Modeling the effects of leaks on measured parameters in a water distribution system

Alexandru Predescu, Mariana Mocanu, Ciprian Lupu
Faculty of Automatic Control and Computers, University POLITEHNICA of Bucharest
mircea.predescu@stud.acs.upb.ro, mariana.mocanu@cs.pub.ro, ciprian.lupu@acse.pub.ro

***Abstract*—This paper describes a method for leak detection in a water distribution system based on detecting discrepancies between measured data and estimations obtained from the simulation of a water distribution network model. Using this method, the effect of a leak on measured data can be estimated at each node and the deviation from nominal values can be used to find the most probable location of the leak in the network. The main aspects which are described in this paper consist of data modeling for the water distribution system according to the INSPIRE directive for spatial data specification, and the algorithm for model simulation and leak detection.**

***Index Terms*—fault detection, model identification, leak sensitivity analysis**

Abbreviations

SCADA Supervisory control and data acquisition

I. Introduction

The water distribution system in present infrastructure requires regular maintenance and repairs, which lead to an increase in operational cost.

In the context of recent developments in IoT and data services, estimating the state of the network has become feasible by installing devices for monitoring real-time parameters in a distributed system as in the case of SCADA systems. Therefore, extensive research has been done on the theoretical and practical approaches for detection and prevention of leaks and problems in the water network.

There are many approaches which propose various detection methods:

- Hardware based methods: using acoustic sensors, gas detectors, negative pressure detectors, and/or infrared thermography
- Software based methods: flow/pressure change detection and mass/volume balance

By using pressure sensors installed in network nodes, the steady state of the system is analyzed and pressure discrepancies above a certain threshold can signal a leak along the pipe. A more advanced approach takes into consideration the low pass filter characteristics of a long pipeline and pressure disturbances which can signal a leak in the case of a high rate of change.

By using flow sensors, the flow rate of change is compared to a statistical model for flow changes within a specific time period.

In model based systems the steady state pipe flow is described mathematically using laws for conservation of mass and energy. Leaks are detected when there are discrepancies between calculated and measured values. [1]

In this paper, we propose a method for leak detection in a water distribution system. The informational background and requirements are described for a future implementation according to the INSPIRE directive which provides specifications for spatial data and facilitates public access to spatial information across Europe. [2] [3]

In the second section we presented some results in the area of water distribution systems and some observations regarding the possible approaches for accurate measurements.

In the third section we describe the informational model of the system and the mathematical model and algorithm used for simulating the model parameters. Then we define a measure of the effects of leaks on measured parameters which is used in the Results section for comparing the two possibilities for measurements and leak detection (using flow sensors or pressure sensors).

II. Related work

Related work in the area of water distribution systems include leakage detection methods based on pressure measurements and their estimates and optimal sensor placement as described in [4]. A method which is designed to reduce cost of the system by optimizing the number of installed sensors is presented in [5], [6]. It is based on a fault sensitivity matrix which emphasizes the discrepancies in the hydraulic model generated by an arbitrary location of a leak in the network.

A similar approach is described in [7], where the effect of leaks on the pressure in each node is defined. The use of pressure sensors is considered to be more practical in terms of costs, as the installation of flow sensors can be more expensive. [8]

On the other hand, accurate pressure measurements are required as discrepancies in the case of a leak can be negligible in the case of large networks. From this point of view, flow measurements can be a more accurate indicator for possible leaks.

III. Our approach for leak detection

We propose an approach which takes into consideration the requirements for compatibility with the standards for open data and the fundamental laws and means of representation in the domain of water distribution systems.

A. The informational model

Data services which are compatible with the INSPIRE directive, can be used to provide information about the network configuration and parameters. This can be implemented in ArcGIS using vector layers which describe the features (pipes, nodes, flow sensors, pressure sensors) and can be used in mapping applications. [9]

Public access is another specification which is implemented by using a relational database where the geographic data is stored as well as the measured, estimation and configuration data.

In this section, the data model is presented, which is created considering the requirements given by the mathematical model. The main task is the identification and localization of leaks in a water distribution system, with further possibilities of bypassing the affected section in the network. This is accomplished by collecting and processing data from pressure and flow rate sensors in the network, and comparing the data to a reference model based on network parameters and estimated data. The structure of the system is described in figure 1.

