Leakage Isolation using Pressure Sensitivity Analysis in Water Distribution Networks: Application to the Barcelona case study

Ramon Pérez*, Vicenç Puig**, Josep Pascual*, Joseba Quevedo*, Edson Landerosº, Antonio Peralta⁺

*Advanced Control Systems Group Universitat Politècnica de Catalunya (UPC) **Institut de Robòtica i Informàtica Industrial (CSIC-UPC) °CETAQUA Water Technological Center +AGBAR Barcelona Water Company

Abstract: Leaks are present to some extent in all water-distribution systems. This paper proposes a leakage localisation method based on the pressure measurements and pressure sensitivity analysis of nodes in a network. The sensitivity analysis using analytical tools is not a trivial job in a real network because the huge non-explicit non-línear systems of equation that describe its dynamics. Simulations of the network in presence and absence of leakage may provide an approximation of this sensitivity. This matrix is binarised using a threshold independent of the node. The binary matrix is assumed as a signature matrix for leakages. However, there is a trade-off between the resolution of the leakage isolation procedure and the number of available pressure sensors. In order to maximise the isolability with a reasonable number of sensors, an optimal sensor placement methodology, based on genetic algorithms, is also proposed. This methodology has been developed for Barcelona Network using Piccolo simulator. The sensor placement and the leakage detection and localization methodologies are applied to district management areas (DMA).

Abstract: Pressure Sensitivity, Leakage Localisation, Sensor Placement.

1. INTRODUCTION

Leakage monitoring may be done on a routine basis or when major losses are suspected between night and day water demands (Lambert, 1994). Methods for locating leaks range from ground-penetrating radar to acoustic listening devices (Farley, 2003). Some of these techniques require isolating and shut down part of the system. Techniques based on locating leaks from pressure monitoring devices allow a more effective and less costly search in situ. This paper proposes a leakage localisation method based on the pressure measurements and pressure sensitivity analysis of nodes in a network when a leak is present in a node (Pudar, 1992).

The analytical calculation of sensitivity is not a trivial job in a real network because the huge non-explicit non-línear systems of equation that describe its dynamics. Simulation of the network in presence and absence of leakage provides an approximation of this sensitivity. The approximation is used to generate a sensitivity matrix that is binarised using a threshold independent of the node. The binary matrix produced is assumed as a signature matrix for leakages. Model based fault detection and isolation techniques described in *Section 2* are used for the leakage detection and

location. The sensor placement is a critical issue for maximising discriminability in the localisation of leakages. The optimal sensor placement is obtained using a cost based on the size (maximal) of the sets of possible nodes containing a leakage. This cost is minimised using genetic algorithms in a process described in *Section 4*.

The signature matrix is generated for the set of sensors selected. This matrix has to be compared with the signature obtained comparing the model and the real measurements. From this comparison, the leakage should be located in a set of possible nodes. This methodology is presented in *Section 5*. The objective of this work is to explore the possibilities of this technique for leakages due to burst but not especially huge. Leakages about one litre per second may not be evident and stay for long periods of time representing an important water loss counted in volume.

In the analysis of water-distribution systems, demands typically are assumed to occur in the nodes. In this paper, it will also be assumed that leaks occur at the nodes. Then, under such assumption, leaks can be seen as additional demands but with unknown location and quantity. This methodology has been developed within PROFURED project. This project has been lead by CETAQUA, water technological center. Data were provided by the Water Company of Barcelona. Simulations based on PICCOLO are used in this network and it has been used in this work.

In Figure 1, the case study DMA used for illustrating the methodology is presented. Its model contains 1600 nodes and 41.153m of pipes. Simulated leaks introduced in the network are of 1 l/s more or less 3% of the total demand of the sector (in the night time). The demand distribution all over the network is the most variable parameter of the model. Some uncertainty in the demand has also been included in order to test the robustness of the method.



