Knowledge-based neurocomputing for operational decision support

Andrzej Bargiela[(1)], Corneliu Arsene[(1)], Masahiro Tanaka[(2)]

[(1)]Intelligent Simulation and Modelling
The Nottingham Trent University
Burton Street, Nottingham NG1 4BU, UK
Tel: +44(0)1158486016 – Fax: +44(0)1158486518
Email: andre@doc.ntu.ac.uk, Corneliu.Arsene@ntu.ac.uk

[(2)]Intelligent Systems Laboratory
Konan University
8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan
Tel: +81 (0)78 435 2527 -- Fax +81 (0)78 435 2540
Email: m_tanaka@konan-u.ac.jp

Abstract. This paper discusses a neurocomputing system for operational decision support in water distribution networks. An analog neural network is used to calculate 'loop-equations based' state estimates. This is followed by the Confidence Limit Analysis (CLA) of the calculated estimates and the General Fuzzy Min-Max neural pattern classification/clustering (PC). The latter emulates the process of experience-building by human operators of water systems. We refer to the resulting neurocomputing system as CLA/PC.

Water distribution systems are representative of a large and important class of systems that are primarily driven by external, incompletely defined stimuli yet, the operation of which needs to be optimised according to some well defined criteria. The operational control of such systems presents considerable challenge to operators who need to develop the ability to `distil' the overall system state from a large number of fuzzy system snapshots (measurements and estimates) in order to decide on the appropriate control action. The software has been applied to a real-life water distribution networks and the representative subset of the results is provided.

1. Introduction

The operational control of water systems (and indeed many real-life systems) is a challenging task because it requires that operators develop a mental model of operation of a large-scale non-linear system which is subject to random disturbances (fluctuations of consumption) and which is monitored using relatively few measurements. The non-linearity of the system makes the outcome of every control action hugely dependent on the current operating state and the inaccuracy (or unavailability) of the various measurements compounds the difficulty of system control.

Given that the critical operating states (i.e. those that pose health or safety risks) occur relatively infrequently, the process of developing satisfactory level of expertise by system operators is quite lengthy and is dependent on the favourable combination of circumstances. It is for this reason that simulations have long been recognised as offering invaluable opportunity to generate scenarios that allow operators to be trained in the context of critical system states. One of the first real-time simulation systems for water industry was reported in [13], [2], [4] and was successfully utilised for the purpose of developing intuitive understanding of the variabilities of flows and pressures in water systems in a broad spectrum of operating conditions. The experience gained by trainee operators through interactive simulation was deemed to be comparable to the understanding gained through extensive experimentation, involving opening fire hydrants at various network location, as documented in [6].

The CLA/PC software system described here is largely based on the original TCLAS software [3] and it preserves its modular design, the realistic modelling of noisy and erroneous measurement data, the ability to modify metering configuration, the ability to switch between the various state estimation procedures in real-time, the ability to evaluate Confidence Limit intervals on state variables and the flexibility of graphical user interface. The new software incorporates all these features and introduces additional modules that provide computationally efficient neural network based Confidence Limit Analysis, [7], and system state classification and clustering, [10], which renders it applicable to large-scale systems. The overall software architecture is presented in Section 2.

It is important to emphasise that unlike in off-line applications, such as optimal network design or pump schedule optimisation, in on-line decision support the challenge is to estimate the system state in real time and preferably an order of magnitude faster so as to allow the `what-if' type of study by operators. The neural state estimation module provides an efficient computation engine while ensuring that all of the meters are used to their full advantage. This set of facilities is described in section 3.1.

One must recognise however that, although the mathematical model built into the state estimation may be accurate, the state estimates are based on input data that contain significant amount of uncertainty. This uncertainty has an impact on the accuracy with which the state estimates can be calculated. It is important, therefore that the system operators are given not only the values of flows and pressures in the network at any instant of time but also that they have some indication of how reliable these values are. The procedure for the quantification of the inaccuracy of the state estimates caused by the input data uncertainty has been developed in the late 80's and has been termed the Confidence Limit Analysis (CLA) [3]. Rather than a single deterministic state estimate, the CLA enables the calculation of a set of all feasible states corresponding to a given level of measurement uncertainty. The set is presented in the form of upper and lower bounds for individual variables, and hence provides limits on the potential error of each variable.

Once the pressure and flow estimates together with their confidence limits (interval estimates) are found it becomes possible to validate the measurements and the network topology. In this paper a neuro-fuzzy recognition system, utilising information from the state estimation and confidence limit analysis procedures, is proposed as a means of fault detection and identification in water distribution networks. The general fuzzy min-max (GFMM) neural network for clustering and classification [9] combines the ability of fuzzy systems to cope with uncertain and ambiguous data with the computational efficiency, learning and pattern recognition ability of neural nets. The neuro-fuzzy recognition approach to fault detection and identification is described in Section 3.2.

