Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 48-21 (2015) 214–219

Sensor Placement for Leak Location in Water Distribution Networks using the Leak Signature Space

Myrna V. Casillas* Luis E. Garza-Castañón* Vicenç Puig**

* *Supervision and Advanced Control Chair, Tecnológico de Monterrey, Campus Monterrey, Av. Eugenio Garza Sada 2501, 64849 Monterrey, México., (e-mail: mv.casillas.phd.mty, legarza @itesm.mx).*
** *Advanced Control Systems Research Group, Technical University of Catalonia, Rambla Sant Nebridi 10, 08222 Terrassa, Spain., (e-mail: vicenc.puig@upc.edu)*

Abstract: In this paper, a sensor placement approach to improve the leak location in water distribution networks is proposed. The sensor placement problem is formulated as an integer optimization problem where the criterion to minimize is the number of overlapping signature domains computed from the leak signature space (LSS) representation. A stochastic optimization process is proposed to solve this problem, based on either a Genetic Algorithms (GA) or a Particle Swarm Optimization (PSO) approach. Experiments on two different DMAs are used to evaluate the performance of the resolution methods as well as the efficiency achieved in the leak location when using the sensor placement results.

Keywords: Water distribution networks, Leak isolation, Sensor placement, Leak signature space.

1. INTRODUCTION

Currently, in a world struggling to satisfy the water demands of a growing population, leaks are estimated to account up to 30% of the total amount of extracted water. To face these challenges, encouraging new technologies arose during the last decades (Colombo et al., 2009), achieving higher levels of efficiency and that came together with novel methods for leakage management.

Despite these first results, the performance obtained until now is still far from allowing the detection of leaks in Water Distribution Networks (WDN) with only a few sensors in a robust and fast way. A major limitation is that the performance achieved is highly dependent on the location of the sensors installed in the network. The development of a sensor placement strategy has been an extensive subject of research. In the context of water systems, some work has been done regarding sensor placement for leak location. Sarrate et al. (2012), defines an isolability index and use it to place the sensors in order to maximize the number of isolable node pairs. Closer to this work, Pérez et al. (2011) used thresholds on the differences of pressures measured to obtain binary matrices that were used to translate the sensor placement problem to an integer programming optimization problem. In Casillas et al. (2013), a new approach for sensor placement for leak location in WDN is proposed that is based on the projection-based location scheme proposed in Casillas et al. (2012, 2014a). Actually, leak location and sensor placement should be considered together since the best placement depends on the method that is used to locate the potential leaks and the efficiency of the leak location depends on the sensor placement.

This paper introduces a model-based optimization method for near-optimal sensor placement to detect leak nodes in WDN. It relies on the so-called Leak Signature Space (LSS), an original representation where a specific signature is associated to each leak location that minimizes the dependence with its magnitude (Casillas et al., 2014b). The proposed approach allows to place adequately the sensors in a WDN in order to take the best benefit of the LSS based leak detection method.

The paper is organized as follows. In Section 2, the LSS method that is used for the leak location is introduced. Section 3 formulates the sensor placement optimization problem while Section 4 shows the methods proposed to solve it. Section 5 evaluates the performance of the approach on a real WDN through several scenarios. Finally, Section 6 summarizes the contribution and indicates points that deserve further attention.

2. LSS BASED LEAK LOCATION METHOD

This section presents the LSS method proposed in Casillas et al. (2014b) that is used to locate the leaks. It is based on the linear approximation of the dependency between the leak magnitude and the pressure residuals. Such model, is used to perform a transformation that allows to represent node leaks by means of points in the LSS, independently of the leak magnitude. Thus, by means of the LSS method the potential leak node can be characterized by a given signature.