Fig. 1: Case study network: Plaça del Diamant

2. LEAKAGE DETECTION AND ISOLATION

The methodology of leakage localisation used in this work is mainly based on standard theory of model-based diagnosis described for example in (Gertler, 1998). Model based diagnosis can be divided in two subtasks: fault detection and fault isolation. The principle of model-based fault detection is to check the consistency of observed behaviour while fault isolation tries to isolate the component that is in fault. The consistency check is based on computing residuals, $\mathbf{r}(k)$, obtained from measured input signals $\mathbf{u}(k)$ and outputs $\mathbf{y}(k)$ using the sensors installed in the monitored system and the analytical relationship which are obtained by system modelling:

$$\boldsymbol{r}(k) = \boldsymbol{\Psi}(\boldsymbol{y}(k), \boldsymbol{u}(k)) \tag{1}$$

where Ψ is the residuals generator function that depends on the type of detection strategy used (parity equation (Gertler, 1998) or observer (Chen and Patton, 1999)). At each time instance, k, the residual is compared with a threshold value (zero in ideal case or almost zero in real case). The threshold value is typically determined using statistical or set-based methods that take into account the effect of noise and model uncertainty (Blanke, 2006). When a residual is bigger than the threshold, it is determined that there is a fault in the system; otherwise, it is considered that the system is working properly. In practice, because of input and output noise, nuisance inputs and modelling errors affecting to the considered model, robust residuals generators must be used. The robustness of a fault detection system means that it must be only sensitive to faults, even in the presence of modelreality differences (Chen and Patton, 1999).

Robustness can be achieved at residual generation (active) or evaluation phase (passive). Most of the passive robust residual evaluation methods are based on an adaptive threshold changing in time according to the plant input signal and taking into account model uncertainty either in the time or frequency domain. In this paper, a passive method in time domain has been proposed for robust fault detection, where the detection threshold has been obtained using the method described in Section 3. Robust residual evaluation allows obtaining set of fault а signatures $\boldsymbol{\phi}(k) = \left[\phi_1(k), \phi_2(k), \dots, \phi_{n_{\phi}}(k) \right],$ where each indicator of fault is obtained as follows:

$$\phi_i(k) = \begin{cases} 0 \text{ if } |\mathbf{r}_i(k)| \le \tau_i(k) \\ 1 \text{ if } |\mathbf{r}_i(k)| > \tau_i(k) \end{cases}$$
(2)

where τ_i is the threshold associated to the residual $r_i(k)$.

Fault isolation involves identifying the faults affecting the system. It is carried out on the basis of fault signatures, ϕ , (generated by the detection module) and its relation with all the considered faults, $f(k) = \{f(k)_1, f_2(k), \dots, f_{n_f}(k)\}$. The method most often applied is a relation defined on the Cartesian product of the sets of faults $FSM \subset \phi \times f$, where *FSM* is the *theoretical signatures matrix* (Gertler, 1998). One element of that matrix FSM_{ij} will be equal to one, if a fault $f_j(k)$ is affected by the residual $r_i(k)$, in this case the value of the fault indicator $\phi_i(k)$ must be equal to one when the fault appears in the monitored system. Otherwise, the element FSM_{ij} will be zero.

3. SENSITIVITY MATRIX

The effect of a leakage on the pressure in a node is evaluated by the sensitivity matrix (Pudar, 1992):

$$S = \begin{pmatrix} \frac{\partial p_1}{\partial f_1} & \cdots & \frac{\partial p_1}{\partial f_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial p_n}{\partial f_1} & \cdots & \frac{\partial p_n}{\partial f_n} \end{pmatrix}$$
(3)

It is extremely difficult to calculate S analytically in a a real network. Implicit equation and a too huge number of

variables are involved. The sensitivity matrix depends on the working point (that is, the demand and boundary conditions).

In this work, the sensitivity matrix has been generated by simulation using increments of pressure and maintaining the leakage flow constant. It has been verified that the analytical and the simulated sensitivity converge for small leakages.

In Figure 2, the sensitivity matrix for the case study network of Figure 1 is shown graphically. It has been ploted for 15 nodes distributed homegenously in the DMA as illustration.



Fig. 2: Sensitivity matrix

Some sensors are much more sensitive to all leakages than others. Thus, a normalisation of sensitivity is needed. Each row is divided by the maximum value (that corresponds to the leakage most important for that node, on the diagonal). This *Normalised Sensitivity Matrix* is shown in Figure 3 for the considered example.

In leakage localisation, each element FSM_{ij} is equal to zero when leakage *j* does not affect pressure in node *i* and it is equal to 1 when leakage *j* affects node *i*. The aim is to generate the signature matrix from the Normalised Sensitivity Matrix. In Figure 3, it can be seen that all leakages affect all pressures.

