2013 Conference on Control and Fault-Tolerant Systems (SysTol)
October 9-11, 2013. Nice, France

Optimal sensor placement for leak location in water distribution networks using genetic algorithms

Myrna V. Casillas[1], Vicenç Puig[2], Luis E. Garza-Castañón[1] and Albert Rosich[3]

Abstract— **In this paper, a new approach for sensor placement in water distribution networks (WDN) is proposed. The sensor placement problem is formulated as an integer optimization problem. The optimization criterion consists in minimizing the number of non-isolable leaks according to the isolability criteria introduced. Because of the non-linear integer and large-scale nature of the resulting optimization problem, genetic algorithms (GA) are used as solution approach. To validate the results obtained, they are compared with exhaustive search methods with higher computational cost proving that GA allow to find near-optimal solutions with less computational load. The proposed sensor placement algorithm is combined with a projection-based isolation scheme. However, the proposed methodology does not depend on the isolation method chosen by the user and it could be easily adapted to any other isolation scheme. Experiments on a real network allow to evaluate the performance of such approach.**

I. INTRODUCTION

Leak location is of great importance for water distribution network systems and represents an important factor for quality in service. In these systems, it is obvious that only a limited number of sensors can be installed due to budget constraints. Thus, the development of a sensor placement strategy has become an important research issue in recent years. Ideally, a sensor network should be configured to facilitate fault detection and maximize diagnosis performance.

Several works have been published on leak detection and isolation methods for Water Distribution Networks (WDN). Andrew et al. [1] present a review of existing transient-based leak detection methods. Model-based leak detection and isolation techniques have also been studied starting with the seminal paper of Pudar and Liggett [2] which formulates the leak detection and isolation problem as a least-squares estimation problem. However, in such non-linear models the parameter estimation of the water networks is not an easy task.

Alternatively, in [3], a method based on pressure measurements and leak sensitivity analysis is proposed. This methodology consists in analyzing the residuals on-line, i.e. the difference between the measurements and their estimation using the network hydraulic models, regarding a given threshold that takes into account the model uncertainty and the noise. When some of the residuals violate their threshold, they are correlated against the leak sensitivity matrix in order to discover which possible leak is present. Although this approach has good efficiency under ideal conditions, its performance decreases in presence of nodal demand uncertainty and noise in the measurements. An improved technique has recently been developed [4] where an extended time horizon analysis of pressure measurements is considered and a comparison between the performances depending on the metric used is performed.

The main objectives of sensor placement are leak detectability, isolability and localization. Leak detectability is the ability of monitoring a variation in pressure due to a loss of water occurring in the network. Leak isolability concerns the capacity of distinguishing between two possible occurrences, whereas leak localization refers to finding the node where the leak is occurring. There are some works devoted to sensor placement for fault detection and isolation (FDI). Some approaches propose to locate sensors based on diagnosticability criteria according to the study of structural matrices [5]. In [6], an optimization method based on binary integer linear programming searches for an optimal set of sensors for model-based FDI.

Each of the previously mentioned works is used in the general framework of FDI. However, there are several contributions dedicated to sensor placement in water distribution networks. In [7], the problem of deploying sensors in a large water distribution network is considered in order to detect the malicious introduction of contaminants. A strategy based on the diagnosticability maximization [8] allows to locate optimally the sensors in distribution networks based on the structural model of the system under consideration. Closer to our research, in [3], an optimal sensor placement is formulated as an integer programming problem where each decision variable x_i associated to a node v_i of the network takes the value 0 or 1 according to the presence or the absence of a sensor installed on this node. This binary representation for sensor placement is used in the latest leak detection works.

This paper proposes a new approach for sensor place-

This work is supported by the Research Chair in Supervision and Advanced Control of Instituto Tecnolgico y de Estudios Superiores de Monterrey and by a CONACYT studentship as well as by the Fond National de la Recherche, Luxembourg (CO11/IS/1206050).

This work has been partially grant-funded by CICYT SHERECS DPI-2011-26243 and CICYT WATMAN DPI- 2009-13744 of the Spanish Ministry of Education and by EFFINET grant FP7-ICT-2012-318556 of the European Commission.