2405-8963
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2015.09.530

2.1 Leak magnitudes linear dependency approximation

Here, let us assume that the behavior of the WDN follows the models described by Todini and Pilati (1988), and that it consists of m nodes, f pipe flows and n pressure sensors located at the nodes (typically $n \ll m$). Let us also define the vectors $\mathbf{p}$, $\mathbf{p}^*$, $\mathbf{q}$, $\mathbf{d}$ which are respectively the vectors of pressure in the junction nodes, pressure in reservoirs, flows trough the pipes and demands:

$$\begin{aligned} \mathbf{p} &= (p_1, \cdots, p_m)^{\mathbf{T}}, \\ \mathbf{p}^* &= (p_1^*, \cdots, p_u^*)^{\mathbf{T}}, \\ \mathbf{q} &= (q_1, \cdots, q_f)^{\mathbf{T}}, \\ \mathbf{d} &= (d_1, \cdots, d_m)^{\mathbf{T}}, \end{aligned} \tag{1}$$

with u corresponding to the number of reservoirs supplying the WDN. According to the representation proposed in Todini and Pilati (1988), the water network model can be solved numerically, using a Newton-Raphson iterative scheme, where the iteration $k+1$ is given by the following set of equations:

$$\begin{aligned} \mathbf{q}^{k+1} &= (I - N^{-1})\mathbf{q}^k - N^{-1}A_{11}^{-1}(\mathbf{q}^k)(A_{12}\mathbf{p}^k + A_{10}\mathbf{p}^*), \\ \mathbf{p}^{k+1} &= -(A_{21}N^{-1}A_{11}^{-1}(\mathbf{q}^k)A_{12})^{-1} \cdot \\ &\quad \left(A_{21}N^{-1}(\mathbf{q}^k + A_{11}^{-1}(\mathbf{q}^k)A_{10}\mathbf{p}^*) + (\mathbf{d} - A_{21}\mathbf{q}^k)\right), \end{aligned} \tag{2}$$

where N is a diagonal matrix such that $N = \text{diag}(\gamma_i)$, $i \in [1, \cdots, f]$, $A_{12} = A_{21}^T$, $A_{10} = A_{01}^T$ with A_{21}, A_{01} being incidence matrices, $A_{11}(\mathbf{q}) = \text{diag}(c_i|q_i|^{\gamma_i})$, $|q_i|$ is the absolute value of the flow q_i, c_i is a constant parameter which depends on the diameter, the roughness and the length of the pipe, and γ_f is the flow exponent parameter. It is important to note that this resolution approach is commonly employed, as e.g. in the EPANET simulator (Rossman, 2000) where large WDN can be simulated efficiently.

The solution of the system of equations (2) corresponds to the case where an equilibrium point has been reached for the network, i.e. the flows and pressures are constant along the time which means $\mathbf{p}^{k+1} = \mathbf{p}^{\mathbf{k}} = \mathbf{p}$ and $\mathbf{q}^{k+1} = \mathbf{q}^{\mathbf{k}} = \mathbf{q}$. A thorough representation of a WDN would theoretically involve a graph structure where each possible leak is assumed to be located in a graph node. However, a leak could possibly appear at any point of any network pipe. For this reason, the exact modeling of any possible leak becomes unfeasible in practice. To mitigate this issue, it is usually assumed that leaks only appear in the network nodes (see, e.g. Pudar and Liggett (1992) among others). With such assumption, the leak can be written as a vector of extra demands $\mathbf{\Delta d}$ and the new demand $\mathbf{d}'$ can be expressed such as:

$$\mathbf{d}' = \mathbf{d} + \mathbf{\Delta d}, \tag{3}$$

where $\mathbf{\Delta d}$ is a m dimensional vector with zeros everywhere except at the node's index where the leak occurs. Now, assuming that the network flow equilibrium has also been reached in presence of leak, the new pressure can be expressed as:

$$\begin{aligned} \mathbf{p}' &= -(A_{21}N^{-1}A_{11}^{-1}(\mathbf{q}')A_{12})^{-1} \cdot \\ &\quad \left(A_{21}N^{-1}(\mathbf{q}' + A_{11}^{-1}(\mathbf{q}')A_{10}\mathbf{p}^*) + (\mathbf{d} + \mathbf{\Delta d} - A_{21}\mathbf{q}')\right) \end{aligned} \tag{4}$$

Then, we propose to represent the residual $\mathbf{r}$ (c.f. Pérez et al. (2011)) as the difference between the nominal pressure obtained using the model without leaks and the pressure determined using the model in case of a leak. Assuming that the flow is approximately the same with and without leak ($\mathbf{q} \cong \mathbf{q}'$), we have:

$$\mathbf{r} = \mathbf{p} - \mathbf{p}', = (A_{21}N^{-1}A_{11}^{-1}(\mathbf{q})A_{12})^{-1}\mathbf{\Delta d}, = \mathbf{S} \cdot \mathbf{\Delta d}. \tag{5}$$