2. The simulation software

The original modular design of TCLAS software [2] and [4] has shown to offer flexible framework for the introduction of additional functionality associated with operational decision support. The main computational modules of the integrated CLA/PC environment are presented in Figure 1.

At the root of the package there is a **Network Simulation** module whose role is to mimic the actual behaviour of the water distribution network. In this module, mathematical models of elements like pipes, valves, pumps etc. combine exact information about the topology and the physical parameters of the network (such as length, diameter, c-values etc.) with the deterministic values of inflows and consumptions and allow calculation of network's pressures and flows. In doing so, the network simulation module provides a facility for pressure/flow study of the network without recourse to physical experimentation. Using this module, leakages or pipe blockages can be simulated by updating the topology information rather than opening hydrants or opening/closing flow control valves.

The **Telemetry Simulation** module is intended to instil a degree of realism into the simulated readings of flows and pressures. Although the network simulation results are accurate in mathematical sense, it is quite clear that they reflect the idealised situation where unlimited number of measurements are available and the metering devices are of absolute accuracy. In practice neither is true so it is necessary to be explicit about meter positioning and to model both Gaussian and gross measurement errors. Indeed by doing so one can explore the relationship between the number, accuracy and the positioning of meters and the quality of subsequent state estimations. The telemetry simulation module can also be used to switch between the actual and simulated measurements so that the software system can be adapted to on-line operation.

Network Simulation
V
Exact flows and pressures
V
Telemetry simulation
V
Realistic flows and pressure measurements (with noise)
V
State estimation with CLA
V
Estimates of flows and pressures with error bounds
V
Pattern classification
V
Decision support messages

Figure 1. Outline of the software modules and the resulting data sets.

The **State Estimation and Confidence Limits Analysis** module processes the telemetered data in two stages. First it finds the state estimate that minimises the discrepancies between the calculated and the actual meter readings (according to either least-squares or least-absolute-values criterion) and subsequently it quantifies the confidence limits on the calculated states. The combined state estimates with confidence limits provide a more robust indication of system performance in presence of uncertain input data. As it is the case with on-line state estimation, no information about any anomalous event simulated in the simulation/telemetry modules is being made available to the state estimation module.

Having obtained state estimates with confidence limits the **Pattern Classification** module attempts to abstract from the detailed numerical results by categorising the operating state into various sub-classes. This is equivalent to the process of accumulating experience by human operators and the progressively more robust classification provides basis for generating decision support messages. The state classification functionality is particularly valuable in the context of emergency states such as leakages in the distribution network.

An access to simulation/estimation/classification results is provided through a powerful graphical user interface implemented in the Visualisation module. The simulations are executed in cycles that are triggered individually or are set to repeat at pre-specified time interval. Each cycle involves the input data update, the execution of the four modules, and the output data update. The input data update takes into account all the changes (e.g. change of meter configuration, addition of a leakage, etc.) made by the user before the beginning of each new cycle. The output data update is concerned with the change of graphical display and suitable data fields representing the results of state estimation, confidence limit analysis and classification modules

3. Neural system modelling and decision making

3.1 State estimation

The water system state estimation is, in essence, a process of solving an over-determined set of non-linear mass-balance or energy-conservation equations. The non-linearity of these equations means that it is necessary to linearise the equations around the current state and to apply Newton-Raphson iterations to find the solution to the original set of equations.
A computationally efficient neural scheme has been proposed in [7] and is depicted in Fig. 2.

