi-method search approach for optimising water distribution network design. Both these approaches have focused on the use of hyper-level methods for multi-objective rather than single objective optimisation as is shown here.

1.5. Sequence Analysis Based Selection Hyper-heuristic

A detailed description of this algorithm can be seen in [7] and only an overview is provided here. The sequence analysis based approach using a hidden Markov model (HMM) as its foundation where the states are defined as low level heuristics at each point in the optimisation process. The optimisation process is essentially treated as a sequence of moves in the search space, with attendant changes in the performance of the solution. Elements of the sequence are the application of LLHs to the optimisation problem and the acceptance decisions are made on the basis of the move acceptance criteria.

A matrix of state transition probabilities exist for the transition between LLHs and emission probabilities exist which determine which acceptance strategy to use. One model applies a selected heuristic to a candidate solution without being evaluated and the other accepts strictly better solutions and worse solutions with some small threshold. The former allows the algorithm to explore the space with no optimality criteria whereas the latter

applies optimality as a criterion for acceptance. Uniform probabilities are assigned to the transition matrix and emission probabilities and are updated only when a new best solution is found by the algorithm. When this occurs, the probability of application of the sequence of LLHs and acceptance strategies that led to the new best solution are increased. The selection of LLHs to apply in the next timestep is then determined by the Viterbi algorithm which will decide which LLH to apply next given the previous LLH used. The power of the algorithm arises from the learning process not just in terms of transition from one LLH to another, but also from the ability to choose more exploratory or exploitative modes through the acceptance strategy.

An example of how the method works is illustrated in Figure 1 and has been shown to work exceptionally for benchmarks [7] from the operational research community, but has not previously been trialled on real-world problems.
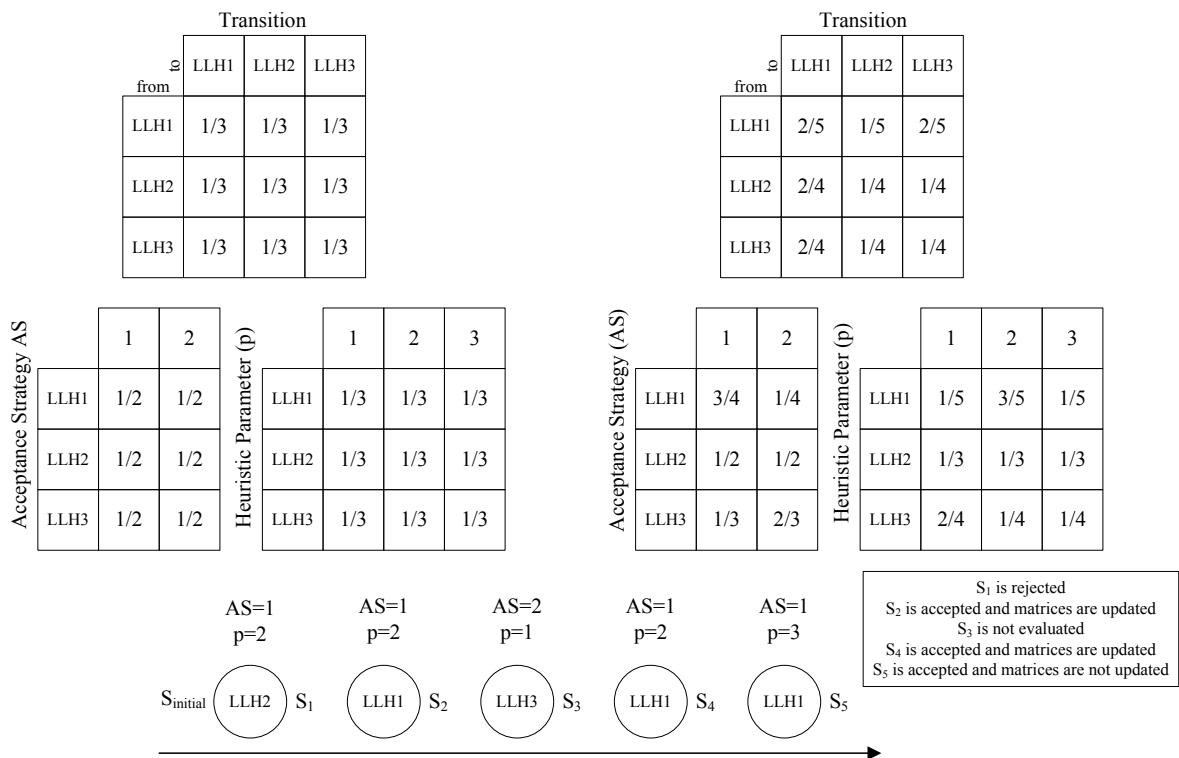
Transition

