

Sensor Placement for Leak Monitoring in Drinking Water Networks combining Clustering Techniques and a Semi-Exhaustive Search

Ramon Sarrate, Joaquim Blesa and Fatiha Nejjari

Abstract—This paper presents an optimal sensor placement strategy based on pressure sensitivity matrix analysis and a semi-exhaustive search strategy that maximizes some diagnosis specifications for water distribution networks. A mean average worst leak expansion distance has been proposed as a new leak location performance measure. The approach is combined with a clustering technique in order to reduce the size and the complexity of the sensor placement problem. The strategy is successfully applied to determine the location of a set of pressure sensors in a district metered area (DMA) in the Barcelona water distribution network (WDN).

I. INTRODUCTION

Water loss due to leaks in pipelines has negative operational, economic, public health and social impacts for both water companies and society. Leaks in WDNs that can happen due to damages and defects in pipes, lack of maintenance or increase in pressure must be detected and located as soon as possible to minimize their effects. Continuous improvements in water loss management are being applied, and new technologies are developed to achieve higher levels of efficiency [1].

The leak location activity is usually supported by ground-penetrating radar or acoustic listening devices [2] that are expensive and not very effective in some kind of pipes like the plastic ones or with large diameters. On the other hand, the leak location techniques based on pressure/flow monitoring devices allow a more effective and less costly search in situ. The need to identify the location of leaks has promoted the development of several techniques based on inverse problems and solving it using pressure or flow measurements. These techniques are based on the sensors installed in the network. Ideally, a sensor network should be configured to facilitate leak detection and location and maximize diagnosis performance under cost limit. Since improper selections may seriously hamper diagnosis performance, the development of sensor placement strategies has become an important research issue in

This work has been funded by the Spanish Government (MINECO) through the project CICYT ECOCIS (ref. DPI2013-48243-C2-1-R), by MINECO and FEDER through the project CICYT HARCRCIS (ref. DPI2014-58104-R), by MINECO through the grant IJCI-2014-2081, by the DGR of Generalitat de Catalunya (SAC group Ref. 2014/SGR/374), by the AGAUR of Generalitat de Catalunya through the grant 2014PDJ00102 and by EFFINET grant FP7-ICT-2012-318556 of the European Commission.

R. Sarrate and F. Nejjari are with the Automatic Control Department, Universitat Politècnica de Catalunya, Rambla de Sant Nebridi, 10, 08222 Terrassa, Spain. email: {ramon.sarrate, fatiha.nejjari}@upc.edu

J. Blesa is with the Institut de Robòtica i Informàtica Industrial (CSIC-UPC). Carrer Llorens Artigas, 4-6, 08028 Barcelona. email: joaquim.bleesa@upc.edu

recent years. Different works that deal with sensor placement for leak location purposes have been published in the last few years. [3] selected the best locations by identifying the best trade-off between reliability and deployment costs. In [4] an algorithm based on sensitivity matrix analysis has been developed to determine where to install a specific number of pressure sensors among hundreds of possible locations in the WDN in order to maximize the capability of detecting and isolating leaks using structural analysis combined with clustering techniques. In [5] a robustness analysis of the sensor placement problem in WDNs has been addressed. The study has been achieved by optimal sensor placement strategies for different leak magnitudes and DMA operating points and evaluated through a robustness percentage index.

Model-based leak location techniques are based on comparing the data gathered by the sensors with the data simulated using a hydraulic model of the district metered area. The sensor placement problem can be roughly stated as choosing a subset, from a given candidate sensor location set, such that some diagnosis performance is guaranteed or at least maximized. Since installing sensors will involve a cost for the drinking water management company, economic constraints must be additionally taken into account in the choice. Sensor placement entails formulating a combinatorial optimization problem. A WDN may easily involve several thousands of candidate sensor locations, which poses a severe optimization challenge. Thus, the sensor placement methodology should be able to cope with such complexity issues.