[1]Myrna V. Casillas and Luis E. Garza-Castañón are with Instituto Tecnolgico y de Estudios Superiores de Monterrey, Monterrey, 64849 México mv.casillas.phd.mty at itesm.mx legarza at itesm.mx

[2]Vicenç Puig is with Advanced Control Systems (SAC), Technical University of Catalonia (UPC), Pau Gargallo 5, 08028 Barcelona, Spain. vicenc.puig at upc.edu

[3]Albert Rosich is with Interdisciplinary Centre for Security, Reliability and Trust in the University of Luxembourg, Luxembourg albert.rosich at uni.lu

ment for leak location in WDN that can be used with the projection-based location scheme proposed in [4]. The sensor placement problem is formulated as an integer optimization problem. The optimization criterion consists in minimizing the number of non-isolable leaks according to the isolability criteria introduced. Because of the non-linear integer and large-scale nature of the resulting optimization problem, genetic algorithms (GA) are used as solution approach. The obtained results are compared with exhaustive search methods with higher computational cost proving that GA allow to find near-optimal solutions with less computational load. Another advantage is that our methodology does not dependent of the isolation method chosen by the user. Experiments in a real network allow to evaluate the performance of such approach.

The rest of the document is organized as follows: Section II presents the leak localization methodology in which our work is based. Section III describes the problem formulation. Sections IV and V present the sensor placement algorithms proposed in this work while in Section VI we show the application and the results obtained in a real water distribution network. Finally, Section VII concludes this work.

II. LEAK LOCATION METHODOLOGY

The leak location methodology used in this paper has been introduced in [4] as an extension of the methodology proposed in [3]. The methodology will be summarised here since it is the basis on the top of which the sensor placement algorithm proposed in this paper will be formulated.

The leak location methodology aims to detect and isolate leaks in a water distribution network using pressure measurements and their estimation using the hydraulic network model. Let us consider a water distribution network with m demand nodes and n pressure sensors. The leak detection methodology is based on the computation of the residual vector $\mathbf{r} = [r_1 \ \ldots \ r_n]^T$ where the residual $r_i \in \mathbf{r}$ is the difference between the pressure measurements p_i and its corresponding estimation $\hat{p}_i$ obtained from the simulation of the hydraulic model with no leak, i.e.

$$r_i = p_i - \hat{p}_i \tag{1}$$

for $i = 1, \ldots, n$.

The leak isolation method relies on analyzing the residual vector (1) using a sensitivity analysis which is based on the evaluation of the effect on the available pressure measurement sensors caused by all possible leaks. To perform such sensitivity analysis the following sensitivity vectors are derived from simulated leak scenarios [3]:

$$\mathbf{s}_j = \begin{bmatrix} \frac{\hat{p}_1^{f_j} - \hat{p}_1}{f_j} \\ \vdots \\ \frac{\hat{p}_n^{f_j} - \hat{p}_n}{f_j} \end{bmatrix} \qquad j = 1, \cdots, m \tag{2}$$

where $\hat{p}_i^{f_j}$ and $\hat{p}_i$ are the pressure estimation obtained from the hydraulic model simulation under leak f_j scenario and leak-free scenario, respectively. For the sake of simplicity and without loss of generality, m possible leaks (one for each node) have been assumed. Then, leak isolation is based on the analysis of the residual vector, together with the sensitivity vectors in order to determine which node has the highest plausability to present a leakage. A variety of metrics can be used to perform this isolability analysis [9]. In this work, a method presented in [4] based on projections between residual and sensitivity vectors will be used. According to the mentioned study the angle method (projection considering the inverse of cosine function) presents the best performance for the isolation task. However, it is important to note that the sensor placement approach proposed in this paper could also be applied using any other isolability method based on sensitivity analysis.