from \ to	LLH1	LLH2	LLH3
LLH1	1/3	1/3	1/3
LLH2	1/3	1/3	1/3
LLH3	1/3	1/3	1/3

Transition

from \ to	LLH1	LLH2	LLH3
LLH1	2/5	1/5	2/5
LLH2	2/4	1/4	1/4
LLH3	2/4	1/4	1/4

Acceptance Strategy AS

	1	2
LLH1	1/2	1/2
LLH2	1/2	1/2
LLH3	1/2	1/2

Heuristic Parameter (p)

	1	2	3
LLH1	1/3	1/3	1/3
LLH2	1/3	1/3	1/3
LLH3	1/3	1/3	1/3

Acceptance Strategy (AS)

	1	2
LLH1	3/4	1/4
LLH2	1/2	1/2
LLH3	1/3	2/3

Heuristic Parameter (p)

	1	2	3
LLH1	1/5	3/5	1/5
LLH2	1/3	1/3	1/3
LLH3	2/4	1/4	1/4

S_1 is rejected
S_2 is accepted and matrices are updated
S_3 is not evaluated
S_4 is accepted and matrices are updated
S_5 is accepted and matrices are not updated

AS=1 p=2   AS=1 p=2   AS=2 p=1   AS=1 p=2   AS=1 p=3

$S_{initial}$ (LLH2) S_1 (LLH1) S_2 (LLH3) S_3 (LLH1) S_4 (LLH1) S_5

Fig. 1. Sequence-based selection hyper-heuristic operates on three low level heuristics each with three possible heuristic parameters

1.6. Parameter Settings

Unlike evolutionary algorithms, there are very few parameters to set for the sequence analysis based approach described above. The only parameter that is required to be set is the threshold T at which worse solutions are accepted in the move acceptance method. This has been set at 0.03 of the cost of the best recorded solution in hand, for the following experimentation. As a single solution algorithm, there is no population size or selection parameters to set. Although some of the LLHs use a parameter to influence their behaviour (e.g. such as a mutation rate) these are learned by the algorithm during optimisaton.

1.7. Objective Function

The objective function is simply calculated as in equation (3) above where alpha is set to 100,000,000 to ensure that the discovered solutions are feasible. The head deficit is calculated as in equation (2) and the cost as per equation (1).

1.8. Low Level Heuristics

For a selection hype-heuristic to work effectively, a number of low level heuristics (LLHs) must be provided for the method to control. The LLHs used in this work are as follows:

LLH0: change only one pipe diameter randomly
LLH1: swap two pipe diameters at random
LLH2: increase or decrease a randomly selected pipe diameter by one pipe size.
LLH3: 'ruin' several pipes and rebuild randomly. The number of pipes to be changed is a parameter (P) that takes a value in the range [1,5] (inclusive).
LLH4: shuffle several pipes (i.e. makes several swaps). Again, the number of pipes to be changed is a parameter (P) that takes a value in the range [1,5].

1.9. New York Tunnels Water Distribution Network

The New York Tunnels water distribution network is a well known benchmark rehabilitation problem consisting of 21 pipes and 16 possible diameters. The task is to determine the additional pipework required to meet a new set of head demands across the network. Figure 2 shows a schematic of the network. A number of near-optimal solutions have been proposed in the literature depending on the setting of the Hazen Williams coefficient, but the generally accepted best solution for this network under standard conditions is 38.64m dollars with zero head deficit.

Figure 1 - New York Tunnels problem schematic

2. Results

2.1. Function Evaluations

The sequence-analysis based method was applied to the New York Tunnels problem for 10 trials, and the termination criterion is set to 20,000 evaluations. The proposed method discovered the result of $38.64m cost and zero head deficit in 7 out of 10 runs and obtained $38.81 cost with zero head deficit in the three other runs. These results are obtained within an average of 10793.7 objective function evaluations which compares favourably with differential evolution (SDE, DDE), and much cheaper than evolutionary algorithm formulations (ALCO-GA, CGA and SGA) as shown in Table 1. The table below provides the mean number of evaluations (out of 10 runs for our experiments) required to obtain the best solution:
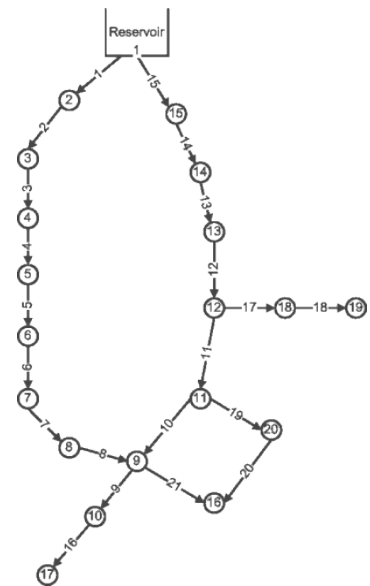
Table 1. An indication of the range of evaluations required for each algorithm type