In this paper, a methodology combining a clustering technique and a semi-exhaustive search through the fault sensitivity matrix concept is proposed for the sensor problem in WDNs. The proposed technique consists in solving a combinatorial optimization problem where the best leak location performance can be achieved by installing a given number of sensors. Since an exhaustive search is highly inefficient a semi-exhaustive search is used to improve the efficiency of the sensor placement techniques. On the other hand, some leak location performance metrics to assess the uncertainty degree of the leak location diagnosis in terms of distance among indiscernible leak locations are defined. The fault sensitivity matrix can be obtained by convenient manipulation of model equations as long as leak effects are included in them [6]. Alternatively, it can be obtained by sensitivity analysis through simulation.

The paper is organized as follows: Section II introduces the model-based leak detection and location approach. The

sensor placement problem formulation is presented in Section III. In Section IV, the sensor placement methodology is applied to a real DMA network in Barcelona. Finally, some conclusions and remarks are given in Section V.

II. MODEL-BASED LEAK DETECTION AND LOCATION IN WDN

A. Fault sensitivity matrix

Model-based fault diagnosis techniques are applied to detect and locate leaks in WDNs. In model-based fault diagnosis [7] a set of residuals are designed based on a process model. Fault detection and isolation is achieved through the evaluation of residual expressions under available measurements. A threshold-based test is usually implemented in order to cope with noise and model uncertainty effects. At the absence of faults, all residuals remain below their given thresholds. Otherwise, when a fault is present the model is no longer consistent with the observations (known process variables). Thus, some residuals will exceed their corresponding thresholds, signalling the occurrence of a fault.

Residual fault sensitivities are a key issue for fault diagnosis. Given a set of m target faults $f_j \in \mathbf{F}$ and a set of n residuals $r_i \in \mathbf{R}$, residual fault sensitivities are collected in the *Fault Sensitivity Matrix* (FSM)

$$\mathbf{\Omega} = \begin{pmatrix} \frac{\partial r_1}{\partial f_1} & \cdots & \frac{\partial r_1}{\partial f_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_n}{\partial f_1} & \cdots & \frac{\partial r_n}{\partial f_m} \end{pmatrix}. \quad (1)$$

The FSM can be obtained by convenient manipulation of model equations as long as fault effects are included in them [6]. Alternatively, it can be obtained by sensitivity analysis through simulation [8]. The latter approach is used in the present paper. Just primary residuals are regarded. Primary residuals are obtained by comparing each actual pressure measurement p_i to the corresponding estimated value in the fault free case \hat{p}_{i0}

$$r_i = p_i - \hat{p}_{i0} \quad (2)$$

A model of the WDN is used by a simulation engine to produce the estimated node pressure. An approximate procedure to obtain the FSM involves using as well the simulator to estimate pressure measurements \hat{p}_{ij} for every node i under fault condition f_j

$$\mathbf{\Omega} \approx \begin{pmatrix} \hat{p}_{11} - \hat{p}_{10} & \cdots & \hat{p}_{1m} - \hat{p}_{10} \\ \vdots & \ddots & \vdots \\ \hat{p}_{n1} - \hat{p}_{n0} & \cdots & \hat{p}_{nm} - \hat{p}_{n0} \end{pmatrix}. \quad (3)$$

Thus, every FSM column corresponds to an estimation of the residual vector in every leak condition. The same nominal leak magnitude is assumed in all simulations. This leak magnitude is not considered in the FSM since it has a scaling factor role.

B. Leak detection and location

A fault can be detected as long as there exists at least a residual sensitive to it. However, isolating faults requires more than one residual being sensitive to them. Fault isolation is achieved by matching the evaluated residual vector pattern to the closest residual fault sensitivity vector pattern (i.e., FSM column vector).

Sometimes a binary version of the FSM is used in the leak location procedure [8]. Then, leak location is achieved by looking for the smallest Hamming distance between FSM columns and the binarized actual residual vector. This and several other alternative leak location methods are compared in [9].