Let $\mathbf{r}$ be the residual vector (1) obtained from the pressure sensors installed in the network, then its normalized projections, ψ_j, onto each sensitivity vector are computed as

$$\psi_j = \frac{\mathbf{r}^T \mathbf{s}_j}{|\mathbf{r}||\mathbf{s}_j|} \tag{3}$$

for $j = 1, \ldots, m$. Then, the largest projection will determine the candidate node that presents a leak, i.e. a leak in node k is located if

$$\psi_k = \max(\psi_1, \cdots, \psi_m) \tag{4}$$

III. PROBLEM FORMULATION

The objective of this work is to develop an approach to place a given number of sensors, n, in a water distribution network in order to obtain a sensor configuration with a maximized leak isolabitily performance for a given leak detection and isolated scheme. Here, for illustrative purposes, the one presented in previous section.

It should be noted that the length of the sensitivity and residual vectors depends on the number of sensors n installed in the network. In fact, according to (1) and (2), these vectors will have as many elements as installed sensors. In order to find a sensor configuration that presents maximum isolability performance regarding all the possible leak scenarios, the following residual vectors derived from simulated leak scenarios will be computed:

$$\mathbf{r}_k = \begin{bmatrix} \hat{p}_1^{f_k} - \hat{p}_1 \\ \vdots \\ \hat{p}_n^{f_k} - \hat{p}_n \end{bmatrix} \qquad k = 1, \cdots, m \tag{5}$$

where $\hat{p}_i^{f_k}$ and $\hat{p}_i$ are the pressure estimation obtained from the hydraulic model simulation under leak f_k scenario and leak-free scenario, respectively. Note that the magnitude of the leaks used to compute the sensitivity vectors in (2) and the one used to compute the residual vectors in (5) are chosen different ($f_j \neq f_k$) in order to increase the robustness of the method. Taking into account the mentioned residual and sensitivity vectors, the sensitivity matrix S and the residual matrix R are computed as follows

$$S = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_m \end{bmatrix} \tag{6}$$

$$R = \begin{bmatrix} \mathbf{r}_1 & \cdots & \mathbf{r}_m \end{bmatrix} \tag{7}$$

Notice that, the S and R matrices have been computed assuming all the nodes measured.

To select a configuration with n sensors, the next binary vector is defined

$$\mathbf{q} = \left[\begin{array}{ccc} q_1 & \cdots & q_m \end{array} \right] \quad (8)$$

where $q_i = 1$ if the pressure in node i is measured, and $q_i = 0$ otherwise. In turn, a diagonal matrix $Q(\mathbf{q})$ is constructed from the vector $\mathbf{q}$ as

$$Q(\mathbf{q}) := diag(q_1, \cdots, q_m) \quad (9)$$

Then, given the vector $\mathbf{q}$ denoting which sensors are installed, the corresponding sensitivity and residual vectors can be determined as

$$\mathbf{s}_j(\mathbf{q}) = Q(\mathbf{q})\mathbf{s}_j, \qquad \mathbf{r}_k(\mathbf{q}) = Q(\mathbf{q})\mathbf{r}_k \quad (10)$$

for $j = 1, \cdots, m$, where $\mathbf{s}_j$ and $\mathbf{r}_k$ are the sensitivity and residual vectors obtained with all nodes measured (i.e. $m = n$, the vectors $\mathbf{s}_j$ and the vectors $\mathbf{r}_k$ contains m elements each). Finally, the projections in (3) can be computed depending of the sensors regarded in $\mathbf{q}$ as

$$\psi_{kj}(\mathbf{q}) = \frac{\mathbf{r}_k^T Q(\mathbf{q})\mathbf{s}_j}{|Q(\mathbf{q})\mathbf{r}_k||Q(\mathbf{q})\mathbf{s}_j|} \quad (11)$$

for $j = 1, \cdots, m$. Note that the property $Q(\mathbf{q}^T)Q(\mathbf{q}) = Q(\mathbf{q})$ has been used in (11).