A process inspired in the *ε-method* proposed by Sezer and Siljak (1986) is proposed with the aim to identify the strongest relations between leaks and pressure measurements. In this process it is absolutely essential to choose conveniently the threshold that controls if a leak has or not an effect on a given pressure. The process proceeds as follows: those leaks that have an effect less than the given threshold are considered as a '0' in the leak signature matrix (2). Otherwise, their effect is considered as a '1'. In this way, the sensitivity matrix is binarised based on selected threshold. Normalisation allows using a unique threshold for all sensors but the choice of the threshold is most relevant in the process. For small thresholds, all binarised matrix elements are 1 and

only detection is possible. As the threshold increases more 0's appear. When threshold approaches 1 then only diagonal is 1 and localisation is perfect (or almost perfect, simulation precision makes some nodes equally sensible to some leakages) that but all sensors are needed. Figure 4 shows the evolution of the number of signatures present in the matrix and the maximum number of leakages with the same signature. It corresponds to the 1613 nodes of the network in Figure 6. Theoretically with 11 sensors (rows) there may be 2047 (2^{11} -1; signature with all 0 is discarded as detection is imposed) different signatures for leakages (columns). In order to get maximum number of signatures a necessary condition is to have in each row 2^{n-1} 1's, where *n* is the number of sensors (rows).



Fig. 3: Normalised sensitivity matrix



Fig. 4: Evolution of the signature matrix depending on threshold

This necessary condition is fulfilled for the threshold where both lines in Figure 4 cross (~ 0.1). This is the threshold used.

4. SENSOR PLACEMENT

An optimal sensor placement is defined as a sensor configuration that achieves the minimum economical cost (number of sensors) while observing pre-specified performance criteria (groups of nodes that are not isolable with a minimum number of elements).

A model of water network can be represented as a graph G = (V, E), where E is the set of edges that represent the pipes and V is the set of vertices (nodes) where pipes meet. Vertices can represent sources, such as reservoirs or tanks, where water is introduced or sinks (demand points) where water is consumed. Each pipe connects two vertices v_i and v_j and usually is denoted as (v_i, v_j) .

Using the graph representation, the problem of optimal sensor placement can be formulated as an integer programming problem, where each decision variable X_i associated to a node v_i of the network can be 1 or 0, meaning respectively that the sensor will be or not installed in this node (Bagajewicz, 2000). The starting point of the algorithm is the leakage sensitivity matrix obtained by simulation. Each element of this matrix contains the sensitivity that presents a node (in rows) to the considered possible leakages (in columns), assuming a single leakage at time.

The objective function is the number of elements for the largest set of leaks with the same signature. In order to increase isolability this cost should be minimized but at the same time keeping the economical cost reasonable, that is installing the less number of sensor that is possible. The problem is solved for a number of sensors; this number is increased till the cost does not decrease. A constraint is included such that all leaks should be detected. It is introduced by forcing that signature with all 0's is not accepted. Thus the cost of non-detection has not to be into account.

This optimization problem can be solved using either deterministic method based for example in Branch and Bound or heuristic methods based for example in Genetic Algorithms. The first type of methods guarantee the optimal solution but the computation time tends to be exponential with the number of nodes/faults (Sarrate *et al.*, 2007). On the other hand, the second type of methods just guarantees a suboptimal solution that tends to the optimal one when the size of considered population tends to infinity. Besides the formulation of solutions in series of 1's and 0's are most convenient for a GA.

In Figure 5, the evolution of cost function is presented. The cost has been taken as the number of nodes in the biggest group of possible leakage isolated with a number of sensors and a threshold between 0.8 and 0.9. It can be seen that the best choice is threshold 0.8 and that 8 sensors are as good as 11. Therefore only 8 sensors are used.



Fig. 5: Evolution of the cost function depending on number of sensors and threshold

In Figure 6, the different groups of nodes with the same leakage signature are shown. There are 39 groups and the hugest contains 190 nodes. The localization of the sensors after the optimization process is in the last figure.



Fig. 6: Groups of nodes with the same leakage signature with 8 sensors and placement of sensors

In an ideal situation with a well calibrated network model a leakage should be searched in one of these regions instead of the whole sector.

5. LEAKAGE LOCALISATION

The localisation of leakages is based on the isolation techniques presented in *Section 2*. The binarised sensitivity matrix is used as a signature matrix for all leakages. In *Section 3*, the signatures were used to obtain an optimal distribution of the sensors and in Figure 6 the groups of nodes that generate the same signature are presented.