In the present paper, a projection based method is considered. Let $\mathbf{r} = [r_1 \cdots r_n]^T$ be the actual residual vector corresponding to all pressure measurement points, and $\omega_{\bullet j}$ be the column of $\mathbf{\Omega}$ corresponding to leak j . Then, leak location is achieved by solving the problem

$$\arg \max_j \frac{\omega_{\bullet j}^T \cdot \mathbf{r}}{\|\omega_{\bullet j}\| \|\mathbf{r}\|}, \quad (4)$$

where $\|\mathbf{v}\|$ stands for the Euclidean norm of vector \mathbf{v} . Thus, the biggest normalized projection of the actual residual vector on the fault sensitivity space is sought.

In order to assess the leak detectability property, the detectable leak set \mathbf{F}_D was defined in [10] in terms of sensitivity analysis as follows: given a set of sensors \mathbf{S} , a set of leaks \mathbf{F} and the corresponding leak (fault) sensitivity matrix $\mathbf{\Omega}$, the set of detectable leaks $\mathbf{F}_D(\mathbf{S})$ is defined as

$$\mathbf{F}_D(\mathbf{S}) = \{f_j \in \mathbf{F} : \exists r_i \in \mathbf{R} : |\omega_{ij}| \geq \epsilon\}, \quad (5)$$

where ϵ is a threshold to account for noise and model uncertainty.

Regarding the leak locatability performance, considering leak location implemented by means of (4), a uniform projection angle $\bar{\alpha}$ defined as the average between the residual leak sensitivity vectors for all leak pairs was proposed in [10].

The resulting sensor locations led to a maximal uniform projection angle $\bar{\alpha}$. In an ideal case, all pairs of leak sensitivity vectors in the FSM (columns of $\mathbf{\Omega}$) should satisfy this uniform projection angle. This uniform angular separation between leak pairs would allow for a successful leak location method applying (4), even when residuals are affected by modeling errors, sensor noise and other uncertainties.

Nevertheless, in a real case the angle between leak pairs is not uniformly distributed. Some leaks can have similar leak sensitivity vectors, which introduces uncertainty in the leak location results when applying (4). This can become a critical issue for water network utilities, especially when this uncertainty involves distant leak locations, i.e. two distant leaks that have similar leak sensitivity vectors. So, distances between nodes with a similar leak sensitivity vector should be considered in the optimal sensor placement methodologies. In order to take

into account these distances, the following properties are defined.

Definition 1 (Leak expansion set). Given a leak $f_j \in \mathbf{F}$ and a projection angle threshold α_{th} , the leak expansion set $\mathbf{F}_j^{\alpha_{th}}$ is defined as

$$\mathbf{F}_j^{\alpha_{th}} = \{f_i \in \mathbf{F} : \frac{\omega_{\bullet j}^T \cdot \omega_{\bullet i}}{\|\omega_{\bullet j}\| \|\omega_{\bullet i}\|} > \cos(\alpha_{th})\}. \quad (6)$$

$\mathbf{F}_j^{\alpha_{th}}$ contains the set of leaks whose correlation with leak f_j is bigger than $\cos(\alpha_{th})$. If $f_i \in \mathbf{F}_j^{\alpha_{th}}$, it follows that $f_j \in \mathbf{F}_i^{\alpha_{th}}$.

Definition 2 (Correlated leak pairs ratio). Given the leak expansion sets $\mathbf{F}_j^{\alpha_{th}}$ $j = 1, \dots, |\mathbf{F}|$, the correlated leak pairs ratio $\eta^{\alpha_{th}}$ is defined as

$$\eta^{\alpha_{th}} = 100 \frac{\sum_{j=1}^{|\mathbf{F}|} |\mathbf{F}_j^{\alpha_{th}}| - |\mathbf{F}|}{2 \binom{|\mathbf{F}|}{2}}. \quad (7)$$

$\eta^{\alpha_{th}}$ provides the percentage of leak pairs from \mathbf{F} whose mutual correlation is bigger than $\cos(\alpha_{th})$ and $\binom{|\mathbf{F}|}{2}$ represents the total number of leak pairs.

Definition 3 (Leak node distance matrix). Given the geographical coordinates of every leak node, the leak node distance matrix $\mathbf{D} \in \mathbf{R}^{|\mathbf{F}| \times |\mathbf{F}|}$ is defined as the matrix whose coefficients d_{ij} are the geographical distance between nodes i and j .