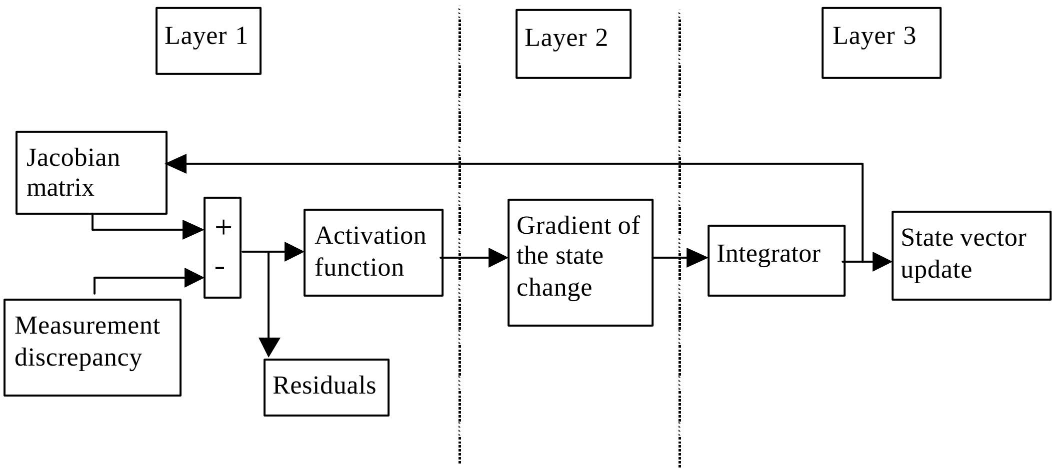

Figure 2. Neural network for solving sets of linear equations in Newton-Raphson iterations.

The detailed form of the Jacobian matrix and the activation function depends on the choice of state variables. Here a set of flows in the spanning tree of the network has been used as state variables. The resulting loop equations gave rise to a robust state estimation [1].

Once the state estimate for the average values of measurements is found, an interval enclosure on the system states is calculated using interval representation of measurements. A direct application of interval arithmetic results in very loose bounding so we adopt here a sensitivity analysis approach that produces much tighter bounds [3], [14]. This stage is referred to as Confidence Limit Analysis (CLA). The state estimates with their confidence limits are now the basic reference for operational decision making.

3.2 Learning of state estimate patterns

The various operational states of the water network result in certain patterns of flows and pressure distribution in the network. For example a leakage in a particular pipe will result in the reduction of water pressure in the neighbourhood of the leaking pipe. Similarly, closing a specific valve will have an effect of re-distribution of the flow that was associated with this valve to the neighbouring pipes. So, in principle, discovering the new patterns of behaviour of the distribution network and creating operational control links for transition between the specific patterns of behaviour is what constitutes development of expertise by the system operators. In this work we adopt the Fuzzy Min-Max clustering/classification approach originally proposed by [11], [12] and generalised in [9]. The neural network that implements the Generalised Fuzzy Min-Max (GFMM) clustering/classification is a three-layer feedforward neural network that grows adaptively to meet the demands of the problem. The network is presented in Figure 3. Each node in the second layer represents a hyperbox fuzzy set. The connections of the first and second layer are the min-max points and the transfer function is the hyperbox membership function. Each node in the third (output) layer represents a class of behaviour of the system. The connections between the nodes of the second and third layer are binary values assuming 1 if the second layer hyperbox fuzzy set is a part of the class represented by the output layer node and 0 otherwise. The output of the third layer node is found by finding the maximum membership value of the constituent second layer hyperbox fuzzy set.

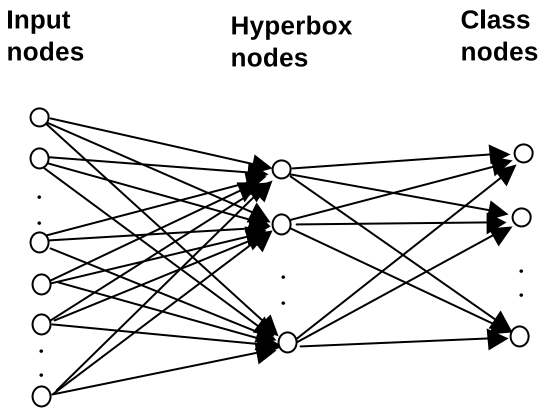

Figure 3. Three-layer neural network implementing GFMM classification/clustering

To improve the efficiency of the learning process the system has been designed as a hierarchical, two-level system. The first level consists of a GFMM neural network, which selects one of the n second level "experts". Input to the first level NN, I_1, is a set of water supplies that characterise the time-of-day period of operation. The second level of the recognition system consists of n NNs. They are called "experts" since each of them is trained using only a part of training set and covers a distinctive part of the 24-hours operational period. The structure of the two-level hierarchical classification is presented in Figure 4. Input to the second level NNs, I_2, comprises all the variables of the input pattern vector which are sensitive to the occurrence of an anomaly (e.g. heads in the load nodes or residuals representing mass balances at water network nodes). The output of the second level NNs is the classification of the water network state denoted by C.

