Expert Systems with Applications 39 (2012) 13214–13224

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Decision support system for water distribution systems based on neural networks and graphs theory for leakage detection

Corneliu T.C. Arsene *, Bogdan Gabrys [1], David Al-Dabass [2]

School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom

ARTICLE INFO

Keywords:
Decision support system
Operational control of water distribution systems
Loop corrective flows equations
Modeling and simulation
Neural network
Graph theory

ABSTRACT

This paper presents an efficient and effective decision support system (DSS) for operational monitoring and control of water distribution systems based on a three layer General Fuzzy Min–Max Neural Network (GFMMNN) and graph theory. The operational monitoring and control involves detection of pipe leakages. The training data for the GFMMNN is obtained through simulation of leakages in a water network for a 24 h operational period. The training data generation scheme includes a simulator algorithm based on loop corrective flows equations, a Least Squares (LS) loop flows state estimator and a Confidence Limit Analysis (CLA) algorithm for uncertainty quantification entitled Error Maximization (EM) algorithm. These three numerical algorithms for modeling and simulation of water networks are based on loop corrective flows equations and graph theory. It is shown that the detection of leakages based on the training and testing of the GFMMNN with patterns of variation of nodal consumptions with or without confidence limits produces better recognition rates in comparison to the training based on patterns of nodal heads and pipe flows state estimates with or without confidence limits. It produces also comparable recognition rates to the original recognition system trained with patterns of data obtained with the LS nodal heads state estimator while being computationally superior by requiring a single architecture of the GFMMNN type and using a small number of pattern recognition hyperbox fuzzy sets built by the same GFMMNN architecture. In this case the GFMMNN relies on the ability of the LS loop flows state estimator of making full use of the pressure/nodal heads measurements existent in a water network.

1. Introduction

Two broad categories of faults occurring in water distribution systems are considered in this work. The faults because of malfunctioning of transducers and telecommunication equipment which are referred to as the measurement errors. And the faults due to leakages and wrong status of valves, invalidating the system model used in the state estimation which are referred to as the topological errors (Arsene & Bargiela, 2001; Gabrys, 1997; Gabrys & Bargiela, 1999; Arsene, 2004; Arsene, 2011; Carpentier & Cohen, 1993). This paper addresses the topological errors introduced by a leakage in a pipe while the recognition of the wrong status of a valve can be dealt with in a similar way by using the concepts described herein.

It is obvious that in the absence of accurate real measurements, the topological errors not only pose a much greater danger to the safety of water network operation but also are more difficult to locate and eradicate even when reliable and efficient state estimators are available. Depending on the topology of the distribution network and the state estimator used (Gabrys, 1997; Arsene, 2004; Arsene, Bargiela, & Al-Dabass, 2004a; Arsene, Bargiela, & Al-Dabass, 2004b), the topological class of errors form characteristic patterns that can be utilized to classify the state of the water network.

The classification of the state of the water network it has been investigated (Gabrys, 1997; Gabrys & Bargiela, 1999; Gabrys & Bargiela, 2000) in the context of the Least Squares (LS) state estimator based on the nodal heads equations (Gabrys, 1997; Gabrys & Bargiela, 1999). The respective approach for diagnosis of leakages and other operational faults occurring in water networks was based on the examination of patterns of state estimates (i.e. nodal pressures, pipe flows) or LS nodal heads residuals by a General Fuzzy Min–Max Neural Network (GFMMNN) (Gabrys & Bargiela, 1999; Gabrys & Bargiela, 2000). It was shown that both the LS nodal heads state estimates with their confidence limits and the LS nodal heads residuals with their confidence limits can

* Corresponding author. Address: School of Computing and Informatics, Nottingham Trent University, Nottingham NG11 8NS, UK. Tel.: +44 (0)1202 965298; fax: +44 (0) 1202 965314.

E-mail addresses: arsenecorneliu@yahoo.co.uk (C.T.C. Arsene), bgabrys@bournemouth.ac.uk (B. Gabrys), david.al-dabass@ntu.ac.uk (D. Al-Dabass).

[1] Computational Intelligence Research Group School of Design, Engineering & Computing Bournemouth University, Poole House Talbot Campus, Fern Barrow Poole, BH12 5BB, United Kingdom Tel.: +44 (0)1202 965298; fax: +44 (0) 1202 965314.

[2] School of Computing and Informatics, Nottingham Trent University, Clifton lane, Nottingham NG11 8NS, UK.

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.05.080

be successfully used to train the GFMMNN recognition system (Arsene & Bargiela, 2001; Arsene, 2004; Belsito, Lombardi, Andreussi, & Banerjee, 1998).

This paper presents the application of the GFMMNN to the classification of the state of the water distribution system based on patterns of data obtained with a LS loop flows state estimator (Arsene and Bargiela, 2001; Arsene et al., 2004a; Arsene et al., 2004b) and Confidence Limits Analysis (CLA) implemented with the same state estimator (Arsene, 2004; Arsene, 2011; Arsene et al., 2011). The investigation has two aims: first, to build an effective and efficient Decision Support System (DSS) (Bargiela et al., 2002) for fault detection and identification in water networks by using (a) the LS loop flows state estimator, (b) the CLA algorithm based on the same LS loop flows state estimator and (c) the GFMMNN system (Bargiela et al., 2002). The second aim is to compare the novel DSS with the initial system described in Gabrys (1997), Gabrys and Bargiela (1999), Gabrys and Bargiela (2000) and based on the LS nodal heads equations. Recently in other works it was tackled the fault detection and identification in water networks by using machine learning technique such as various other types of neural networks (Caputo and Pelagagge, 2003; Shinozuka et al., 2005; Belsito et al., 1998; Feng and Zhang, 2006; Izquierdo et al., 2007) or Support Vector Machine (SVM) (Mashford et al., 2009). However, the use of the same pattern recognition architecture, such as the GFMMNN, with different simulation, state estimation and CLA algorithms is very scarce.