Matrix \mathbf{D} is a symmetric matrix ($d_{ij} = d_{ji}$), with diagonal coefficients equal to zero ($d_{ii} = 0$). This matrix will be used to compute distances in leak expansion sets.

Definition 4 (Worst leak expansion distance). Given a leak expansion set $\mathbf{F}_j^{\alpha_{th}}$ and the leak node distance matrix \mathbf{D} , the worst leak expansion distance $R_j^{\alpha_{th}}$ is defined as

$$R_j^{\alpha_{th}} = \max_{f_i \in \mathbf{F}_j^{\alpha_{th}}} d_{ij}, \quad (8)$$

$R_j^{\alpha_{th}}$ provides the maximum Euclidian distance between the node of leak f_j and the nodes of leaks whose correlation with leak f_j is bigger than $\cos(\alpha_{th})$. This metric is next used to compute the following overall leak location uncertainty index in terms of leak node distances.

Definition 5 (Average worst leak expansion distance). Given a set of leaks \mathbf{F} and a projection angle threshold α_{th} , leak expansion sets $\mathbf{F}_j^{\alpha_{th}}$ with $j = 1, \dots, |\mathbf{F}|$ can be computed applying (6). Then, the average worst leak expansion distance can be computed as

$$\bar{R}^{\alpha_{th}} = \frac{1}{|\mathbf{F}|} \sum_{j=1}^{|\mathbf{F}|} R_j^{\alpha_{th}}, \quad (9)$$

$\bar{R}^{\alpha_{th}}$ provides the average of the worst leak expansion distances considering all the possible leaks in \mathbf{F} .

As was discussed in [11], the greater the threshold α_{th} , the greater the uncertainty in terms of leak expansion distance and number of correlated leak pairs. The choice of this threshold should take into account implementation requirements of the leak location software module, as well as practical issues concerning the water utility maintenance procedures. On the one hand, the leak location software module will have to deal with sensor measurement noise and network modeling uncertainty. Therefore, the bigger the threshold is, the better the performance of the leak location procedure is. On the other hand, for the water utility maintenance department, a smallest leak location result uncertainty is preferable. Indeed, upon the occurrence of a leak, the leak location software module will provide a set of leak node candidates to the maintenance department, which then will undergo leak field-testing. Thus, the smaller the leak expansion distance is, the better it is, which involves specifying a smaller projection angle threshold. Therefore, a tradeoff exists between both criteria.

In order to find a good balanced solution, it is expected to find a sensor placement solution suitable for a range of projection angle thresholds as proposed in [11].

Definition 6 (Mean average worst leak expansion distance). Given a set $\mathbf{A}_{th} = \{\alpha_{th}^1, \dots, \alpha_{th}^{n_\alpha}\}$ that covers a suitable range of projection angle thresholds. Then, the mean average worst leak expansion distance over this set can be computed as

$$\bar{\bar{R}}^{\alpha_{th}} = \frac{1}{|\mathbf{A}_{th}|} \sum_{\alpha_{th} \in \mathbf{A}_{th}} \bar{R}^{\alpha_{th}}, \quad (10)$$

where the average worst leak expansion distance $\bar{R}^{\alpha_{th}}$ is computed applying (9) to every projection angle threshold in \mathbf{A}_{th} .

III. SENSOR PLACEMENT METHODOLOGY

A. Problem statement

The sensor placement problem for leak location involves solving a combinatorial optimization problem where the best leak location performance that can be achieved installing the cheapest number of sensors is sought. Additionally, a budget constraint is imposed by the water utility, which must be taken into account in the sensor placement methodology.

Let \mathbf{S} be the candidate pressure sensor set and m_p the maximum number of pressure sensors that can be installed in the network according to the budget constraint. Although fewer pressure sensors could be installed in the network, in this work it is assumed that m_p sensors should be installed. Then, the problem can be roughly stated as the choice of a configuration of m_p pressure sensors in \mathbf{S} such that the best diagnosis performance is attained. This diagnosis performance depends on the set of sensors installed in the network $S \subseteq \mathbf{S}$ and will be stated in terms

of the detectable leak set and the mean average worst leak expansion distance, i.e., $\mathbf{F}_D(S)$ and $\bar{R}^{\alpha_{th}}$.