Now we are able to compute the projection matrix Ψ as

$$\Psi(\mathbf{q}) = \left[\begin{array}{ccc} \psi_{11}(\mathbf{q}) & \cdots & \psi_{1m}(\mathbf{q}) \\ \vdots & \ddots & \vdots \\ \psi_{m1}(\mathbf{q}) & \cdots & \psi_{mm}(\mathbf{q}) \end{array} \right] \quad (12)$$

In order to infer how good a sensor configuration is to locate a leak at node $i \in \{1, \cdots, m\}$, the next performance index is introduced:

$$\varepsilon_i(\mathbf{q}) = \begin{cases} 0 & \text{if } \psi_{ii}(\mathbf{q}) = \max(\{\psi_{i1}(\mathbf{q}), \cdots, \psi_{im}(\mathbf{q})\}) \\ 1 & \text{otherwise} \end{cases} \quad (13)$$

such that the performance index $\varepsilon_i = 0$ as long as leak in node i is perfectly located, and $\varepsilon_i = 1$ otherwise.

As the objective is to maximize isolability regarding leaks in all network nodes, the performance index is computed to account for all nodes leak as

$$\bar{\varepsilon}(\mathbf{q}) = \sum_{i=1}^{m} \frac{\varepsilon_i(\mathbf{q})}{m} \quad (14)$$

Notice that $\bar{\varepsilon}(\mathbf{q}) = 0$ as long as a sensor configuration is chosen such that all possible leaks can be perfectly located. Indeed, $\bar{\varepsilon}(\mathbf{q}) * 100$ is the percentage of non isolable leaks.

Based on the vector $\mathbf{q}$ and the extended performance index $\bar{\varepsilon}(\mathbf{q})$ the sensor placement problem is casted as an optimization problem formulated as

$$\begin{array}{l} \min_{\mathbf{q}} \bar{\varepsilon}(\mathbf{q}) \\ \text{s.t.} \\ \\ \sum_{i=1}^{m} q_i = n \end{array} \quad (15)$$

where $\mathbf{q}$ is defined in (8) and $n \in \{1, \ldots, m\}$ is the number of sensors we want to place.

Remark. It should be noticed that the solution of the previous optimization algorithm provides the best sensor location when the leak size that is wanted to be located is close to the value used for evaluating residuals (5). If the leak size is smaller or larger that this value, the optimal sensor location could vary. Moreover, the obtained leak isolation error could be larger that the minimum value obtained as the solution of the optimization problem (15). This motivates for enhancing the sensor placement method suggested in this paper by introducing some robustness against the leak magnitude.

IV. SOLUTION USING A SEMI-EXHAUSTIVE SEARCH

As stated in Section III, the problem of sensor placement involves finding an n-sensor configuration among a set of m candidate nodes. One trivial approach to solve the problem would be to check all the $\binom{m}{n}$ sensor configurations. Here, we propose a first algorithm as an alternative to this trivial methodology in order to ensure the optimal location in a benchmark network.

This method involves the search of the best configuration based on every possible combination but reducing the computation cost by rejecting configurations that is proved they can not be candidates for the optimum. The method is described in Algorithm 1.

The goal of this algorithm is to find the optimal sensor configuration taking into account all the possible combinations of sensors and considering the method that will be used to perform the leak location. First, the algorithm initiates the minimum number of non localizable (NL) leaks min_{NL} found so far to m (line 1). Then, a loop is performed over each possible combination k of sensors configuration (line 2). The binary vector $\mathbf{q}^k$ is evaluated which allows to compute the updated sensitivity and residual matrices, i.e., $\hat{S}^k$ and $\hat{R}^k$ respectively (line 3 and 4) and the current number of NL leaks is initiated to 0 (line 5). Then, the algorithm checks for each potential leak α if it can be located with the current sensor configuration. It evaluates the element (α, α) of the matrix Ψ and for each other column β of row α, it tests if the projection gives a higher score (line 10). If it is the case, then the number of NL leaks is augmented (line 11) and the other columns of the Ψ matrix do not need to be tested (line 12). When the number of NL leaks is higher than the minimum number of NL leaks found so far, i.e. $nb_{NL}^k \geq min_{NL}$, then the current configuration cannot be optimal and the algorithm aborts the evaluation and continues with the next configuration (line 16), improving in this way the computational efficiency of the algorithm. Otherwise, the minimum number of NL leaks is updated by the current number of NL leaks (line 20) and the index of the configuration is taken as best index found so far (line 21). This algorithm performs a semi-exhaustive search in the sense that all the configurations are considered but useless computations are avoided as much as possible.