2. Numerical algorithms

Three main numerical algorithms are used in this paper for the generation of the training and testing data of the GFMMNN recognition system: a simulator algorithm, a state estimator and a CLA algorithm. The GFMMNN pattern recognition system was developed initially in the context of the simulation and state estimation of water network based on the nodal heads equations, the Newton–Raphson numerical technique (Jeppson & Davis, 1976) and the LS optimization criterion.

2.1. Simulator algorithm

Modeling and simulation of water distribution system consists of two main ingredients: the set of independent equations that describe the water network and the numerical optimization method used to calculate the nodal heads and the pipe flows. In Fig. 1 is shown a water network where the edges are the pipes that distribute the water to the consumers which are represented by the nodes (e.g. 1,2,3, etc.). A simulator algorithm is defined as a solution of the water network equations for a given set of nodal demands. The nodes represented with a square in the figure below are nodes with fixed head/pressure.

The simulator algorithm used here is based on the loop corrective flow algorithm defined for a water distribution system with n-nodes, l-loops, and p-pipes. The continuity equation must be satisfied, that is the flow entering a node equals the nodal consumption plus the flow exiting the respective node. Therefore, an initial pipe flows solution Q_i that satisfies the continuity equation is calculated as:

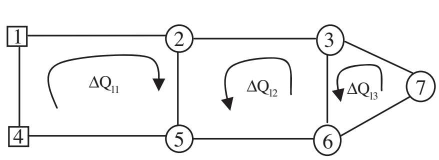

Fig. 1. Example of water network.

$$A_{np}Q_i = d \tag{1}$$

where d are the nodal demands and A_{np} is the topological incidence matrix.

The topological incidence matrix A_{np} has a row for every node and a column for every pipe of the water network. The entries for each row +1 and -1 indicate that the flow in a pipe enters or leaves the node (Arsene et al., 2004a; Arsene et al., 2004b). The energy equation is solved with the Newton–Raphson method and states that sum of the pipe head losses around each loop equals 0:

$$M_{lp}h = 0 \tag{2}$$

where h represents the pipes head losses calculated by the Hazen–Williams equation (Epp & Fowler, 1970) and M_{lp} is the loop incidence matrix.

The loop incidence matrix has the property that the entries +1 and -1 corresponds to the flow in a pipe being for example clockwise in a loop or anti-clockwise in the same loop, while 0 means that a pipe does not belong to a loop (Arsene et al., 2004a; Arsene et al., 2004b).

The loop corrective flows ΔQ_l at the step $k+1$ of the Newton–Raphson iteration method which solves (2) are:

$$\Delta Q_{l_{k+1}} = \Delta Q_{l_k} - \left[\frac{\partial \Delta H}{\partial \Delta Q_{l_k}}\right]^{-1} \Delta H \tag{3}$$

where ΔH are the residual loop head losses (i.e. $\Delta H = M_{lp}h$).

The Jacobian matrix $\frac{\partial \Delta H}{\partial \Delta Q_{l_k}}$ in (3) can be expressed as:

$$J = M_{lp}AM_{pl} \tag{4}$$

where M_{pl} is the transpose of loop incidence matrix and A is a diagonal matrix with a special property.

$$A = \begin{pmatrix} us_1|Q_1|^{u-1} & 0\ldots & 0 \\ 0\ldots & us_2|Q_2|^{u-1} & 0 \\ 0 & 0\ldots & us_p|Q_p|^{u-1} \end{pmatrix} \tag{5}$$

where $s_{1,2,\ldots p}$ is the pipe head loss coefficient and u is the exponent in the Hazen–Williams equation (Arsene, 2011; Carpentier & Cohen, 1993; Arsene et al., 2004a; Arsene et al., 2004b; Kumar, Narasimhan, & Bhadllamudi, 2008; Jeppson & Davis, 1976).

The final pipe flow solution $\widetilde{Q}$ for each pipe is:

$$\widetilde{Q} = Q_i + M_{lp}^T \Delta Q_l \tag{6}$$

where $\widetilde{Q}$ are the final pipe flows calculated at the end of the Newton–Raphson method (Arsene et al., 2004a; Arsene et al., 2004b; Jeppson & Davis, 1976).

The loop simulator requires the computation of the loop incidence matrix M_{lp} and the initial pipe flows Q_i. This problem is in general based on the decomposition of the water network into a spanning tree (e.g. Fig. 5) starting from a node which becomes the main root node with fixed value pressure. A spanning tree contains all the vertices and the edges of a connected and undirected graph except for the edges which form the cycles (i.e. loops) of the graph (Arsene et al., 2004a; Arsene et al., 2004b). Different search strategies can be employed in order to search the water network. The Depth First (DF) search from the graph theory is one of the possible choices for finding the loops in a water network. The DF search has the property that always a pipe that does not belong to the spanning tree called a chord pipe or a co-tree pipe, connects a node with one of its predecessor in the tree. Based on the spanning tree, the topological incidence matrix A_{np} can be split in a tree $T(n\times n)$ incidence matrix which defines the incidence of the tree pipes, which are the pipes situated in the spanning tree, and a co-tree incidence matrix $C(n\times l)$ which contain the co-tree pipes that are not in the spanning tree and form the loops (i.e. A_{np} = [T