To solve the sensor placement problem, some network model information is also required. On the one hand, the leak node distance matrix \mathbf{D} is assumed to be previously obtained based on the geographical coordinates of every leak node. On the other hand, the leak sensitivity matrix Ω corresponding to the complete set of candidate sensors is assumed to be previously computed, following the methodology described in Section II-B. Additionally, a set of projection angle thresholds \mathbf{A}_{th} must be provided according to the criteria specified in Section II-B. Hence, the optimal sensor placement for leak diagnosis can be formally stated as follows:

GIVEN a candidate sensor set \mathbf{S} , a leak node distance matrix \mathbf{D} , a leak sensitivity matrix Ω , a leak set \mathbf{F} , a set of projection angle thresholds \mathbf{A}_{th} and a number m_p of pressure sensors to be installed.

FIND the m_p -pressure sensor configuration $S \subseteq \mathbf{S}$ such that:

- 1) all leaks in \mathbf{F} are detectable, i.e. $\mathbf{F}_D(S) = \mathbf{F}$ according to (5), and
- 2) the mean average worst leak expansion distance is minimized, i.e. $\bar{R}^{\alpha_{th}}(S) \leq \bar{R}^{\alpha_{th}}(S^*)$ for any $S^* \subseteq \mathbf{S}$ such that $|S^*| = m_p$.

B. Optimal sensor placement algorithm

As the optimization problem stated in the previous section cannot be solved by efficient branch and bound search strategies, a suboptimal two-step hybrid methodology that combines clustering techniques with an exhaustive search was proposed in [10]. The cluster analysis allowed reducing the initial candidate sensor set cardinality such that the exhaustive search could be solved in a reasonable time. Such reduction involved partitioning the initial candidate sensor set \mathbf{S} into ℓ clusters, and selecting N candidate sensors of each cluster. In particular, a number of clusters equal to the number of sensors to be installed was imposed, i.e., $\ell = m_p$. Thus, the exhaustive search was applied to a reduced candidate sensor set of $N \times m_p$ sensors.

However, since an exhaustive search is highly inefficient, in this paper a semi-exhaustive search will be proposed in order to improve efficiency. Both methods are next stated, and they will be compared in Section IV.

Exhaustive method

Step 1 Clustering techniques are applied to reduce the initial set of candidate sensors \mathbf{S} to \mathbf{S}' , with $|\mathbf{S}'| = N \times m_p$.

Step 2 An exhaustive search is applied to the reduced candidate sensor set \mathbf{S}' . This search implies that the diagnosis performance must be evaluated $\binom{|\mathbf{S}'|}{m_p}$ times.

Semi-exhaustive method

Step 1 Clustering techniques are applied to reduce the initial set of candidate sensors \mathbf{S} to \mathbf{S}' , with $|\mathbf{S}'| = N \times m_p$.

Step 2 A semi-exhaustive search is applied to the reduced candidate sensor set \mathbf{S}' so that m_p -sensor configurations involving just one sensor from every cluster are considered. This search implies that now the diagnosis performance must be just evaluated $\binom{|\mathbf{S}'|}{m_p}$ times.

Thus, in the semi-exhaustive method more confidence is provided to the clustering procedure within the sensor placement methodology. Next section recalls several clustering approaches developed in previous works that solve Step 1 in both methods.