As one can see, it is possible to use a linear approximation of the relation between the residual (and consequently with the pressure measurement) and the leak through a $\mathbf{S}$ matrix factor, under equilibrium assumptions, that is known as the *sensitivity matrix*. However, in presence of a leak, a change will occur in the flow due to the extra demand occurring in one node of the network. Fortunately, the flow changes will be small for the problem addressed by our leak location method. In our case, we focus on medium size leaks, i.e., leaks ranging from 2 to 6 lps, which corresponds to medium ranges in leaks that occurs in the studied networks. Such range is chosen since smaller leaks are masked by uncertainties in the network and larger leaks usually reach the surface rapidly. This approximation involves to include an error factor $\varepsilon_{\mathbf{q}}$ that deviates from this linear approximation relation as $\mathbf{r} = \mathbf{S} \cdot \mathbf{\Delta d} + \varepsilon_{\mathbf{q}}$. This error is not modeled in the theoretical approximation but it will be taken into account in the realistic cases analysed in the experiments. The $\mathbf{S}$ matrix has been used in a variety of works (Pérez et al., 2011; Casillas et al., 2014a, 2012).

The pressure in a network at a given time instant can be represented by a point in the m-dimensional space of the pressure measurements. In case of a leak, equation (5) indicates that the position of this point is located on a line passing through the origin and whose direction depends on the node where the leak occurs. Moreover, the position of the point on the line depends on the leak magnitude. However, in practice, pressure measurements are accessible in only a limited number of nodes n, that correspond to locations where the sensors are placed. Fortunately, the n-dimensional space of the sensors is a subspace (a projection) of the m-dimensional space of the pressure measurements. Thus, this linear dependency is also valid in this projected space.

2.2 Leak signature space

The linear dependency presented above is such that for any pair of residual vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ corresponding to different leak magnitudes but occurring in the same node j, it can be stated that:

$$\mathbf{r}_2^j = \alpha \mathbf{r}_1^j, \tag{6}$$

with α proportional to the leak magnitude. Thus, any residual corresponds to a direction vector of the line representing the leak at a specific node. Based on the sensor representation, it is possible to use the projection of the direction vector onto a selected hyperplane of the n-dimensional space. For simplicity, let us assume for now that the last coordinate is chosen to form the hyperplane, such that for a given residual $\mathbf{r}^j = [r_1^j, \cdots, r_n^j]^{\mathbf{T}}$ the projection vector $\bar{\mathbf{r}}^j$ is computed as:

$$\bar{\mathbf{r}}^j = [\frac{r_1^j}{r_n^j}, \cdots, \frac{r_{n-1}^j}{r_n^j}, 1]^{\mathbf{T}}. \tag{7}$$

Thus, there is a unique expression of such projection vector for a linear representation of the residuals. Consequently, it is possible to associate to the leak in node j and independently of its magnitude, a unique point $\tilde{\mathbf{r}}^j = [\bar{r}_1^j, \cdots, \bar{r}_{n-1}^j]$ in this $(n-1)$-dimensional hyperplane called the Leak Signature Space (LSS).

To represent the leak associated to a given node j in the network with n sensors, f different leaks magnitudes ${}^k l$ are simulated, with $k \in [1, \cdots, f]$. Then, the associated f residual vectors ${}^k\mathbf{r}^j = [{}^k r_1^j, \cdots, {}^k r_n^j]^{\mathbf{T}}$ are computed and if the projection is performed with respect to the last coordinate, their projection onto the LSS are given by the $(n-1)$-dimensional points of coordinates ${}^k\tilde{\mathbf{r}}^j = [\frac{{}^k r_1^j}{{}^k r_n^j}, \cdots, \frac{{}^k r_{n-1}^j}{{}^k r_n^j}]^{\mathbf{T}}$. Then, the point $\tilde{\mathbf{r}}^j$, corresponding to the leak j, is taken as the barycenter of these f partial signatures ${}^k\tilde{\mathbf{r}}^j$ built from the different leak magnitudes:

$$\tilde{\mathbf{r}}^j = \left[\frac{1}{f} \sum_{k=1}^{f} \frac{{}^k r_1^j}{{}^k r_n^j}, \cdots, \frac{1}{f} \sum_{k=1}^{f} \frac{{}^k r_{n-1}^j}{{}^k r_n^j} \right]. \tag{8}$$