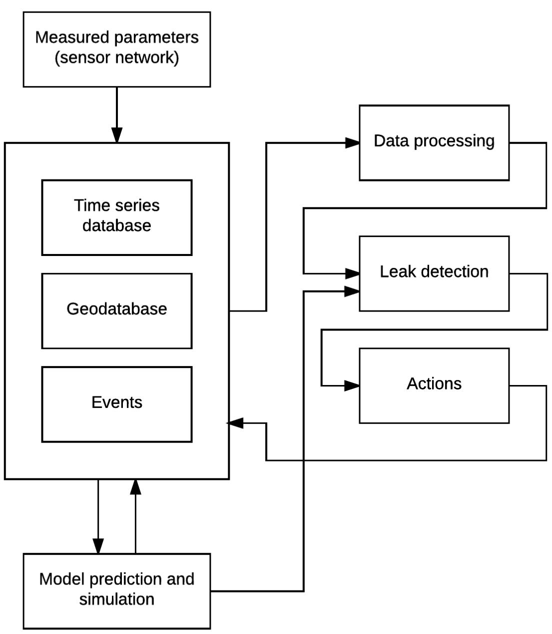

Fig. 1. Block diagram of the system

The following modules are used to define the conceptual data model:

- Water network: data about the infrastructure (pipes, nodes)
- Node types: defines the node type (producer, consumer) and properties
- Monitoring stations: location, type
- Sensor types: technical specification for installed sensors
- Measured and estimated data: data from sensors or estimated data as time series
- Processed data: processed data (estimated water loss)
- Events: information about detected events (leaks, actions)

The database which is used for measured and simulated data is shown in figure 2.

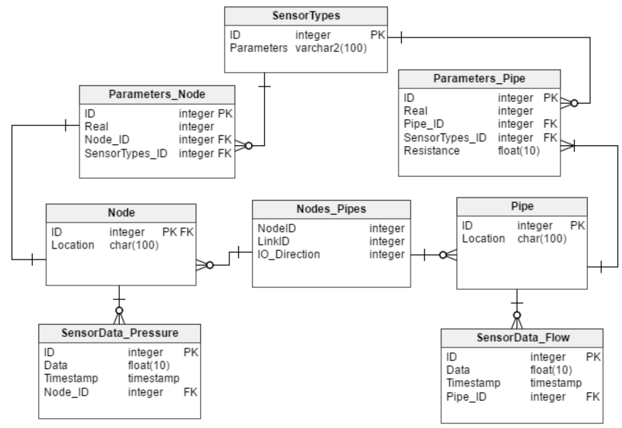

Fig. 2. Database design

The data model can be translated to the model described in the INSPIRE specification: Module - Feature dataset, Object - Feature class (Attribute tables), Attribute - Data dictionary attribute, Relation - Relation class, Additional information - Metadata

B. The mathematical model

The network model of the water distribution system can be defined as a graph (as in graph theory). The fundamental laws which define the process are used to simulate the state of the network based on the available data. In the case of water distribution systems, graph nodes define the junctions and graph edges define the pipes. Based on the mass and energy conservation laws, the state of the network can be estimated and compared to the actual data. This can be an indicator of the possible leaks in the network which translate into a disturbance in the measured parameters. The mass conservation law state that the sum of flows into a node is equal to the sum of flows out of the node. If there is a disturbance in the form of an additional demand (such as a leak), the law can be written as follows:

$$\sum q_{in} - \sum q_{out} = q_{ext} \tag{1}$$

The hydraulic model of the water distribution system can be defined in analogy to the electrical circuits and can be used to estimate the flow through the pipes, the head in each node and the water level in storage tanks. With this analogy, the model can be written in a similar way as the Kirchhoff laws and Ohm's law, for which numerical methods can be used to calculate the unknown variables.