In the process of leak localization, the signature generated in case of having 8 sensors installed with a fault is compared

with signature matrix. If the model were perfect, the leak should be localized with one measurement. Assuming that some parameters may be uncertain, the test has been done during 15 days of simulation (only the lowest consume hour is used each day) and then three options are used to assign the observed leakage signature to a group:

- Mean of the sensitivities
- Mean of binarised sensitivities
- The voting (all days the leak is assigned to a group and that with more assignation is the elected)

Results without uncertainty were not good using any of the three decision criteria. It was due to the changing boundary conditions that affected very much the sensitivity matrix. It is necessary to generate it each day when boundary conditions are known. When a new signature matrix for each day is generated the two first approaches are useless and the third one is tested. It provided perfect results without uncertainty, 100% localisation. It means that each day the group that was signalled suitable to have a leakage contained the node with leakage. Of course these groups were all different. Signature matrix is adapted to boundary conditions. Thus, there are different probabilities of having a leak in a node. This appears on Table 1. It shows the number of nodes with that have been signalled 0-15 times (each for each day). The shadowed line cell corresponds to the one that contains the node with leakage. In this case, it corresponds always to the 15. It has been done for the 39 groups (one leakage for each) that appeared in the first day.

The number in each cell corresponds to the nodes that have been signalled n times in the 15 days, being n the number of the row. Columns correspond to the 39 cases tested. In Figure 7, the nodes are presented in grey scale representing the times that have been signalled to be suitable of containing a leakage. The one that contained it appears in the black area.



Fig. 7: Localisation of a leak in the correct zone with adapting signature matrix

Table 1: Results using voting criteria adapting signature matrix

		Leak												
		1	2	3	4	5	6	7	. 8	9	10	11	12	13
Number of detections	0	529	316	503	316	986	798	782	884	1245	489	1253	1343	363
	1	629	505	639	519	438	615	491	518	325	761	131	64	778
	2	175	559	311	545	48	147	95	52	64	279	15	9	345
	3	126	88	12	88	18	37	76	5	3	9	29	12	59
	4	32	57	11	51	17	15	32	22	0	17	1	6	22
	5	19	39	21	37	30	0	63	12	0	15	0	7	33
	6	54	35	9	11	11	15	30	11	0	9	16	2	13
	7	31	8	5	- 39	7	0	14	5	0	4	117	6	10
	8	9	2	15	3	3	0	13	9	0	5	2	9	2
	9	1	1	31	0	12	0	2	6	1	5	11	94	4
	10	10	2	56	1	18	0	6	17	0	1	5	46	9
	11	3	1	15	2	10	2	7	55	0	26	11	7	0
	12	1	6	8	1	6	0	6	33	1	0	2	10	0
	13	9	18	2	9	0	1	15	10	0	5	4	2	0
	14	5	0	1	0	29	6	6	0	0	6	8	20	0
	15	7	3	1	18	7	4	2	1	1	9	35	3	2
								Lea	k					
		14	15	16	17	18	19	20	21	22	23	24	25	26
Number of detections	0	357	1311	1363	354	489	539	477	490	1396	1430	996	959	807
	1	869	66	44	853	552	575	778	512	122	92	372	347	533
	2	304	21	13	317	156	179	255	182	12	11	82	132	134
	3	50	83	5	57	170	59	65	47	55	52	100	90	53
	4	14	18	6	14	10	70	4	141	3	5	39	15	50
	5	12	119	9	11	7	53	5	4	7	8	2	23	1
	6	9	5	5	11	5	95	12	25	3	7	4	8	7
	7	9	2	35	7	25	18	33	31	4	1	2	19	1
	8	1	0	11	1	29	12	0	17	3	11	5	18	23
	9	0	0	23	0	53	9	2	25	11	3	1	11	9
	10	2	4	0	3	69	4	1	21	2	0	6	14	14
	11	4	8	9	4	11	8	4	46	2	5	15	1	0
	12	6	1	0	2	36	3	1	35	0	0	2	1	0
	13	2	1	17	4	9	8	0	23	9	11	8	1	6
	14	0	0	44	1	15	1	1	24	0	0	2	0	0
	15	1	1	56	1	4	1	2	17	11	4	4	1	2
		27	20	20	20	21	22	Lea	K 24	25	26	27	20	20
	0	2/ 506	28	29	504	507	502	544	502	400	400	502	38 977	569
	1	390	709 595	202	394	397	393 502	344	393	490	490	393	257	508
	2	362	202	293 90	200	70	100	246	100	182	163	116	357	158
1	2	100	200	90 07	209	115	76	240	76	102	105	110 Q1	45	1.38
1	3	56	45 1	77	120	121	/0	12/	/0	4/	50	01 /0	34	40
ons	4	0.	2	3/	120	121	42	1/0	42	140	12	40	54	30
umber of detecti	6	9	2	21	32	43	44	20	40	25	43	34	11	50
	7	1	5	21	35	33	19	24 16	-40	35	70	13	4	35
	8	2	2	8	33	11	51	2	43	18	16	27	0	18
	9	6		10	14		9	2	16	22	101	23	2	.0
	10	1	1	4	23	13	6	5	6	38	9	9	40	14
ź	11	6	5	2	5	2	4	23	3	25	79	27	139	8
	12	1	7	2	16	6	11	6	23	30	3	7	13	30
	13	1	0	5	8	0	40	4	22	38	2	27	1	10
	14	8	0	0	1	1	12	6	21	7	1	5	6	0
	15	5	1	1	5	1	15	2	9	26	1	11	1	1