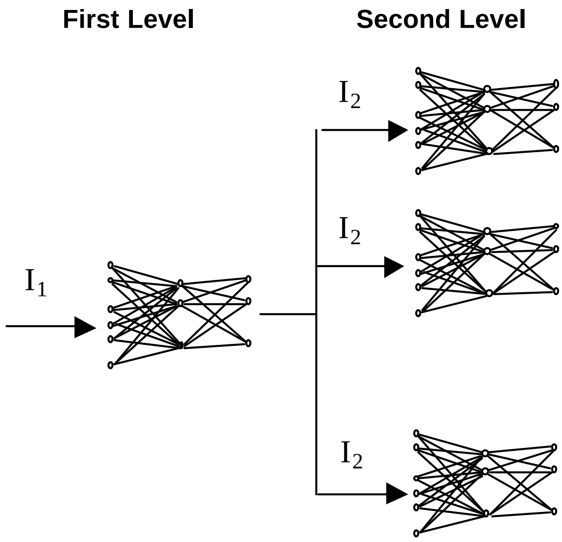

Figure 4. Two-level neural classification scheme.

4. Simulation-based study

The neurocomputing system outlined above has been applied to a real-life water distribution network that has been reduced to 26 nodes and 38 pipes. The 'normal' operating state was defined as 'no leakage in the system' and 38 'special cases' were defined as 'leakage occurring in one of the system pipes'. No combined leakages in two or more pipes were considered. The system state was estimated 10 times for 10 different levels of leakages in each of the 38 system pipes in each of the 24-hour periods. This gave rise to 24*38*10*10=91200 estimations plus 24*10 estimations for the 'normal' operating state.

The results of these estimations were used as a training data for the classification neural network. The trained neural networks have been assessed on random input patterns representing various leakages in the system. The neural system demonstrated a robust performance with almost no misclassifications for leakages of 10 l/s or larger and some 10% misclassification for leakages of 5 l/s or larger.

5. Conclusions

This paper demonstrates how the basic water network simulation program has been integrated with confidence limits analysis and neural classification/clustering modules to deliver decision support functionality to water system operators. The pattern recognition approach to high-level system state diagnosis mimics the human information processing and provides a powerful abstraction tool that reduces the vast amount of detailed numerical information to qualitative description of system state. Although the software environment is primarily intended for operational decision support, it can also be used for operator training purposes. The package has been written entirely in MATLAB environment. The modular structure of this package enables future extensions/modifications of individual components without affecting the integrity of the overall software suite.

References

[1] Arsene, C., Bargiela, A. (2001) "Decision support for forecasting and fault diagnosis in water distribution systems", ICWI,2001, DeMontfort University, Leicester, UK.
[2] Bargiela, A. and D.Al-Dabass (1988) "A simulated real-time environment for verification of advanced water network control algorithms", Int. J. Systems Science, Vol. 14, No. 3, pp. 51-62.
[3] Bargiela, A. and G.Hainsworth (1988) "Pressure and flow uncertainty in water systems", ASCE Journal of Water Resources, Planning and Management, vol.115, no. 2, pp.212-229
[4] Bargiela, A. (1989) "A graphical interactive software tool for confidence limit analysis", Proc. AWWA Computer Specialty Conference, Denver Col., pp.193-204.
[5] Bargiela A. (1998). "Uncertainty - A Key to Better Understanding of Systems", Keynote Lecture, Proc. 10th Eur. Simulation Symposium, ISBN 1-56555-147-8, pp.11-19.
[6] Carpentier, P. and G.Cohen. (1993). "Applied mathematics in water supply network management", Automatica, vol.29, no.5, pp.1215-1250.
[7] Cichocki, A. and A.Bargiela. (1997). "Neural networks for solving linear inequality systems", Parallel Computing, vol.22, no.11, pp.1455-1475
[8] Cichocki, A. and R.Unbehauen. (1992) "Neural networks for solving systems of linear equations and related problems", IEEE Trans. Circuits Syst., vol.39 (Feb.), pp.124-138.
[9] Gabrys, B. and A.Bargiela. (2000) General fuzzy min-max neural network for clustering and classification, IEEE Trans. Neural Networks, Vol. 11, 3, pp.769-783.
[10] Gabrys, B. and A.Bargiela. (1999) "Neural networks based decision support in presence of uncertainties", ASCE Journal of Water Resources Planning and Management, vol. 125, no. 5, pp. 272-280.
[11] Simpson, P.K., (1992), Fuzzy min-max neural networks – Part 1 Classification, IEEE Trans. Neural Networks, vol. 3, pp.776-786.
[12] Simpson, P.K., (1993), Fuzzy min-max neural networks – Part 2 Clustering, IEEE Trans. Fuzzy Systems, vol. 1, pp.32-45.
[13] Sterling, M.J.H and A.Bargiela. (1984). "Minimum norm state estimation for computer control of water distribution systems', IEE Proceedings, vol.131 (March), no.2, pp. 57-63.
[14] Kurzhanski, A.B., Tanaka, M., (1989), On a unified framework for deterministic and stochastic treatment of identification problems, IASA Working Paper, WP-89-13.