Algorithm	Best Solution Cost	Evaluations
Sequence-Analysis Hyperheuristic	**38.64m**	**10,794**
CGA [10]	38.64m	44,324
SGA [10]	38.64m	54,789

| SDE [10] | 38.64m | 12,855 |
| DDE [10] | 38.64m | 13,214 |

The performance of the sequence-based selection hyper-heuristic is shown to be markedly better than the modified evolutionary algorithm approaches and offers small improvements over the differential-evolution approaches. The performance of the hyper-heuristic is of course dependent to a certain degree on the LLHs that it has access to and further work investigate the use of differential evolution-type operators to potentially improve performance, although a population based approach would then be necessitated as the differential evolution requires a population.

2.2. LLH Usage Statistics

A key benefit of using the sequence-analysis based approach is that the probability matrices are available for scrutiny at any point in the optimisation process. This can show the reliance of an algorithm on a particular LLH or LLH type and can reveal trends of LLH usage over time. These statistics can then be used to analyse the performance of the algorithm and to better understand the characteristics of the underlying search space as seen by the low level heuristics.

Figure 3 shows the probability of selection of each of the low level heuristics aggregated over the optimisation run. This reveals that LLH2, the pipe increment/decrement, is used mostly during the optimization run and approximately 70% of the time. LLH0, LLH1 and LL4 are the next most used and are used approximately equally. The ruin and recreate approach of LLH3 is used only very sparingly, as would be expected of such an operator with the potential to radically change solutions.
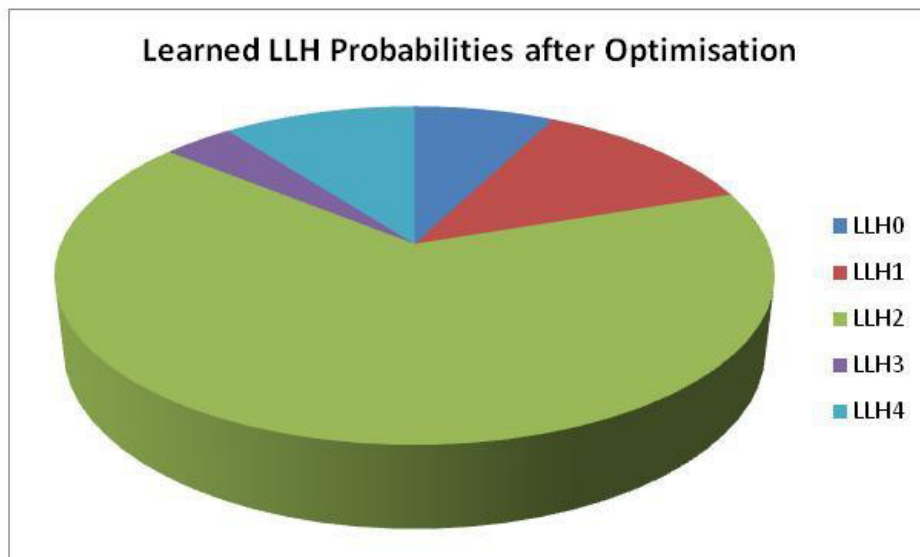

