Model-Based Monitoring Techniques for Leakage Localization in Distribution Water Networks

Jordi Meseguer\textsuperscript{a,*}, Josep M. Mirats-Tur\textsuperscript{a}, Gabriela Cembrano\textsuperscript{b,a}, Vicenç Puig \textsuperscript{b,c}

\textsuperscript{a} CETaqua, Water Technology Center, Carr. Esplugues 75, Cornellà de Llobregat, 08940, Spain
\textsuperscript{b} Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain
\textsuperscript{c} Dpt. Enginyeria de Sistemes, Automàtica i Informàtica Industrial (UPC), Barcelona, Spain

Abstract

This paper describes an integrated model-based monitoring framework for leakage localization in district-metered areas (DMA) of water distribution networks, which takes advantage of the availability of a hydraulic model of the network. The leakage localization methodology is based on the use of flow and pressure sensors at the DMA inlets and a limited number of pressure sensors deployed inside the DMA. The placement of these sensors has been computed using an optimal sensor placement method based on a Genetic Algorithm optimization, which integrates the direct modelling approach (simulation) used to identify the location of leaks. The application of the resulting monitoring framework in a certain DMA of the Barcelona distribution network is provided and discussed using simulated leakage scenarios. The obtained results show that leakage detection and localization may be performed efficiently, reducing the required time for detection/localization, by following a simple procedure.

Keywords: Water Networks; Hydraulic Models; Leakage Localization; Optimal Sensor Placement; Pressure Sensor; Genetic Algorithm

1. Introduction

Continuous improvements on water loss management applied to large class of water distribution networks are being applied, based on the use of new available technologies. Nonetheless, the whole leakage localization process may still require long periods of time (i.e. weeks, months) with an important volume of water wasted before the leak is found [1]. To avoid this inconvenience, leakage detection and localization based on mathematical models may be
used [2] which can “compare” the data gathered by installed sensors in the network with the data obtained by a model of this network. The use of flow and pressure sensors together with hydraulic models of the water network for leak detection and localization is a suitable approach for the on-line monitoring of water balance [3][4][5]. Thereby, Pérez et al. [1] presents a straightforward direct modelling methodology for leakage detection and localization in district metered areas (DMAs) of water distribution networks which is inspired by the binary model-based fault diagnosis theory [6] and takes benefit from those available DMA hydraulic models used by water operators. In [7], the method proposed in [1] was extended to work with non-binary fault signatures enhancing the overall performance of the method.

Despite the successful results, this type of monitoring techniques for leakage localization applied to large class of water distribution networks presents some limitations. Firstly, the number of sensors installed is usually limited because of budget constraints. And secondly, the sensor devices need to be properly located in order to enhance the performance of the real-time leakage detection module [8]. Therefore, a strategy that optimizes the number and placement of sensors is required. However, the performance of the leakage localization method is very sensitive to the number of available sensors and their placements [1][9][10]. Consequently, the optimal sensor placement and the leakage localization problems can not be considered separately. Since the objective of the optimization is to enhance the performance of the model-based leakage diagnosis, the sensor placement must consider requirements set for this process. In this line, the sensor deployment should provide:

1. High distinguishability among potential leaks. In general, these techniques are based on the fact that a leak at a certain location causes a characteristic effect across the entire network, which is actually measured by sensors. Then, if the effects of different leaks are similar (i.e. different leaks produce the same sensor measurements), these cannot be isolated.

2. Strong robustness in front of model-reality divergences and other uncertainties. In real applications, there is always a mismatch between models and reality, sensor measurements may be affected by disturbances, or there may be unknown system inputs.

When an accurate hydraulic model of the network (i.e. DMA) is available, the combined solution of the sensor-placement and leakage-detection problems is very straightforward: the optimal sensor placement problem can be formulated as an optimization-simulation method [8], while leakage localization can be achieved through direct-modelling techniques based on the use of the hydraulic model [7][11]. Regarding the optimal sensor placement problem, global search techniques based on Genetic Algorithms (GA) have been shown to be a suitable approach to carry out the optimization process selecting the potential solutions (i.e. sensor locations) following an evolutionary procedure [12] and assessing their fitness using a simulation environment aiming to compute an estimation of the overall leakage localization performance [8][9].

This paper describes the overall monitoring framework for leakage detection and localization, stressing the importance of coordinating sensor placement and the leakage localization. To this end, the sensor placement method presented by [9] is considered which is developed to enhance the performance of the model-based leakage localization method presented by [7]. As a case study, a real DMA of the Barcelona distribution network is taken into account.