C. Candidate sensor set reduction by clustering analysis

The use of clustering analysis to group pressure candidate sensors in WDN for reducing the number of candidate sensors was initially proposed in [4]. The k-means clustering algorithm [12] was applied to the normalized rows of leak sensitivity matrix Ω defined in (1) to group the initial number of candidate sensors $|\mathbf{S}|$ into as many clusters as the chosen cardinality of the reduced sensor candidate set ($|\mathbf{S}'|$). Once the candidate sensors were partitioned into clusters, the sensors associated with the normalized rows with the smallest Euclidean distance to the cluster centroids were selected as a member of the reduced candidate sensor set \mathbf{S}' . In [10], the Evidential C-Means (ECM) algorithm [13] and normalized rows of leak sensitivity matrix were used to group the initial number of candidate sensors $|\mathbf{S}|$ into as many clusters as the number of sensors to be installed (m_p). The difference between the aforementioned clustering algorithms is that k-means belongs to closed data sets methods, which do not allow overlapping of clusters. While the ECM algorithm is based on the belief functions framework [14] and provides the centroids of the clusters and the degree of membership (plausibility) of every element (sensor) to every cluster. The plausibility of the elements was chosen in [10] to select the N more representative elements of every group, providing the $N \times m_p$ elements of the reduced candidate set \mathbf{S}' . Finally, in [11] the same clustering strategy as in [10] was proposed. But in order to maximize the leak detectability, in addition to the plausibility of the elements, the Euclidean norm of the rows of the fault sensitivity matrix was taken into account in [11] to choose the N most representative sensors of every cluster. This last clustering-based method will be the one used in this paper to reduce the number of sensor candidates.

IV. APPLICATION TO A REAL WDN

A. DMA case study

The considered DMA is located in Barcelona city with 883 nodes, 927 pipes and two inflow inputs modeled as reservoir nodes. In the network, we consider the existence of 448 potential leaks, corresponding to dummy nodes, that should be detected and located. In order to reduce the problem complexity, a subset of node pressures, corresponding to 311 nodes with demand, is chosen as the

candidate sensor set. It is also assumed that there is no sensor already installed in the network before solving the sensor placement problem.

A fault sensitivity matrix has been obtained using the EPANET hydraulic simulator. Leaks are simulated in EPANET through the corresponding emitter coefficient, which is designed to model fire hydrants/sprinklers, and it can be adapted to provide the desired leak magnitude in the network.

Given a set of boundary conditions EPANET software has been firstly used to estimate the steady-state pressure at the 311 candidate sensor nodes. Next, 448 leaks have been simulated in the dummy nodes and the steady-state pressure has been estimated again in the 311 sensor nodes. Finally, a fault sensitivity matrix has been obtained as the pressure difference between the fault free case and each faulty situation. In this case, the fault sensitivity matrix has been computed for a leak magnitude of 1.5 lps (liters per second).

B. Sensor placement analysis

Assume that the water distribution company has assigned a baseline budget for investment on instrumentation that just makes it possible to install 5 pressure sensors. Hence, 5 pressure sensors should be chosen out of 311 such that all leaks are detectable and the mean average worst leak expansion distance is minimised. Recall from Section III-A that, in addition to the fault sensitivity matrix and the leak set, a leak node distance matrix and a set of projection angle thresholds should be specified. Matrix \mathbf{D} has been obtained from geographical data contained in the EPANET model and the following projection angle threshold set has been considered: $\mathbf{A}_{th} = \{10, 20, 30, 40, 50, 60\}$.

According to the methodology described in Section III-B, an exhaustive and a semi-exhaustive search will be applied to a reduced set \mathbf{S}' of candidate pressure sensors. A cardinality of 25 pressure sensors has been considered for this DMA.

In order to reduce the number of candidate pressure sensors from 311 to 25, clustering techniques have been applied to the data set (311 normalized rows of the fault sensitivity matrix $\mathbf{\Omega}$) as described in [11]. First, ECM clustering algorithm has been used to classify the data set into $\ell = 5$ different clusters (since $m_p = 5$). The algorithm takes 30 seconds in the classification procedure to obtain 5 clusters.

With the plausibility matrix obtained from the clustering algorithm, a hard partition has been obtained by assigning each element to its highest plausibility cluster. Fig. 1 depicts in different colors the 5 different network node clusters, where the closest node to the centroid have been highlighted in every cluster. Finally, the most N representative sensors of every cluster have been chosen according to the criteria proposed in [11] with $N = 5$, such that $N \times \ell = |\mathbf{S}'|$.