Computing such barycenters for the m possible leak nodes, m leak signatures $\tilde{\mathbf{r}}^j$ are obtained, with $j \in [1, \cdots, m]$, that can be used to perform leak location. It should also be remarked that when the number of sensors increases, the dimension of the hyperplane in the LSS also increases and thus increases the chances to discriminate the different leak signatures.

For each leak signature computed as the barycenter of the partial signatures, a domain of influence can be determined as a $(n-1)$-dimensional sphere whose radius corresponds to the largest Euclidean distance from the barycenter to the partial signatures. Thus, the radius for each leak signature domain is defined as:

$$\mathbf{rad}^j = \max_{k \in [1, \cdots, f]} \rho(\tilde{\mathbf{r}}^j, {}^k\tilde{\mathbf{r}}^j), \tag{9}$$

where $\rho(\tilde{\mathbf{r}}^j, {}^k\tilde{\mathbf{r}}^j)$ are the Euclidean distances between the barycenter of the node j and each of the partial signatures obtained by simulating the f different magnitudes. Therefore, if two domains overlap, it means that the two leak signatures are very similar and thus, there is a risk of confusing one leak with the other.

2.3 Leak location method

To estimate the leak location based on the LSS representation, first, the m leak signatures of the network nodes are computed with the method presented above. Then, when a leak occurs, its signature in the LSS is determined, i.e. the $(n-1)$-dimensional point $\tilde{\mathbf{r}}^*$ representative of this leak from the real residual measurements is computed. Finally, the Euclidean distance ρ between $\tilde{\mathbf{r}}^*$ and the various leak signatures $\tilde{\mathbf{r}}^j$ in the LSS is computed. The leak node is then estimated as the one whose signature is the closest to the current leak signature i.e the index id of the leak is such that:

$$id = \arg \min_{j \in [1, \cdots, m]} \rho(\tilde{\mathbf{r}}^*, \tilde{\mathbf{r}}^j), \tag{10}$$

where arg min is the argument of the minimal distance evaluated.

2.4 Incorporating a time horizon analysis

In practice, for a given network, the demand usually varies along the time and it is important to carry out the leak location analysis taking into account a given time horizon (Casillas et al., 2014a). To address this point, it is proposed to record, for each potential leak node j, the various signatures it has along a time horizon that corresponds to one cycle of the demand pattern. Since the demand varies along this pattern, the position of the leak signatures changes accordingly. Then, in presence of a leak, a comparison is made between the positions of its signatures in the LSS along the day and the positions of the reference leak signatures. Such comparison is performed by summing the Euclidean distances for each time instant considered. Thus, if T instants of time are considered, there are T residuals ${}_t\tilde{\mathbf{r}}^*$, $t \in [1, \cdots, T]$ obtained from the measured pressures which are compared with T leak signature references ${}_t\tilde{\mathbf{r}}^j$. The index of the leaky node is then estimated as the one that minimizes the sum of the distances between the node signature and the current leak signature along the time horizon:

$$id = \arg \min_{j \in [1, \cdots, m]} \left(\sum_{t \in [1, \cdots, T]} \rho({}_t\tilde{\mathbf{r}}^*, {}_t\tilde{\mathbf{r}}^j) \right). \tag{11}$$

3. SENSOR PLACEMENT: PROBLEM FORMULATION

The objective is to find for a WDN and a given number of sensors n, the sensor locations which maximize the efficiency of the leak location when using the LSS based leak location method. Here, we propose an explicit formulation of such optimization problem.