This paper is organized as follows: Section 2 presents the model-based monitoring framework for leakage localization considered in this paper. Section 3 recalls the model-based leakage localization method proposed by [7]. Then, in Section 4, the optimal pressure sensor placement method proposed to maximize the performance of the leakage localization procedure is described. The influence of the pressure sensor placements on the performance of the leakage localization method is illustrated in Section 5 using the proposed case of study. Finally, in Section 6, the main conclusions are presented.

2. Model-Based Monitoring Techniques for Leakage Localization in Distribution Water Networks

This section describes the considered optimization-simulation framework to solve the sensor placement problem enhancing the overall performance of the leakage localization in terms of leakage detectability and isolability (Fig 1). The considered methodology relies on a GA-based optimization-simulation process which takes advantage of the global search capacity of GAS [12] to evolve a population of potential solutions (i.e. sensor locations) to the (near) optimal ones [13]. The fitness of every proposed potential solution (i.e. sensor locations) calculated in the ‘Fitness
Evaluation’ step is based on estimating the overall performance of the considered leakage localization method achievable with this sensor configuration using simulation. The ‘Leakage Localization Performance Simulator’ is based mainly on the considered leakage localization method and should consider those representative network situations that could appear in a real operation (i.e. representative operating points, mismatches between the network hydraulic model and the reality or between the demand model and the reality).

3. Model-Based Leakage Localization Method for Distribution Water Networks

3.1. Mathematical Modeling

The method proposed by [7] works with steady-state models in an extended period simulation (EPS) [13] where the governing laws are determined by the conservation of mass and energy [2] and a demand model is also considered. Thereby, the demand of node \( i \) was defined by the nodal base demand \( b_i \) and demand pattern \( p_{a,i} \) which may be estimated using the billing information provided by the water operator [7]. Then, leaks are assumed to be located in the nodes and simulated as an emitter coefficient \( C_j \) generating a leakage size depending on the pressure of that node ([14], [15]):

\[
f_j = C_j p_j^\gamma
\]

where \( f_j \) is the leak size; \( C_j \) is the associated emitter coefficient; \( p_j \) is the pressure at node \( j \); and \( \gamma \) is an exponent in the range of 0.5 (Hazen-Williams, Darcy-Weisbach, Chezy-Manning formulas).

In this method, the DMA EPS hydraulic model is implemented in EPANET [16] updating the boundary conditions of the network at a given time instant \( k \) using the measurements of the DMA total inflows and pressure at the DMA inlets at this time instant.

3.2. Model-Based Leakage Localization method

The method proposed by [7] is based on comparing, at every time instant \( k \), 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 which are obtained using the DMA hydraulic model. Thereby, the residual set, \( r(k) \in \mathbb{R}^{ns} \), is determined by the difference between the measured pressure at certain network nodes, \( p(k) \in \mathbb{R}^{ns} \), and the predicted pressure at these nodes considering a leak-free scenario, \( \hat{p}_{\text{m}}(k) \in \mathbb{R}^{ns} \):

\[
r(k) = p(k) - \hat{p}_{\text{m}}(k) = (p_{1}(k) - \hat{p}_{\text{m1}}(k)) \cdots (p_{ns}(k) - \hat{p}_{\text{mns}}(k))
\]

The size of the residual vector \( r \), \( ns \), depends on the number of inner pressure sensors of the DMA network.
Regarding the number of potential leaks, \( f = \{ f_1, f_2, \cdots, f_n \} \), it is equal to the number of network nodes, \( np \), since from the modeling point of view, leaks are placed in these locations (single-leak scenario assumption). On the other hand, the theoretical pressure disturbances caused by all potential leaks are stored in the theoretical fault signature matrix, \( FSM(k) \in \mathbb{R}^{ns \times np} \), with as many rows as DMA inner pressure sensors, \( ns \), and as many columns as potential leaks (DMA network nodes), \( np \). This matrix can be obtained from a sensitivity-to-leak analysis which evaluates the theoretical effect of all potential leaks \( f \) in the pressure of all the monitored nodes, \( p_i(k) \) [7]:

\[
FSM(k) = \begin{bmatrix}
\hat{p}_{f_1}(k) - \hat{p}_{0f_1}(k) & \cdots & \hat{p}_{f_n}(k) - \hat{p}_{0f_n}(k) \\
\vdots & \ddots & \vdots \\
\hat{p}_{nf_1}(k) - \hat{p}_{0nf_1}(k) & \cdots & \hat{p}_{nf_n}(k) - \hat{p}_{0nf_n}(k)
\end{bmatrix}
\]  

(3)

where \( \hat{p}_{if_i}(k) \) is the predicted pressure in the node where the pressure sensor \( i \) is placed when a nominal leak of size \( f \) is forced in node \( j \) and \( \hat{p}_{0if_i}(k) \) is the predicted pressure associated with the sensor \( i \) under a scenario free of leaks. Both the matrix \( FSM(k) \) and the vector \( r(k) \) depend on the demand and boundary conditions [17] and must be computed at every analysis time step, \( k \).