The reduced set \mathbf{S}' with 25 candidate pressure sensors, partitioned according to their corresponding cluster, is dis-

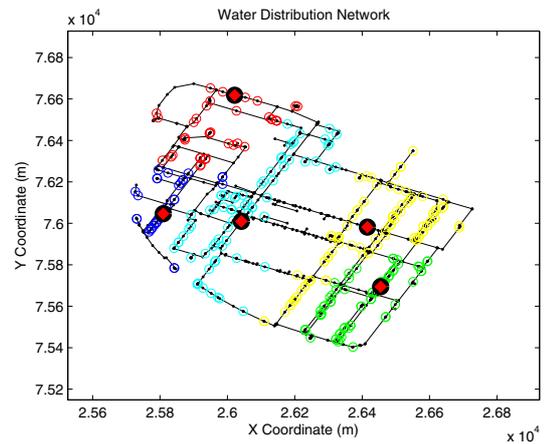


Fig. 1. Clustering results

played in Fig. 2. In the figure, a different color corresponds to each cluster.

The exhaustive search is first applied to solve the sensor placement problem, providing the set of 5 pressure sensor places signaled as magenta starred nodes in Fig. 2. Installing these pressure sensors, all 448 leaks are detectable and the mean average worst leak expansion distance amounts to 698.23 m. In the exhaustive search all $\binom{25}{5} = 53130$ possible sensor configurations are evaluated requiring more than 21 h of computation time.

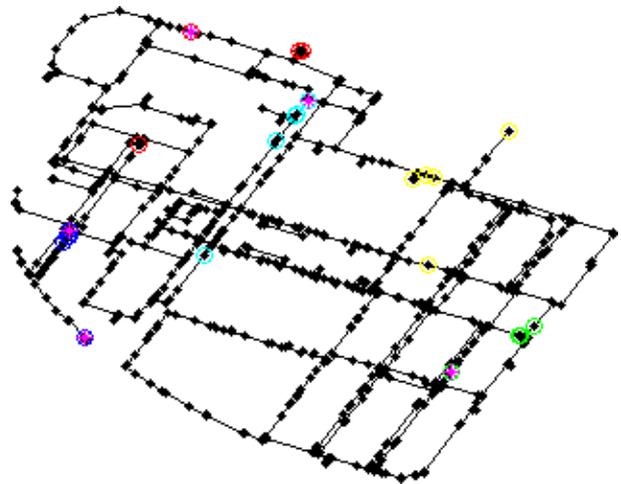


Fig. 2. DMA network sensor placement results (exhaustive search)

The semi-exhaustive search is next applied to solve the sensor placement problem, providing the set of 5 pressure sensor places signaled as magenta starred nodes in Fig. 3. As expected, the solution contains one sensor from each cluster. Installing these pressure sensors, all 448 leaks are detectable and the mean average worst leak expansion distance amounts to 717.42 m. In the semi-exhaustive search just $\binom{25}{5} = 3125$ possible sensor configurations are evaluated requiring approximately 76 min of computation time.

Remark that the semi-exhaustive method cannot guaran-

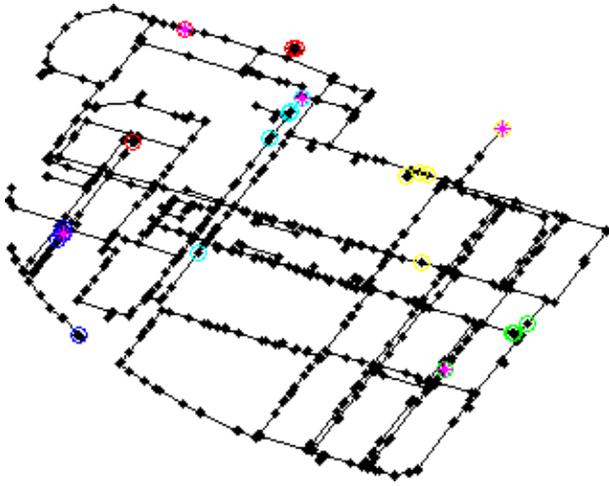


Fig. 3. DMA network semi-exsensor placement results (semi-exhaustive search)

tee global optimality. In fact, the performance index of the solution increases 2.75% in comparison to the exhaustive method. However, the time required to find a solution has drastically decreased (i.e, a 94% reduction). This clearly proves the benefit of the semi-exhaustive approach over the exhaustive one.