Algorithm 1 Sensor placement based on semi-exhaustive search

Require: A sensitivity matrix S, a residual matrix R. The number of sensors n, the number of nodes m and a $(d \times n)$ matrix L where $d = \binom{m}{n}$, i.e. each row is a possible combination of sensors position.

Ensure: The optimal sensors configuration of index k_{min} with error $\bar{\varepsilon}_{min}$.

$min_{NL} \leftarrow m$
for $k = 1, \cdots, d$ **do**
  $\mathbf{q}^k \leftarrow eval\_Q(L^k)$ // c.f. (8)
  $\hat{S}^k \leftarrow eval\_S(\mathbf{q}^k, S); \hat{R}^k \leftarrow eval\_R(\mathbf{q}^k, R)$ // c.f. (10)
  $nb^k_{NL} \leftarrow 0$
  for $\alpha = 1, \cdots, m$ **do**
    $\Psi^k_{\alpha\alpha} \leftarrow eval\_\Psi(\hat{S}^k, \hat{R}^k, \alpha)$ // c.f. (11) and (12)
    for $\beta = 1, \cdots, m; \beta \neq \alpha$ **do**
      $\Psi^k_{\alpha,\beta} \leftarrow eval\_\Psi(\hat{S}^k, \hat{R}^k, \alpha, \beta)$
      if $\Psi^k_{\alpha\beta} > \Psi^k_{\alpha\alpha}$ **then**
        $nb^k_{NL} \leftarrow nb^k_{NL} + 1$
        break
      end if
    end for
    if $nb^k_{NL} \geq min_{NL}$ **then**
      break
    end if
  end for
  if $nb^k_{NL} < min_{NL}$ **then**
    $min_{NL} \leftarrow nb^k_{NL}$
    $k_{min} = k$
  end if
end for
$\bar{\varepsilon}_{min} = \frac{min_{NL}}{m}$

V. SOLUTION USING GA

Genetic algorithms (GA) are well known search and optimization tools based on principles of natural genetics and natural selection [10], [11]. Because of their broad applicability, ease of use, and global perspective, GA have been increasingly applied to various search and optimization problems in the recent past. Some fundamental ideas of genetics are borrowed and used artificially to construct search algorithms that are robust and require minimal problem information. GA transform a population of individual objects, each with an associated fitness value, into a new generation of the population using the Darwinian principle of reproduction and survival of the fittest and analogs of naturally occurring genetic operations such as crossover (sexual recombination) and mutation. Each individual in the population represents a possible solution to a given problem. The genetic algorithm attempts to find a very good (or best) solution to the problem by genetically breeding the population of individuals over a series of generations.

The GA can be used in the context of sensor placement in water distribution networks in order to find near-optimal placement of these sensors for leak detection. In that case, a chromosome corresponds to the possible presence or absence of a sensor at a given node.

Here, the sensor placement problem formulated as an optimization problem in Section III is solved using genetic algorithms and implemented using the GA Toolbox of MATLAB. The GA needs to establish a function whose output involves an index to be minimized. In our case, this function corresponds to the evaluation of the error index computed in (14) according to the computation of the projection matrix as in (12). This error depends of the number of maximum values in each row of the matrix that are not elements of the diagonal in the projection matrix.

The pseudo-code of the algorithm is shown in the Algorithm 2. First, we initialize the variables of the GA (line 1) including the number of generations, the bit string type population, the tolerance as 10^{-10}, and the elite count as 1 in order to save one of the previous results analysed. Then, we declare the search restriction (line 2) being that the number of "ones" in the solution corresponds to the number of sensors to install and a seed size z is chosen (line 3). This seed creates an initial matrix with random sensor positions and every location delivered by the GA is tested according to the function declared in the algorithm. The sensor placement is based on the construction of binary vectors where the presence of a "one" represents a sensor located in the correspondent node. This vector allows to select the adequate rows of the matrices S and R in order to compute the projection matrix according to the selected nodes to be measured. Once we have this projection matrix, we look for the maximum value of each row of the matrix, expecting the highest to be in the diagonal position. If it occurs, it means that the leak index equal to the row in question can be located with the selected sensor configuration. Otherwise, the leak cannot be located using this configuration.