Regarding the leakage localization process at time instant \( k \), this is based on a correlation process which compares the residual vector \( r(k) \) (Eq. (2)) with the theoretical signatures of all potential leaks (columns of matrix \( FSM(k) \); Eq. (3)) applying the correlation function\(^2\). The potential leaks whose theoretical signatures have the largest correlation values with the residual vector \( r(k) \) are the best candidates to have the leak

\[
\max_j \left( \rho_{r, FSM_j}(k) \right) \quad j = 1, \ldots, np
\]

(4)

where \( \rho_{r, FSM_j} \) is the obtained correlation between the residual, \( r(k) \) and the \( j^{th} \)–column of the theoretical fault signature matrix, \( FSM_j \), associated with a potential leak in node \( j \).

4. Optimal Pressure Sensor Placement for Model-Based Leakage Localization in Water Distribution Networks

4.1. Introduction

The placement of the DMA inner pressure sensors is crucial in order to achieve a suitable performance of the leakage localization process in terms of leakage detectability and isolability [1][8]. For the leakage localization method described in [7], a binary optimal pressure-sensor placement method was considered [18] assuming that a potential sensor either detects a leak or not (hence binary) and for those detected leaks, the pressure sensitivity to a leak (Eq. (3)) of the potential sensors is the same. Nonetheless, this does not hold for real network where the pressure sensitivity to a leak of a given sensor depends on the place where is located (i.e. placements near DMA inlets have a low pressure sensitivity to the existing leaks given that the pressure is set by the pressure reducing valves). In [9], the previous binary methodology is extended by removing the binary reasoning process to increase the leak distinguishability, obtaining an enhanced optimizer cost function that aims at enhancing both distinguishability and robustness.

\(^2\)Pearson’s correlation coefficient \( \rho_{x,y} \) between two variables \( x, y \), may be defined as \( \rho_{x,y} = \frac{\text{cov}(x,y)}{\sqrt{\text{cov}(x,x)\text{cov}(y,y)}} \), where \( \text{cov}(a,b) = E[(a - \overline{a})(b - \overline{b})] \) is the covariance function between two variables \( a \) and \( b \), being \( \overline{a} = E(a) \) and \( \overline{b} = E(b) \) respectively.
4.2. Optimal Pressure Sensor Placement Method

In consonance with the considered leakage location methodology (Section 3.2), the theoretical fault signature matrix, \( FSM(k) \in \mathbb{R}^{\text{wop}} \) (Eq. (3)), with as many rows as DMA inner pressure sensors, \( ns \), and as many columns as potential leaks (DMA network nodes), \( np \), contains the theoretical signatures of all potential leaks and thus, being the key element to isolate (locate) the leak once the residual set, \( r(k) \in \mathbb{R}^{\text{w}} \) (Eq. (2)) is computed and using the procedure given by Eq. (4).

The objective of the optimal pressure sensor placement algorithm is to estimate the placement of \( ns \) pressure sensors, \( ND \in ND \), which allow to maximize the number of isolable leaks and to enhance the overall leakage localization process. The set \( ND \) is defined as follows:

\[
ND = \left\{ N_{d_v} \in (Z^+)^{\text{w}} | N_{d_v} : \alpha = 1, \ldots, ns; v = 1, \ldots, nN \right\}
\]

where \( ns \) is the number of sensors to be considered \((ns \leq np)\), \( N_{d_v} \) is the node index to place a given pressure sensor \((1 \leq N_{d_v} \leq np)\) and \( nN \) which determines the \( GA \) search-space, in general, is determined by \( \left( \begin{array}{c} np \\ ns \end{array} \right) \).