3.1 Considering a single instant of time

A given sensor placement corresponds to a unique configuration defined as a m-dimensional binary vector $\mathbf{x}$ such that:

$$\mathbf{x} = [\, x_1, \, \cdots, \, x_m \,], \tag{12}$$

where $x_i = 1$ if a sensor is installed (i.e. the pressure is measured) in the node i and $x_i = 0$ otherwise. Then, assuming a projection onto the hyperplane of index $\lambda \in [1 \cdots n]$, as discussed in Section 2, we can compute for each node j its leak signatures $\tilde{\mathbf{r}}^j(\mathbf{x}, \lambda)$ from equation (8) and its leak domains $\mathbf{rad}^j(\mathbf{x}, \lambda)$ from equation (9). For a given node, the domain of its leak signature represents the region where may appear the signature of the measured leak if the leak actually occurred at this node. Thus, when two leak domains intersect, it increases the risk for the two leaks to be non isolable. Based on such analysis, we propose to take as objective function the minimization of the overlapping between leak domains considering the signatures of all network nodes. For each pair of leak nodes (j, j') whose domains may overlap, a given sensor placement represented by the vector $\mathbf{x}$ and a given projection λ, we define the associated error $\varepsilon(\mathbf{x}, \lambda)^{jj'}$ such as:

$$\varepsilon(\mathbf{x}, \lambda)^{jj'} = \begin{cases} 1 \text{ if } \rho\left(\tilde{\mathbf{r}}^j(\mathbf{x}, \lambda), \tilde{\mathbf{r}}^{j'}(\mathbf{x}, \lambda)\right) \leq \mathbf{rad}^j(\mathbf{x}, \lambda) \\ \qquad\qquad\qquad\qquad\qquad\qquad + \mathbf{rad}^{j'}(\mathbf{x}, \lambda) \\ 0 \text{ otherwise} \end{cases} \tag{13}$$

with $j = 1, \cdots, m-1$ and $j' = j+1, \cdots, m$, where ρ is the Euclidean distance in the LSS space. Then, the error index that takes into account all the nodes leaks is computed as:

$$\epsilon(\mathbf{x},\lambda) = \sum_{j=1}^{m-1} \sum_{j'=j+1}^{m} \varepsilon(\mathbf{x},\lambda)^{jj'} \tag{14}$$

and the sensor placement problem is formulated as an optimization problem subject to equality constraints of the form:

$$\begin{aligned} &\min_{\mathbf{x},\lambda} \epsilon(\mathbf{x},\lambda) \\ &\text{s.t.} \\ &\sum_{i=1}^{m} x_i = n. \end{aligned} \tag{15}$$

3.2 Considering a time horizon analysis

As shown before, the leak location can incorporate a time horizon analysis (see Section 2.4). Following such a scheme, it is possible to use a time horizon in the evaluation function, with the objective of increasing the quality of the sensor placement and thus, for a better leak isolation within the network. To achieve that, the error function used in the optimization problem defined in equation (15) is modified in order to work with the mean number of overlaps along the time horizon analyzed and is then computed as:

$$\bar{\epsilon}(\mathbf{x},\lambda) = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t(\mathbf{x},\lambda) \tag{16}$$

where $\epsilon_t(\mathbf{x},\lambda)$ is the error at the time instant t computed from equation (14) and T is the length of the time horizon selected.

4. SENSOR PLACEMENT: RESOLUTION METHODS

In this section, we describe two stochastic methods that can be used to obtain solutions to the sensor placement problem: the Genetic Algorithm (GA) and the Particle Swarm Optimization (PSO) method. Both approaches require the establishment of a cost function that has to be minimized. We propose to work only on the configuration $\mathbf{x}$ defining the sensor placement as variable to optimize, and to deal with the variable λ by considering for a given $\mathbf{x}$ only the best LSS hyperplane projection that is the one which leads to the lowest number of overlaps. Therefore, the cost function considered is the modified error index of equation (14), that becomes:

$$\epsilon(\mathbf{x}) = \min_{\lambda\in[1\cdots n]} \sum_{j=1}^{m-1} \sum_{j'=j+1}^{m} \varepsilon(\mathbf{x},\lambda)^{jj'}, \tag{17}$$

and the optimization problem is now in the form:

$$\begin{aligned} &\min_{\mathbf{x}} \epsilon(\mathbf{x}) \\ &\text{s.t.} \\ &\sum_{i=1}^{m} x_i = n. \end{aligned} \tag{18}$$

The application of PSO to our problem relies on the PSO Toolbox of MATLAB developed by Chen (2009) and offered in the MathWorks page, where a particle can be represented by the possible presence or absence of a sensor at a given node.