VI. REAL CASE STUDY: LIMASSOL NETWORK

The proposed methodology involving GA is applied to a real network. In this paper, the Limassol network in Cyprus is used. This network has a total of 197 demand nodes and is represented in Figure 1. The network model is available in EPANET, as in the case of the Hanoi network. First, results obtained when placing three sensors are achieved by using the semi-exhaustive algorithm. This algorithm is time-demanding in this case since there are more than 1.2×10^6 possible combinations of nodes to be measured. The sensor placement problem is set-up with an $EC = 0.25$ (leak of approximately 1.67 lps) for the sensitivities and an $EC = 0.20$ (leak magnitude= 1.3 lps approximately) for the residuals. The best configuration obtained leads to place sensors in nodes {82, 133, 157} with an $\bar{\varepsilon} = 0.258$. This triplet will serve as reference to evaluate the performance of the GA approach. Then, we apply Algorithm 2 based on GA in order to find the adequate set of sensor configurations for various types of residual and sensitivity matrices when varying the EC according to leak magnitudes within a given range.

Algorithm 2 Sensor placement based on Genetic Algorithms

Require: A sensitivity matrix S, a residual matrix R. The number of sensors n, the number of nodes m and the maximum number of iterations d.

Ensure: A near-optimal sensors configuration $\mathbf{q}_{min}$ with error index $\bar{\varepsilon}_{min}$.

$init \leftarrow InitVarGA()$
$restrict \leftarrow SetRestrictions(\sum_{i=1}^{m} \mathbf{q}_i = n)$
$z \leftarrow ChooseSeed()$
for $k = 1, \cdots, d$ **do**
  Create I^k matrix of size $((z+1) \times m)$ where each row is a random initialization such that:

$$I^k_{ij|i \neq (z+1)} \leftarrow \begin{cases} 1 & \text{if row } i \text{ is with a sensor in node } j \\ 0 & \text{otherwise} \end{cases}$$

$$I^k_{(z+1)j} \leftarrow \begin{cases} \{\} & \text{if } k = 1 \\ q^{k-1} & \text{otherwise} \end{cases}$$

  GA based search:
  Inputs: $init$, $restrict$, R, S, I^k.
  while An optimization criterion is not reached **do**
    $\mathbf{q} \leftarrow getConfig()$
    $\hat{S}(\mathbf{q}) \leftarrow eval\_S(\mathbf{q}, S(\mathbf{q}))$
    $\hat{R}(\mathbf{q}) \leftarrow eval\_R(\mathbf{q}, R(\mathbf{q}))$
    $\Psi(\mathbf{q}) \leftarrow eval\_\Psi(\hat{R}(\mathbf{q}), \hat{S}(\mathbf{q}))$
    $\varepsilon(\mathbf{q}) \leftarrow eval\_\varepsilon(\Psi(\mathbf{q}))$ // c.f. (13)
    $\bar{\varepsilon}(\mathbf{q}) \leftarrow \underset{i}{mean}(\varepsilon_i(\mathbf{q}))$ //c.f. (14)
  end while
  Find $\{\mathbf{q}^k, \bar{\varepsilon}^k\}$ such that $\bar{\varepsilon}^k = \min_{\mathbf{q}}(\bar{\varepsilon}(\mathbf{q}))$
end for
Find $\{\mathbf{q}_{min}, \bar{\varepsilon}_{min}\}$ such that $\bar{\varepsilon} = \min_{k}(\bar{\varepsilon}^k)$

TABLE I

CORRESPONDING ERROR INDEX IN LIMASSOL NETWORK FOR CONFIGURATION IN II

	Residuals EC				
Sensitivities EC	0.15	0.2	0.25	0.3	0.35
0.15		0.324	0.294	0.299	0.314
0.2	0.299		0.284	0.279	0.294
0.25	0.279	0.274		0.243	0.243
0.3	0.309	0.279	0.263		0.258
0.35	0.324	0.279	0.263	0.258	

The decision of how to choose the location of the sensors in the network is based on the Algorithm 2. In order to look for the best configuration according to sensitivities and residuals, we perform the search using every possible combination of sensitivites/residuals considering different leak magnitude sizes (*EC* values). Table I shows the computed error index for each case that was found using the genetic algorithm.