In general, when applying the leakage localization method described in Section 3.2, the theoretical fault signature matrix \( FSM(k) \in \mathbb{R}^{\text{wop}} \) is computed at every time instant considering the existing boundary conditions associated with the operating point of the DMA network (i.e. pressure and flow at the DMA inlets) [11]. However, the optimal pressure sensor placement methods should compute sensor placements enhancing the performance of the leakage localization method for the different existing operating points of the DMA network. To tackle this point, the matrix \( FSM \in \mathbb{R}^{\text{wop} \times \text{nc}} \) can be used which can be constructed by appending \( nc \) theoretical fault signature matrices \( FSM(k) \in \mathbb{R}^{\text{wop}} \) which are associated with representative DMA network scenarios (operating points / situations) (i.e. day-night, seasonality, model-reality discrepancies) of the DMA network:

\[
FSM = (FSM_{1,ij})_a \quad \alpha = 1, \ldots, nc
\]

For a given scenario \( \alpha \), the fitness, \( (FSS_v)_\alpha \), of the potential solution \( N_{d_v} \) may be computed as a function of different indicators [19]: e.g. the number of unique leak signatures \( FSM_{v,\alpha} \) (isolable leaks) contained in \( (FSM)_\alpha \) \((NG_v \in \mathbb{R}: \text{to be maximize})\); the sum of the size of the largest groups of non-isolable leaks \(4 \) \((GS_v \in \mathbb{R}: \text{to be minimize})\); the sum of the biggest geographical areas of the groups of non-isolable leaks \(5 \) \((GSgeo_v \in \mathbb{R}: \text{to be minimize})\). Thus, the fitness, \( (FSS_v)_\alpha \), of the potential solution \( N_{d_v} \) can be computed as a function \((f: \mathbb{R}^3 \rightarrow \mathbb{R})\) of these indicators:

\[
(FSS_v)_\alpha = f(NG_v, GS_v, GSgeo_v)
\]

Then, taking into account the \( nc \) representative scenarios, the overall fitness value \( N_{d_v} \) can be computed as a weighted sum of the fitness values \((FSS_v)_\alpha\) associated with every \( \alpha \) scenario:

\[
FSS_v = \sum_{\alpha=1}^{nc} \beta_\alpha (FSS_v)_\alpha
\]

where \( \beta_\alpha \in \mathbb{R} \) \((0 \leq \beta_\alpha \leq 1; \sum_{\alpha=1}^{nc} \beta_\alpha = 1)\) are the weights used to stress the importance of certain scenarios.

Based on the method proposed by [7] (Section 3.2), [11] shows the procedure to be followed to compute the theoretical fault signature matrix \( (FSM)_\alpha \). Considering this procedure and the sensor placement optimization framework illustrated in Fig 1, Fig 2 shows the details of the considered optimization-simulation framework to solve the optimal pressure sensor placement problem.

---

3 A leak in node \( j1 \) with a leak signature \( FSM_{j1} \), \( j1 \)–column of the theoretical fault signature matrix \( FSM \), is isolable regarding another leak in node \( j2 \) with a signature \( FSM_{j2} \) when the component-wise condition \( FSM_{j1} \neq FSM_{j2} \) holds. This condition could be computed through the Pearson’s correlation coefficient used in Eq. (4).

4 For every unique leak signature \( FSM_j \) contained in \( FSM \), there is a group of non-isolable leaks sharing the same signature.

5 Geographical area containing all the nodes associated with a group of non-isolable leaks.
5. Case Study

5.1. Description

In this paper, the Castelldefels Platja DMA of the Barcelona water distribution network\(^6\) has been used to illustrate the influence of pressure-sensor placement on the performance of the leakage localization method. This DMA, with a network length around 80 km, has two inlets, 2492 nodes and 5111 pipes. Its area is approx. 6 km\(^2\) with a delivered monthly volume of 100,000 m\(^3\) in winter season and 200,000 m\(^3\) in summer season, on average. In Fig 3, the water network of Castelldefels Platja DMA can be seen. In this figure, the two DMA inlets have been highlighted using red triangle symbols. Regarding the instrumentation, this DMA is provided with flow and pressure sensors at each inlet and by 6 inner pressure sensors. In Fig 3, the placements of the 6 inner pressure sensors (green star symbols) computed using the framework explained in Section 4 can be seen which have inspired the real placement implemented in this DMA.

\(^6\) This network managed by Aigües de Barcelona is segmented into 117 pressure levels, and 214 DMAs.
5.2. Performance of the leakage localization method: optimal pressure sensor distribution

In this section, the performance of the leakage localization method (Section 3) is illustrated using the optimal pressure sensor distribution shown Fig 3 and applying two simulated leak scenarios. In every scenario, the simulated leak has been forced using an emitter coefficient $C=0.92$ being $6.5 \text{l/s}$ the average size of the resulting leaks given that the average pressure in this DMA is around $30 \text{ m.w.c.}$ (Eq. (1)). In real applications, there is a mismatch between the hydraulic model and the real behavior of the DMA networks due to modelling errors and the uncertainty about the water demand distribution. Thus, in the simulated leak scenarios, a mismatch between the water demand model integrated into the DMA network hydraulic model (Section 3.1) and the real water demand distribution has been forced. Thereby, the nodal base demands, $b_i$, have been disturbed randomly between $10\%$ and $-10\%$ regarding the ones used by the hydraulic model. In these scenarios, measurements of the DMA inner pressure sensors have been estimated using the EPS hydraulic model implemented in EPANET and using real network loading conditions corresponding to a certain day (see Fig 4 where measurements of DMA inflows and pressure at DMA inlets are presented).

