Control Engineering Practice 55 (2016) 162–173

Contents lists available at ScienceDirect

Control Engineering Practice

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

Leak localization in water distribution networks using a mixed model-based/data-driven approach

Adrià Soldevila [a], Joaquim Blesa [a,b], Sebastian Tornil-Sin [a,b], Eric Duviella [c,d], Rosa M. Fernandez-Canti [a], Vicenç Puig [a,b,*]

[a] Research Center for Supervision, Safety and Automatic Control (CS2AC), Rambla Sant Nebridi, s/n, 08022 Terrassa, Spain
[b] Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Carrer Llorens Artigas, 4-6, 08028 Barcelona, Spain
[c] Mines DOUAI - IA - F-59508 Douai, France
[d] Univ. Lille, F-59000 Lille, France

ARTICLE INFO

Article history:
Received 26 January 2016
Received in revised form
26 April 2016
Accepted 4 July 2016
Available online 27 July 2016

Keywords:
Water distribution networks
Leak localization
Fault diagnosis
Model-based methods
Data-driven methods
Barcelona WDN

ABSTRACT

This paper proposes a new method for leak localization in water distribution networks (WDNs). In a first stage, residuals are obtained by comparing pressure measurements with the estimations provided by a WDN model. In a second stage, a classifier is applied to the residuals with the aim of determining the leak location. The classifier is trained with data generated by simulation of the WDN under different leak scenarios and uncertainty conditions. The proposed method is tested both by using synthetic and experimental data with real WDNs of different sizes. The comparison with the current existing approaches shows a performance improvement.

1. Introduction

Water leaks in a Water Distribution Network (WDN) can cause significant economic losses in fluid transportation leading to increase reparation costs that finally generate an extra cost for the final consumer. In many WDNs, losses due to leaks are estimated to account up to 30% of the total amount of extracted water (Puust, Kapelan, Savic, & Koppel, 2010). This is a very important amount in a world struggling to satisfy water demands of a growing population (Valipour, 2014a, 2014b, 2014c, 2014d).

The traditional approach to leakage control is a passive one, whereby the leak is repaired only when it becomes visible. Recently developed acoustic instruments (Khulief, Khalifa, Mansour, & Habib, 2012) allow to locate also invisible leaks, but unfortunately, their application over a large-scale water network is very expensive and time-consuming. A viable solution is to divide the network into District Metered Areas (DMAs), where the *flow* and the *pressure* at the input are measured (Lambert, Simpson, Vítkovský, Wang, & Lee, 2003; Puust et al., 2010), and to maintain a permanent leakage control-system: leakages in fact increase the flow and decrease the pressure measurements at the DMA entrance. Various empirical studies (Lambert, 2001; Thornton & Lambert, 2005) propose mathematical models to describe the leakage flow with respect to the pressure at the leakage location. Best practice in the analysis of DMA flows consists in estimating the leakage when the flow is minimum. This typically occurs at night, when customers' demand is low and the leakage component is at its largest percentage over the flow (Puust et al., 2010). Therefore, practitioners monitor the DMA or groups of DMAs for detecting (and then repairing) leakages by analyzing the minimum night flow, and also employ techniques to estimate the leakage level (Puust et al., 2010). However, leakage detection may not be easy, because of unpredictable variations in consumer demands and measurement noise, as well as long-term trends and seasonal effects.

Several works have been published dealing with leak location methods for WDN (see Wu & Sage (2006) and references therein). For example, in Colombo, Lee, and Karney (2009), a review of transient-based leak detection methods is offered as a summary of current and past work. In Yang, Wen, and Li (2008), a method is proposed to identify leaks using blind spots based on previously leak detection that uses the analysis of acoustic and vibrations signals (Fuchs & Riehle, 1991), and models of buried pipelines to predict wave velocities (Muggleton, Brennan, & Pinnington, 2002). More recently, Mashford, Silva, Marney, and Burn (2009) have

* Corresponding author at: Research Center for Supervision, Safety and Automatic Control (CS2AC), Rambla Sant Nebridi, s/n, 08022 Terrassa, Spain.
E-mail address: vicenc.puig@upc.edu (V. Puig).

http://dx.doi.org/10.1016/j.conengprac.2016.07.006
0967-0661/

developed a method to locate leaks using Support Vector Machines (SVM) that analyzes data obtained by a set of pressure control sensors of a pipeline network to locate and calculate the size of the leak. Another set of methods is based on the inverse transient analysis (Covas & Ramos, 2001; Kepler, Covas, & Reis, 2011). The main idea of this methodology is to analyze the pressure data collected during the occurrence of transitory events by means of the minimization of the difference between the observed and the calculated parameters. In Ferrante and Brunone (2003a, 2003b), it is shown that unsteady-state tests can be used for pipe diagnosis and leak detection. The transient-test based methodologies use the equations for transient flow in pressurized pipes in frequency domain and then, information about pressure waves is taken into account too.