Table II presents the nodes where the sensors should be placed in order to obtain the minimum isolation error computed via genetic algorithms according to the value selected in the computation of the sensitivities and the magnitude of the leak tested (see Table I). Just for validation, if the same analyses were repeated with the semi-exhaustive search, the same error and configurations as the GA case would be found but with a higher computational cost.

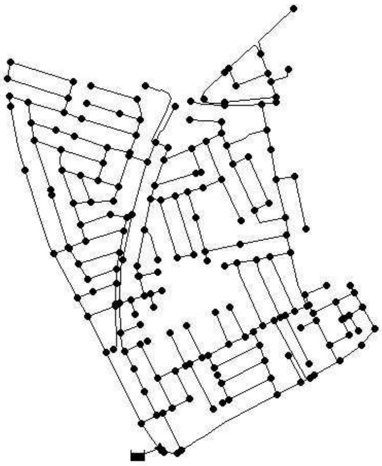

Fig. 1. Water network in Limassol, Cyprus

From previous tables, it can be noticed (as in the case of the Hanoi network) that depending of the emitter coefficient (leak magnitude) selected for the sensitivities and for the residuals, the algorithm finds different configurations. In order to find the best configuration, the following tests are performed:

- *Variation in the tested leak magnitude*: We compute the projection matrix for all the found configurations taking into account every combination of the sensitivities computed with emitter coefficient values of 0.15, 0.2, 0.25, 0.3 and 0.35 and the same number of computation of the residuals.
- *Consideration of a limited sensor precision*: To take into account the limitation of the sensor precision, we truncate the two last decimals of the pressure measurements to compute the residual matrices.
- *Application of random noise in the measurements*: The third test is the application of random noise in the measurements around 0.5% of the expected measurements.

In order to select the adequate configuration of sensors, we propose to perform the experiments described above and look for the combination with the smallest average error index for all the possible leak magnitudes and sensitivities to test. This criterion is analytically established by taking the minimum of the average error indices

$$\min(\frac{1}{L^2}\sum_{j=1}^{L}\sum_{i=1}^{L}\varepsilon^{ij}) \tag{16}$$

where L is the number of leak magnitudes used and ε^{ij} is the error index (c.f. (13) and (14)) obtained with residual and sensitivity with respective indices i and j. In such a way, the search of the best sensor placement is built as a *min-max* optimization problem.

The Table III shows the averages obtained for each configuration when the experiments were performed as well as the

TABLE II

SENSOR PLACEMENT IN LIMASSOL NETWORK INSTALLING 3 SENSORS

	Residuals EC				
Sensitivities EC	0.15	0.2	0.25	0.3	0.35
0.15		40 77 172	25 77 133	76 133 185	76 133 152
0.2	76 133 152		76 86 152	77 124 152	76 110 173
0.25	85 156 196	8 76 150		75 116 157	72 115 150
0.3	72 118 163	76 133 141	77 111 150		75 23 152
0.35	76 128 140	75 120 150	77 115 137	29 112 152	