V. CONCLUSIONS

This work presents an optimal sensor placement approach based on pressure sensitivity matrix analysis and a semi-exhaustive search strategy combined with clustering techniques to reduce the size and complexity of the sensor placement problem. The goal is to characterize and determine a sensor set that guarantees a maximum degree of leak locatability while a budgetary constraint is satisfied. A mean average worst leak expansion distance has been proposed as a new leak location performance measure. This metric aggregates the normalized projection degree between the residual fault sensitivity vectors for all fault pairs.

The strategy is successfully applied to a DMA of the Barcelona WDN using a semi-exhaustive and an exhaustive search. Although the semi-exhaustive method cannot guarantee global optimality, the time required to find a solution has decreased with respect to the exhaustive search which clearly proves the benefit of the semi-exhaustive approach over the exhaustive one. As a future work, genetic algorithms or other types of optimization methods that provide some guarantee regarding the solution optimality will be investigated.

REFERENCES

- [1] R. Puust, Z. Kapelan, D. Savic, and T. Koppel, "A review of methods for leakage management in pipe networks," *Urban Water Journal*, vol. 7, pp. 25–45, 2010.
- [2] T. Farley, *Losses in Water Distribution Networks*. UK: IWA Publishing UK, 2003.
- [3] A. Candelieri, D. Soldi, D. Conti, and F. Archetti, "Analytical leakages localization in water distribution networks through spectral clustering and support vector machines," *Procedia Eng.*, vol. 89, no. 11, p. 1080–1088, 2014.

- [4] R. Sarrate, J. Blesa, F. Nejari, and J. Quevedo, "Sensor placement for leak detection and location in water distribution networks," *Water Science and Technology: Water Supply*, vol. 14, no. 5, pp. 795–803, 2014.
- [5] J. Blesa, F. Nejari, and R. Sarrate, "Robust sensor placement for leak location: analysis and design," *Journal of Hydroinformatics*, vol. 18, no. 1, pp. 136–148, 2016.
- [6] J. Blesa, V. Puig, and J. Saludes, "Robust identification and fault diagnosis based on uncertain multiple input–multiple output linear parameter varying parity equations and zonotopes," *Journal of Process Control*, vol. 22, no. 10, pp. 1890–1912, 2012.
- [7] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control*, 2nd ed. Springer, 2006.
- [8] R. Pérez, V. Puig, J. Pascual, J. Quevedo, E. Landeros, and A. Peralta, "Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks," *Control Engineering Practice*, vol. 19, no. 10, pp. 1157–1167, 2011.
- [9] M. V. Casillas, L. E. Garza, and V. Puig, "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*, Mexico City, Mexico, August 2012, pp. 570–575.
- [10] R. Sarrate, J. Blesa, and F. Nejari, "Clustering techniques applied to sensor placement for leak detection and location in water distribution networks," in *22th Mediterranean Conference on Control and Automation*, Palermo, Italy, June 2014, pp. 109–114.
- [11] F. Nejari, R. Sarrate, and J. Blesa, "Optimal pressure sensor placement in water distribution networks minimizing leak location uncertainty," *Procedia Engineering*, vol. 119, pp. 953–962, 2015.
- [12] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. on Math. Statist. and Prob.*, Berkeley, Univ. of Calif. Press, 2008, pp. 281–297.
- [13] M. Masson and T. Denoeux, "ECM: An evidential version of the fuzzy c-means algorithm," *Pattern Recognition*, vol. 41, pp. 1384–1397, 2008.
- [14] Z. G. Liu, Q. Pan, and J. Dezert, "A new belief-based k-nearest neighbor classification method," *Pattern Recognition*, vol. 46, no. 3, pp. 834–844, 2013.