Figure 3 - Learned LLH Probabilities after Optimisatione

2.3. Transition Probabilities

Figure 4 shows the probability of moving between the various LLHs used in the optimisation and shows the probability of moving to the LLH shown on the X-Axis. This shows how the algorithm is deciding which LLHs to select and highlights pairs of LLHs that work well together. For instance LLH2 has strong connections to itself, and LLH 0, meaning that if either LLH0 or LLH2 have been executed in the last application there is a good chance that the next LLH will be LLH0. LLH3 has predictably very small probabilities which is reflected in its usage characteristics.
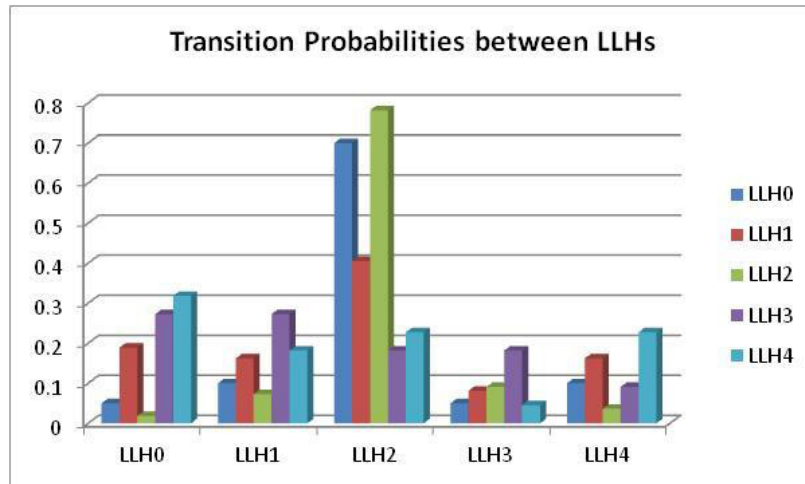

Figure 4 - Probability of transition between LLHs

2.4. Acceptance Strategies

The LLHs will generate a variety of new solutions, but a key part of an algorithm is whether it decides to accept or reject that solution. This approach has two acceptance strategies built in, and the probabilities of using them subsequent to the application of each LLH can be visualised.
Figure 5 shows that the successful LLH2 utilises the exploitation acceptance strategy frequently, meaning that the solutions generated by this LLH are accepted as the next solution providing that they meet optimality criteria.
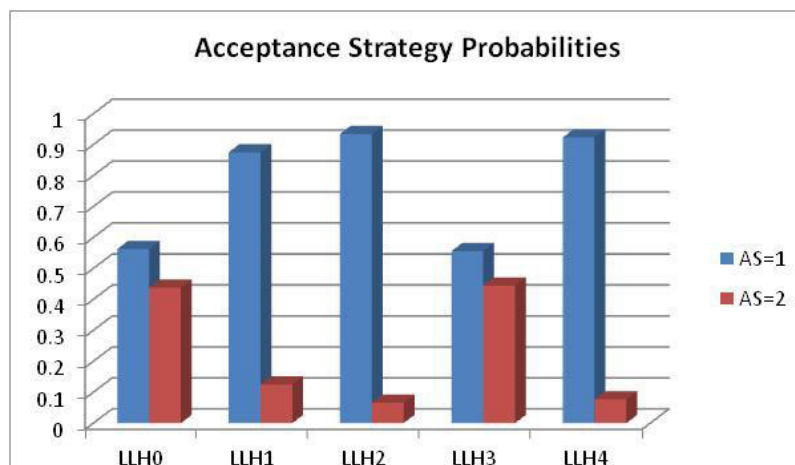

Figure 5 - Probabilities of acceptance strategy use by LLH

LLH0 and LLH3 make more use of the exploration strategy and so their generated solutions are not checked for optimality before being accepted and therefore represent the exploratory elements of the algorithm.   This combination suggests that LLHs 1,2 & 4 are being used predominantly as a greedy approach, making small changes to solutions (by swapping existing material for LLHs 1 & 4 and by incrementing for LLH 2) and applying optimality criteria before new solutions are accepted.   LLHs 0 and 3 make more random perturbations and by using the exploratory acceptance strategy introduce new material and thus explore the space of possible solutions.
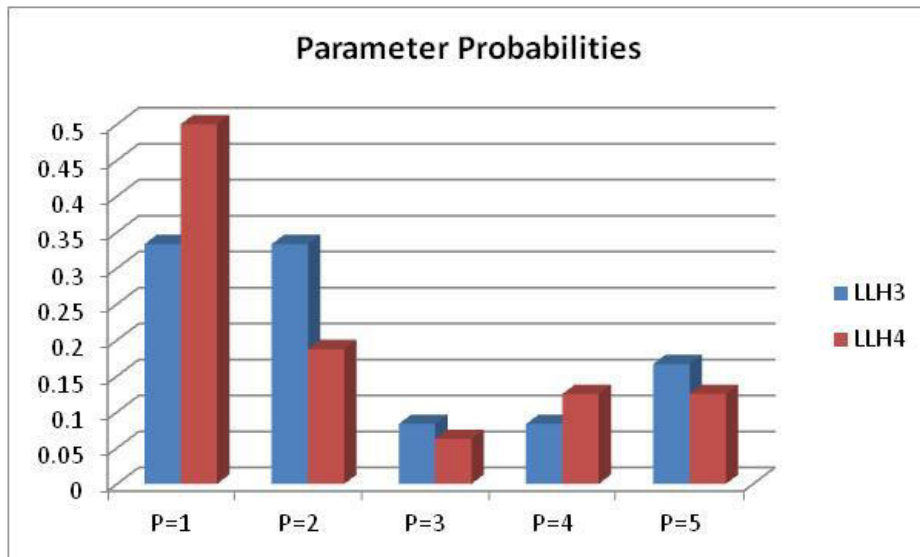
2.5. Parameter Probabilities