Applying the procedure given in Section 4, considering the placements ($N_d$) of the inner pressure sensors highlighted in Fig 3, the number of unique leak signatures is $N_g=275$ and the sum of the size of the 5 largest groups of non-isolable leaks is $G_S=1357$ nodes. In the following table, the size of the 5 largest groups can be seen:

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (nodes)</td>
<td>416</td>
<td>410</td>
<td>217</td>
<td>167</td>
</tr>
</tbody>
</table>

Fig 5 shows the performance of the leakage detection method using two simulated single-leak scenarios (one in each plot) highlighting the leak exact location (red spot), the predicted most correlated location (blue circle) and other predicted high-correlated locations (black spots)($>98\%$ of the highest correlation). In these two scenarios, this method predicts a location close to the exact location of the simulated leaks due to the considered water demand distribution. Nonetheless, the predicted localization is still inside the area of high-correlated nodes which contains the real localization of the leak.
Fig 4 Real network loading conditions used to compute all the simulated leak scenarios: measurements of DMA inflows (left plot) and pressure at DMA inlets (right plot).

5.3. Performance of the leakage localization method: non-optimal pressure sensor distribution

In this section, the procedure followed in Section 5.3 is applied but considering a non-optimal pressure sensor distribution which can be seen in Fig 6 (green stars). For these placements \( N_{d,v} \), the number of unique leak signatures is \( NG_v = 48 \) and the sum of the size of the 5 largest groups of non-isolable leaks is \( GS_v = 2976 \) nodes (see Table 2). Regarding the values presented in Section 5.2, the number of unique leak signatures \( NG_v \) has meaningfully decreased while the sum of the size of the 5 largest groups of non-isolable leaks \( GS_v \) has meaningfully increased. Therefore, this means that when using the pressure sensor placements shown in Fig 6 (green stars), the performance of the leakage localization method is worsened being capable to isolate a lower number of leaks and for every leak which can be isolated, the uncertainty about the true location is bigger given that there is a big number of leaks presenting a similar signature. This effect can be seen in Fig 6 where the performance of the leakage detection method is shown using the two simulated single-leak scenarios (one in each plot) considered in Section 5.2. In this case, the error associated with the prediction of the leak exact location and the area of high-correlated nodes increase meaningfully regarding the case shown in Section 5.2 obtaining a poor performance in terms of leak localization.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (nodes)</td>
<td>1024</td>
<td>682</td>
<td>623</td>
<td>390</td>
</tr>
</tbody>
</table>

6. Conclusions

This paper deals with the importance of coordinating sensor placement and leak detection/localization methodologies to achieve the best performance. It reviews a method for leak detection in DMAs of urban distribution networks based on comparing simulated pressure signals with real pressures measured by sensors in the network, as those pressure difference vectors contain the signatures of the possible faults happening in the network. The performance of this process depends on where the pressure sensors are installed in the network and on the characteristics of these sensors. Then, the process of leakage detection cannot be separated to that of sensor placing and, conversely, the sensor placement methodology cannot be separated from that of leakage localization. The main
contribution in this paper is an interdependent algorithm of both processes. A case study of a DMA in the Barcelona water distribution network is used to show the dependency of the overall leakage detection/localization performance and sensor placement. It is important to remark that one of the underlying hypothesis is that a single fault occurs in the network at a time. Research to consider the case of multiple concurrent faults is currently underway.

Fig 5 Optimal pressure sensor distribution: performance of the leakage detection method using two single-leak scenarios (one in each plot) highlighting the leak exact location (red spots), the most correlated location predicted by the method (blue circle) and other locations presenting also high correlations (black spots) (>98% of the highest correlation).

Fig 6 Non-optimal pressure sensor distribution: performance of the leakage detection method using two single-leak scenarios (one in each plot) highlighting the leak exact location (red spots), the most correlated location predicted by the method (blue circle) and other locations presenting also high correlations (black spots) (>98% of the highest correlation).
Acknowledgements

The authors wish to thank the support received by the EFFINET grant FP7-ICT-2012-318556 of the European Commission and the Spanish Research Grant of project ECOCIS (MINECO) DPI2013-48243-C2-1-R.

References