The stochastic optimization method using either GA or PSO works as follows. First, the variables of the algorithm are initialized including the number of generations, the bit string type population, and the convergence tolerance (set as 10^{-10}). Restrictions to the search are also added, which ensures that the number of sensors placed is correct. Note that in the PSO case, the constraint regarding the number of sensors is in practice converted into a high penalty constraint since equality constraints were not accepted for binary variables. A seed of size z is chosen randomly which allows to create an initial matrix with random sensor positions. Then, the stochastic optimization method evaluates from the initial population matrix and the matrix of residuals, a set of sensor placements and returns the one which corresponds to the lowest error. After *it* iterations, the method returns the sensor configuration that minimizes the error the most, with the corresponding LSS hyperplane of projection λ. Such configuration minimizing the number of leak signature domains overlaps will allow to take the maximum benefit of the LSS based leak location method.

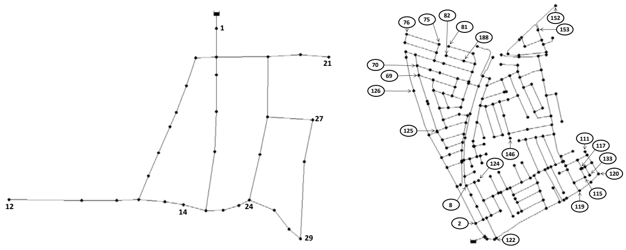
5. EXPERIMENTAL RESULTS

The resolution methods presented in the previous section were applied to two networks. First, the WDN of Hanoi in Vietnam (Fujiwara and Khang, 1990). This academic benchmark has been used in several works, where the objective was either to design the network or to optimize the operations performed within it (Zheng and Simpson, 2013). It consists of 1 reservoir, 31 demand nodes and 34 pipes as shown in Figure 1.a. Second, the WDN of Limassol in Cyprus. This real network has already been used in leak location and sensor placement works (Casillas et al., 2013) and consists of 1 reservoir, 197 demand nodes and 239 pipes as shown in Figure 1.b. All the experiments were performed in Matlab, using a Windows 7 Computer with a Pentium Dual Core processor of 2 GHz, a memory (RAM) of 4 GB and a 64-bit operating system. The network models are simulated using the EPANET software (Rossman, 2000). Leak magnitudes are simulated by changing a node emitter coefficient (EC), that is related to the flow rate through the following equation:

$$EC = \frac{q}{p^{\gamma}} \tag{19}$$

where q is the flow rate, p is the fluid pressure and γ is the pressure exponent. In the Hanoi network, to build the leak signatures, 7 EC values are considered going from 2 to 8 with a step size of 1 which leads to leaks magnitudes varying from 20 to 80 liters per second (lps). For the Limassol network, 7 EC are also used going from 0.3 to 0.9 with a step size of 0.1 which leads to leaks varying from 2 lps to 6 lps. Note that the orders of leak magnitudes are chosen differently from one network to the other because the order of magnitudes of the demands and flows appearing in the networks also differ.

In the following tables, the optimal sensor placements presented are the ones for which the number of overlaps between leak signature domains is minimum. Additionally, the efficiency is provided which corresponds to the percentage of leaks correctly located. It is computed by testing leaks in all possible nodes with all possible magnitude within the range used to build the signature models and domains. The measurements are simulated with a random Gaussian white noise with mean amplitude corresponding to 0.5% of the expected measurement value.



(a) (b)

Fig. 1. The Hanoi (a) and Limassol (b) WDNs used as benchmarks to evaluate the sensor placement method.

5.1 Tests on the Hanoi WDN

Both GA and PSO approaches are applied to the Hanoi example. The parameters for both algorithms were selected after several trial and error tests. In the GA, the seed was set to 100 and 3 iterations were allowed in order to increase the efficiency of the method with a maximum of 3 generations per iteration. In the PSO method, the initialization matrix was set with a size of 50 rows and 10 iterations were allowed in order to increase the efficiency, with a maximum of 50 generations for each of them.

Results are shown in Table 1 and Table 2, where only a single instant of time was considered. As one can see, both methods find optimal solutions in all cases. However, note that sometimes the solution corresponds to a different sensor combination, but resulting in an equivalent efficiency. Finally, PSO tends to be substantially faster on these scenarios.