Figure 6 - Probabilities of parameter selections for LLH3 and LLH4

LLHs 3 and 4 each have a parameter associated with them, the number of times per application that the solution are 'ruined and recreated' and the number of shuffles respectively. As expected the value of 1 is most popular for both LLHs as this is a smaller perturbation of an existing solution, although for LLH3, a parameter value of 2 is also very likely. Both LLHs show a decrease to a parameter value of 3 which has minmimal probability, but with increasing probabilities for higher values of the parameter setting. It is not entirely clear why this should be the case, making higher numbers of swaps and ruin and recreate operations is likely to result in significantly modified solutions and so these moves will be highly exploratory in nature. One possible hypothesis is that the algorithm is using these operations periodically to escape from local minima encountered as a result of the application of the greedy approach described in the previous section.

3. Conclusions

A sequence analysis based hyperheuristic based on a hidden Markov model has been successfully applied to the New York Tunnels benchmark problem. The method is shown to be highly competitive from a computational perspective and to deliver interesting information as to the best low level heuristics to use on this problem. The results show that as expected, methods that make large scale and potentially destructive moves are used sparingly in comparison with more usual small move mutation methods. They also show that the system evolves a heavy reliance on an LLH (2) that makes small adjustments to existing solutions through the incrementing and decrementing of pipe sizes, but resorts occasionally to more destructive operations..   Finally the results also show

that the system is able to evolve probabilities of move acceptance and of parameterization of LLHs that require this during the optimisation.

Although the work presented here is largely theoretical, this approach will be attractive to those working in the field of real-world network optimisation for a number of reasons. Firstly, the ability of the algorithm to learn how best to traverse the topology of the search space means that it can find better solutions more quickly than other algorithms, secondly through the various matrices that it presents, it is able to convey information gained about the search space over time leading to a better understanding of the problem. The final key characteristic is the system's flexibility. Low level heuristics can be generated by engineers, incorporated by the optimisation process along with standard mutation and other perturbation heuristics to improve the solutions created by the algorithm. The use of hyperheuristics in this way will potentially open doors towards a more collaborative approach to optimisation where algorithm and engineer work together to solve difficult problems in water systems research.

Acknowledgements

References

[1] Savic, D. A., & Walters, G. A. (1997). Genetic algorithms for least-cost design of water distribution networks. *Journal of water resources planning and management*, *123*(2), 67-77.

[2] Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y., & Tan, C. L. (2003). Ant colony optimization for design of water distribution systems. *Journal of water resources planning and management*,*129*(3), 200-209.

[3] Montalvo, I., Izquierdo, J., Pérez, R., & Tung, M. M. (2008). Particle swarm optimization applied to the design of water supply systems. *Computers & Mathematics with Applications*, *56*(3), 769-776.

[4] Eusuff, M. M., & Lansey, K. E. (2003). Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of Water Resources Planning and Management*, *129*(3), 210-225.

[5] Burke, E. K., Gendreau, M., Hyde, M., Kendall, G., Ochoa, G., Özcan, E., & Qu, R. (2013). Hyper-heuristics: A survey of the state of the art. *Journal of the Operational Research Society*, *64*(12), 1695-1724.

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

[7] Kheiri A, Keedwell E. A sequence-based selection hyper-heuristic utilising a hidden Markov model, accepted to *Conference on Genetic and Evolutionary Computation (GECCO '15)*.

[8] McClymont, K., Keedwell, E., Savić, D., & Randall-Smith, M. (2013). A general multi-objective hyper-heuristic for water distribution network design with discolouration risk. *Journal of Hydroinformatics*, *15*(3), 700-716.

[9] 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.

[10] Zheng, F., Simpson, A. R., & Zecchin, A. C. (2012). A performance comparison of differential evolution and genetic algorithm variants applied to water distribution system optimization. In *World Environmental & Water Resources Congress (EWRI 2012), Albuquerque, NM*.