Model-based leak detection and isolation techniques have also been studied starting with the seminal paper of Pudar and Liggett (1992) which formulates the leak detection and localization problem as a least-squares parameter estimation problem. Unfortunately, the parameter estimation of water network models is not an easy task (Savic, Kapelan, & Jonkergouw, 2009). The problem of leak localization in WDNs can be addressed as a particular case of the general problem known in the literature as the problem of Fault Detection and Isolation (FDI) in dynamic systems (Blanke, Kinnaert, Lunze, & Staroswiecki, 2006). However, the model of a DMA leads to a non-explicit model that can only be solved using numerical methods and limiting the applicability of most of the current FDI approaches that make an explicit use of the model. Moreover, there exist a high coupling of residuals and leaks plus a reduced number of sensors that as a result they complicate the isolation task. For this reason specific fault diagnosis methods for leak localization should be developed. A first contribution in this line can be found in Pérez et al. (2011, 2014) where a model-based method that relies on pressure measurements and leak sensitivity analysis is proposed. This methodology consists in computing online residuals, i.e. differences between the measurements and their estimations obtained using the hydraulic network model, and checking them against thresholds that take into account the modeling uncertainty and the noise. When some of the residuals violate their threshold, the residuals are matched against the leak sensitivity matrix in order to discover which of the possible leaks is present. Although this approach has good efficiency under ideal conditions, its performance decreases due to the nodal demand uncertainty and noise in the measurements. This methodology has been improved in Casillas, Garza-Castañón, and Puig (2012) where an analysis along a time horizon has been taken into account and a comparison of several leak isolation methods is presented. It must be noticed that in cases where flow measurements are available, leaks could be detected more easily since it is possible to establish simple mass balance in the pipes. See for example the work of Ragot and Maquin (2006) where a methodology to isolate leaks is proposed using fuzzy analysis of the residuals. This method finds the residuals between the flow measurements and their estimation using a model without leaks. However, although the use of flow measurements is feasible in large water transport networks, this is not the case in water distribution networks where there is a dense mesh of pipes with only flow measurements at the entrance of each District Metering Area (DMA). In this situation, water companies consider as a feasible approach the possibility to installing some pressure sensors inside the DMAs, because they are cheaper and easier to install and maintain.

In this paper, a new approach for leak localization in WDNs is presented. This methodology is used once the leak has been detected by means of the analysis of the nightly water demands of the DMA that is used for detecting and estimating the leakage level (Puust et al., 2010), and after the application of the sensor validation and reconstruction described in Cugueró-Escofet et al. (2016). The approach combines the use of pressure models and classifiers. Following a model-based methodology successfully tested in Pérez et al. (2011, 2014), a pressure model of the considered WDN is used in a first stage to compute residuals, i.e. differences between the measured (sensors) and estimated (model) values of the water pressure in nodes of the network, that are indicative of leaks. In a second stage, a classifier is applied to the obtained residuals with the aim to determine the leak location. This on-line scheme relies on a previous off-line work in which the network model is obtained and the classifier is trained with data generated in extensive simulations of the network. These simulations consider leaks with different magnitudes in all the nodes of the network, differences between the estimated and real consumer water demands and noise in pressure sensors. The underlying idea is to obtain a classifier able to distinguish the leak location independently of the unknown real leak magnitude and the presence of uncertainties associated with the water demands and the pressure measurements.

The rest of the paper is organized as follows. Section 2 presents the background on model-based fault leak localization methods based on sensitivity analysis and highlights their limitations. Section 3 presents in detail the proposed method. Section 4 details the application of the method to three WDNs of different sizes and provides a comparison with other well-established approaches. Finally, Section 5 draws the main conclusions of the work and introduces some potential extensions.

2. Background

2.1. Principle of model-based leak location approaches

Model-based approaches aim to locate leaks in a water distribution network by comparing pressure measurements with their estimations obtained using the hydraulic network model. Usually, this methodology is used for locating leaks in a given leak size range defined by the water network management company. The minimum size is related to the sensor resolution and modelling/demand uncertainty and the maximum size is defined as the value such that the leak behaves as burst and the leak can be seen in street. Model-based leak localization methods are based on comparing the monitored pressure disturbances caused by leaks at certain inner nodes of the DMA network with the theoretical pressure disturbances caused by all potential leaks obtained by using the DMA network mathematical model (Pérez et al., 2014). This comparison uses the residual vector, $\mathbf{r} \in \mathbb{R}^{n_s}$ that is determined by the difference between the measured pressure at inner nodes where sensors are installed

$$\mathbf{r}(t) = \mathbf{p}(t) - \hat{\mathbf{p}}_{\mathbf{o}}(t) \tag{1}$$

where $\mathbf{p} \in \mathbb{R}^{n_s}$, and the estimated pressure at these nodes obtained using the network model considering a leak-free scenario, $\hat{\mathbf{p}}_{\mathbf{o}} \in \mathbb{R}^{n_s}$.

The size of the residual vector $\mathbf{r}$, n_s, depends on the number of inner pressure sensors installed in the DMA. In recent years, some optimal pressure sensor placement algorithms have been developed to determine which pressure sensors have to be installed inside the DMA such that with minimum economical costs (number of sensors), a suitable performance regarding leak localization is guaranteed, see Pérez et al. (2011), Casillas, Puig, Garza-Castañón, and Rosich (2013), Sarrate, Blesa, Nejjari, and Quevedo (2014) among others.

The number of potential leaks, $\mathbf{f} \in \mathbb{R}^{n_n}$, is considered to be equal to the number of DMA nodes n_n, since from the modeling point of view, as proposed in Pérez et al. (2011, 2014), leaks are assumed to be in these locations.