Table 1. Sensor placement in Hanoi WDN using GA and a single instant of time.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec)
2	12, 21	5	93.1	69.9
3	12, 21, 27	1	98.6	113.9
4	1, 12, 21, 29	0	100.0	119.9

Table 2. Sensor placement in Hanoi WDN using PSO and a single instant of time.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec)
2	12, 21	5	93.1	17.3
3	12, 14, 21	1	98.6	38.0
4	1, 12, 21, 24	0	100.0	57.5

GA and PSO approaches were also applied to the Hanoi example when considering a 24 hours time horizon to build the LSS model. Results are shown in Table 3 and Table 4, where it can be seen that the efficiency tends to increase compared to the tests based on a single instant of time, but only slightly and at the expense of an important increase in computational cost.

Table 3. Sensor placement in Hanoi WDN using GA and a time horizon analysis.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec$\times 10^3$)
2	12, 21	7	93.5	2.2
3	12, 14, 21	1.08	98.8	4.1
4	1, 12, 21, 27	0.08	100.0	3.7

Table 4. Sensor placement in Hanoi WDN using PSO and a time horizon analysis.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec$\times 10^3$)
2	12, 21	7	93.5	0.5
3	12, 14, 21	1.08	98.8	1.2
4	1, 12, 21, 27	0.08	100.0	1.3

5.2 Tests on the Limassol WDN

GA and PSO approaches where applied to the more complex case of the Limassol network. For the GA a seed size of 100 was chosen with 10 iterations and 5 generations per iteration. For the PSO, a seed size of 50 was selected with 150 iterations and 50 generations per iteration. The results obtained considering only a single instant of time are shown in Table 5 and Table 6.

Table 5. Sensor placement in Limassol WDN using GA and a single instant of time.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec$\times 10^3$)
2	8, 152	253	60.6	11.0
3	75, 120, 152	145	74.1	19.6
4	76, 82, 120, 152	114	78.0	31.2
5	75, 82, 120, 126, 152	106	79.0	36.8

Table 6. Sensor placement in Limassol WDN using PSO and a single instant of time.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec$\times 10^3$)
2	8, 152	253	60.6	3.3
3	76, 115, 152	151	74.0	6.0
4	69, 82, 119, 152	131	76.2	15.1
5	70, 115, 125, 152, 188	122	77.4	17.4

First, let remark that, with our settings, PSO works faster than GA but when the problem increases, i.e. when more sensors are installed and more combinations are possible, GA finds better combinations. From additional experiments not reported in the table, we see that even by increasing the number of iterations, the PSO tends to be trapped in a local suboptimum. It may be explained by the fact that PSO has memory of past successes and tends to explore around these recorded configurations whereas when it is necessary to leap from one region to a distant other region, crossover operations like those in GA are probably preferred.

GA and PSO were also applied to the Limassol network while incorporating a time horizon analysis. However, only the case of 2 and 3 sensors were considered since the computational time becomes prohibitive for a higher number of sensors. Moreover, for the GA method, only 1 iteration *it* is authorized and 15 iterations for the PSO instead of the 10 and 150 iterations used respectively for the single time instant case. Results are summarized in Table 7 and Table 8. It shows that contrary to the Hanoi case, for this more complex scenario, the incorporation of a time horizon significantly improves the efficiency. It is remarkable to see that the LSS based leak location method relying on only 3 sensors and the GA method is able to locate correctly almost 85% of the leaks. However, the time horizon analysis clearly comes at the expense of an important additional computational cost.

Table 7. Sensor placement in Limassol WDN using GA and a time horizon analysis.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec×10^3)
2	124, 153	424.8	77.7	32.3
3	76, 133, 169	289.8	84.9	84.8

Table 8. Sensor placement in Limassol WDN using PSO and a time horizon analysis.