Uncertainty was calculated using the monthly variation for a demand. It was of 18% of the total demand. Uncertainty was introduced as a coefficient multiplied to the demand of each node generated as a random number between 0.8 and 1.2. The global demand has been kept equal. The total demand affects greatly to the sensitivity.

Results are presented in Table 2 and Figure 8. In this case the leaky node is not always exactly in the most signalled group and the dark grey in the figure does not correspond to 15 but to 13 days. In Figure 8 the gray scale is lighter because there are less correct detections due to the uncertainty.





Increasing uncertainty interval, the proposed localization methodology produces poorer results. For a 50% uncertainty leaks were not well localized but they were localized in a neighbour zone.



Fig. 8: Localisation of a leak in the correct zone

6. CONCLUSIONS

A leakage localization method based on the pressure measurements and sensitivity analysis of nodes in a network has been proposed. The leakage localization methodology is founded in standard model based fault diagnosis well established theory.

In order to maximise the isolability with a reasonable number of sensors an optimal sensor placement methodology based on genetic algorithms is also proposed. The objective function in the minimisation process was the size of the maximum group discriminated.

To assess the validity of the proposed approach, it has been applied to a real DMA of Barcelona network. Models and information were provided by the water company. For this sector (DMA) the sensor placement and the leakage detection and localization methodologies have been applied with successful results even in presence of demand uncertainty.

An issue in the process is to recalculate the sensitivity matrix for each boundary condition using the simulation model because of the high dependence of it to global consumption. This approach is being currently developed using linear parameter varying (LPV) models that consider the consumption as a scheduling variable (Vento, 2009).

ACKNOWLEDGMENT

The authors wish to thank the support received by WATMAN ref. DPI2009-13744 of the Spanish Ministry of Education.

REFERENCES

- M. Bagajewicz (2000), *Design and Upgrade of Process Plant Instrumentation*. Lancaster, PA: Technomic Publishers.
- M. Blanke, M. Kinnaert, J. Lunze, M. Staroswiecki (2006), *Diagnosis and Fau lt-tolerant Con trol*, 2° Edition, Springer.
- J. Chen, R. J. Patton (1999), *Robust Model -Based F ault Diagnosis for Dynamic Sy stems*. Kluwer Academic Publishers.
- M. Farley, S. Trow (2003). *Losses in W ater Distribution Networks*. IWA Publishing UK.
- J. J. Gertler (1998), *Fault Det ection a nd Diagnosis i n Engineering Systems*, Marcel Dekker.
- R. S. Pudar (1992) J.A. Ligget Leaks in Pi pe N etworks. Journal of Hydraulic Engineering. Vol. 118, No. 7, July 1992, pp. 1031-1046.
- A. Lambert (1994), Accounting for lo sses: the Bursa and background c oncept. (BABE)IWEM Journal, April 1994, 8(2), 205-14.
- R. Sarrate, V. Puig, T. Escobet, A. Rosich (2007). *Optimal* Sensor Placement for Model-based Fault Detection and Isolation. 46th IEEE Conference on Decision and Control. New Orleans, USA.
- M. E. Sezer and D.D. Siljak. *Nested epsilon-decomposition* and cl ustering of complex systems. Automatica, 22(3):321–331, 1986.
- J. Vento, V. Puig (2009) Leak Detection and Is olation in Pressurized Water Networks Usi ng Interval LPV Models. 7th Workshop on Advanced Control and Diagnosis (ACD'09). Zielona-Gora, Poland.