TABLE III

NORMALIZED ERROR INDEX AVERAGES IN EXPERIMENTS

		Test			
Configuration		**Magnitude**	**Resolution**	**Noise**	**Average**
1	75 116 157	0.3363	0.4124	0.5759	0.4415
2	85 156 196	0.3617	0.4546	0.5967	0.4710
3	72 115 150	0.3449	0.4289	0.5830	0.4523
4	76 110 173	0.3404	0.4089	0.5561	0.4351
5	77 124 152	0.3480	0.4442	0.5807	0.4576
6	76 133 152	0.3183	0.4028	0.5581	0.4264
7	76 86 152	0.3348	0.4208	0.5723	0.4426
8	25 77 133	0.3358	0.4195	0.5688	0.4414
9	76 133 185	0.3338	0.4203	0.5640	0.4393
10	40 77 172	0.3675	0.4475	0.6162	0.4771
11	76 133 152	0.3183	0.4028	0.5462	0.4224
12	8 76 150	0.3558	0.4617	0.6127	0.4767
13	72 118 163	0.3731	0.4429	0.5759	0.4640
14	76 133 141	0.3411	0.4157	0.5703	0.4424
15	76 128 140	0.3553	0.4246	0.5731	0.4510
16	75 120 150	0.3284	0.4310	0.5817	0.4470
17	77 111 150	0.3419	0.4223	0.5660	0.4434
18	77 115 137	0.3299	0.4168	0.5612	0.4360
19	75 123 152	0.3388	0.4206	0.5749	0.4448
20	29 112 152	0.3939	0.4546	0.5896	0.4794

total average for each solution. From this table, the optimal sensor placement configuration is the one with minimum total average error index.

As a conclusion, although the performance decreases in presence of uncertainties related to leak magnitude, a near optimal solution for the sensor placement can be found placing the sensors in nodes 76, 133 and 152 in order to maximize the leak isolability criteria.

VII. CONCLUSIONS

In this paper a new approach to sensor placement for water distribution networks, that maximizes leak isolability, has been proposed. This approach is combined with a projection based isolation scheme. Nevertheless, it could be easily adapted to any other isolation scheme.

The sensor placement problem has been formulated as an integer optimization problem. The optimization criterion is based on minimizing the number of non-isolable leaks according to the isolability criteria introduced. Because of the non-linear integer and large-scale nature of the resulting optimization problem, GA have been proposed. To validate the results obtained, they have been compared with those produced by an exhaustive search method with higher computational cost proving that GA allow to find near-optimal solutions with less computational load. The effect on the unknown leak size and its effect in the sensor placement algorithm have also been studied.

As future work, we propose to perform the analysis based on the behavior of the network along a time horizon taking into account the demand pattern to improve the selection of the best place to locate the sensors. We also want to design a robust algorithm in which all the uncertainties are evaluated inside the optimization function in order to obtain a sensor placement that considers the effect of noise, sensor resolution and different leak magnitudes.

REFERENCES

[1] P. L. Andrew F. Colombo and B. W. Karney, "A selective literature review of transient-based leak detection methods," *Journal of Hydro-environment Research*, pp. 212–227, 2009.

[2] R. S. Pudar and J. A. Liggett, "Leaks in pipe networks," *Journal of Hydraulic Engineering*, vol. 118, no. 7, pp. 1031–1046, 1992.

[3] 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. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0967066111001201

[4] M. V. Casillas, L. Garza-Castañón, 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*. Elsevier, 2012, pp. 570–575.

[5] A. Yassine, S. Ploix, and J.-M. Flaus, "A method for sensor placement taking into account diagnosability criteria," *Int. J. Appl. Math. Comput. Sci.*, vol. 18, no. 4, pp. 497–512, Dec. 2008. [Online]. Available: http://dx.doi.org/10.2478/v10006-008-0044-5

[6] A. Rosich, R. Sarrate, and F. Nejjari, "Optimal sensor placement for fdi using binary integer linear programming," in *20th International Workshop on Principles of Diagnosis, DX09*, 2009, pp. 235–242.

[7] A. Krause, J. Leskovec, C. Guestrin, J. Vanbriesen, and C. Faloutsos, "Efficient sensor placement optimization for securing large water distribution networks," *Journal of Water Resources Planning and Management*, 2008.

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

[9] L. Rokach and O. Maimon, "Clustering methods," in *Data Mining and Knowledge Discovery Handbook*, 2005, pp. 321–352.

[10] J. R. Koza, "Survey of genetic algorithms and genetic programming," in *In In Proceedings of the Wescon 95 - Conference Record: Microelectronics, Communications Technology, Producing Quality Products, Mobile and Portable Power, Emerging Technologies*. IEEE Press, 1995, pp. 589–594.

[11] D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA, Addison-Wesley, 1989.