Sensors	Node indexes	Overlaps	Efficiency (%)	Time (sec×10^3)
2	124, 153	424.8	77.7	27.5
3	75, 122, 117	306.2	84.8	66.4

6. CONCLUSION

This paper describes a new approach to place a given set of sensors in a WDN in order to maximize the efficiency of the leak location. The solution is motivated by the application of an original leak representation method based on what is called the leak signature space. Genetic Algorithms and Particle Swarm Optimization are applied to solve the underlying optimization problem. Results demonstrate the validity of the approach. They also show that PSO tends to give results faster than the GA approach being very effective for small networks or few sensors. However, the solutions found by the GA tend to correspond to better placements and with higher efficiency for larger networks or more sensors. It also appears that a time horizon analysis can significantly improve the performance in case of complex scenarios. Finally, the negative correlation between the optimization function and the efficiency of the leak detection method is verified.

As future work, we would like to extend this work to more complex WDN as the Barcelona network used in (Casillas et al., 2014a), consisting of 3320 nodes. Finally, we have seen that even with the stochastic optimization tools proposed in this paper, the application of the time horizon analysis was not tractable for the Limassol network with more than 3 sensors. We would like to investigate new approaches to tackle this limitation and obtain a better leak location efficiency.

ACKNOWLEDGEMENTS

This work is supported by the Research Chair in Supervision and Advanced Control of Tecnológico de Monterrey, Campus Monterrey and by a CONACYT studentship. This work has also been partially funded by the Spanish Ministry of Science and Technology through the Project ECOCIS (Ref. DPI2013-48243-C2-1-R) and Project HAR-CRICS (Ref. DPI2014-58104-R), and by EFFINET grant FP7-ICT-2012-318556 of the European Commission.

REFERENCES

Casillas, M.V., Puig, V., Garza-Castañón, L.E., and Rosich, A. (2013). Optimal sensor placement for leak location in water distribution networks using genetic algorithms. *Sensors*, 13(11), 14984–15005. doi:10.3390/s131114984. URL `http://www.mdpi.com/1424-8220/13/11/14984`.

Casillas, M.V., Garza-Castañón, L., and Puig, V. (2012). Extended-horizon analysis of pressure sensitivities for leak detection in water distribution networks. In *8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, 570–575. Elsevier.

Casillas, M.V., Garza-Castañón, L., and Puig, V. (2014a). Model-based leak detection and location in water distribution networks considering an extended-horizon analysis of pressure sensitivities. *Journal of Hydroinformatics*, 16(3), 649 – 670.

Casillas, M.V., Garza-Castañón, L., Puig, V., and Vargas-Martinez, A. (2014b). Linear model representation for robust leak location in water distribution networks. In *Memorias del XVI Congreso Latinoamericano de Control Automático pp. 230 - 235.*

Chen, S. (2009). Another particle swarm toolbox. In *MATLAB Central File Exchange*. URL `http://www.mathworks.com/matlabcentral/fileexchange/25986-another-particle-swarm-toolbox`.

Colombo, A.F., Lee, P., and Karney, B.W. (2009). A selective literature review of transient-based leak detection methods. *Journal of Hydro-environment Research*, 212–227.

Fujiwara, O. and Khang, D.B. (1990). A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resources Research*, 26(4), 539–549. doi:10.1029/WR026i004p00539. URL `http://dx.doi.org/10.1029/WR026i004p00539`.

Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., and Peralta, A. (2011). Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. *Control Engineering Practice*, 19(10), 1157 – 1167. doi:10.1016/j.conengprac.2011.06.004. URL `http://www.sciencedirect.com/science/article/pii/S0967066111001201`.

Pudar, R.S. and Liggett, J.A. (1992). Leaks in pipe networks. *Journal of Hydraulic Engineering*, 118(7), 1031–1046.

Rossman, L. (2000). Epanet 2 user's manual. *United States Envionmental Protection Agency.*

Sarrate, R., Nejjari, F., and Rosich, A. (2012). Sensor placement for fault diagnosis performance maximization in distribution networks. In *18th Mediterranean Conference on Control and Automation (MED)*, 1–6.

Todini, E. and Pilati, S. (1988). Computer applications in water supply: vol. 1—systems analysis and simulation. chapter A gradient algorithm for the analysis of pipe networks, 1–20. Research Studies Press Ltd., Taunton, UK.

Zheng, F., Z.A. and Simpson, A. (2013). Self-adaptive differential evolution algorithm applied to water distribution system optimization. *Journal of Computing in Civil Engineering*, 27(2), 148–158.