The first law can be written for each node, which states that the sum of flows into a node is equal to the sum of flows out of the node:

$$\sum_j Q_{ij} - \sum_j D_{ij} = 0, \quad i = 1..n \tag{2}$$

We used the following notations: Q_{ij} - Input flow from node i to node j, D_{ij} - Output flow (demand, leak) from node i to node j. [7]

The equivalent for Ohm's law for smooth flow (laminar flow) of a fluid through a pipe is given by Poisseuille's law which can be written as follows:

$$Q = \frac{H_1 - H_2}{R} \tag{3}$$

$$R = \frac{8\eta l}{\pi r^2} \tag{4}$$

From (3) and (4), the headloss between two adjacent nodes can be calculated:

$$h = \frac{32\eta l}{\pi d^2} Q \tag{5}$$

We used the following notations for the model parameters and the equivalent electrical parameters: H - head (potential), h - headloss (voltage drop), R - resistance to flow (resistance), Q - flow (current), l - pipe length, d - pipe diameter, η - viscosity. [10]

In the proposed scenario, the pressure at the input node and the demand from each external node (flow to each leaf) are known, and the other parameters have to be estimated.

This is a practical scenario, as the demands and supply can result from a statistical model. Then, the indicator which can represent a leak is the deviation of measured head in one or more nodes from the value obtained in simulation.

The first problem is finding a solution which satisfies the equations for every node and edge of the network. A general solution is proposed which separates the unknown variables from the given parameters. The solution is defined for laminar flow conditions and it is based on numerical methods.

The equations are defined for each internal node (i.e. having at least one input edge and one output edge) and for each edge. The known and unknown variables are separated and the problem is defined in matrix form.

We define the following properties for nodes which are used in the algorithm:

- Id: Node index in the graph adjacency list
- H: The value for head
- Location: The coordinates of the node
- Known: Defines whether the node head is given (initial conditions) or has to be calculated

- UNId: Node index in the list of unknown nodes
- INId: Node index in the list of internal nodes

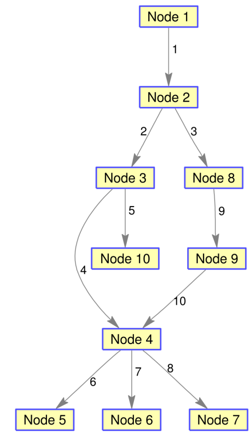

Fig. 3. Graph representation of the water network

- IE: Input edges
- OE: Output edges
- Type: Defines whether the node is real or simulated leak

We also define the following properties for edges which are used in the algorithm:

- Id: Edge index, can be automatically generated
- R: The resistance of the pipe
- Q: Flow through the pipe
- Known: Defines whether the edge value is given (initial conditions) or has to be calculated
- UEId: Edge index in the list of unknown edges
- IN: Input node
- ON: Output node

The following notation is used in the algorithm:

- neq: Number of equations in the model
- nvar: Number of variables in the model
- nun: Number of unknown nodes
- nue: Number of unknown edges
- nin: Number of internal nodes
- nre: Number of real edges (which are not used for simulating a leak)

The equations for nodes and edges are generated using the following general algorithm which can be also used for other scenarios:

1) Initialization
   - Find: nun, nue, nin, nre
   - neq = nin + nre
   - nvar = nue + nun
   - A = zeros(neq, nvar)
   - b = zeros(neq,1)
2) Generating the equations for each internal node (k)
   - for each input_edge (i)
     – if input_edge is known:
       b(k)=b(k)-Q(i)
     – if input_edge is unknown:
       A(k, input_edge.UEId) = 1

- for each output_edge (i)
  - if output_edge is known: b(k)=b(k)+Q(i)
  - if output_edge is unknown: A(k, output_edge.UEId) = -1

3) Generating the equations for each edge (k) defined by $H_i - H_j = Q_k R_k$ where H_i is the head at the input node and H_j is the head at the output node. The notation k(value) is used to define whether the value is known or has to be calculated.
   - $k(H_i) = 0, \quad k(H_j) = 0, \quad k(Q_k) = 0$
     - Equation: $H_i - H_j - Q_k R_k = 0$
     - Matrix form: $A(nin + k, nue + edge.IN.UEId) = 1$ $A(nin + k, nue + edge.ON.UEId) = -1$ $A(nin + k, edge.UEId) = -edge.R$
   - $k(H_i) = 0, \quad k(H_j) = 0, \quad k(Q_k) = 1$
     - Equation: $H_i - H_j = Q_k R_k$
     - Matrix form: $A(nin + k, nue + edge.IN.UEId) = 1$ $A(nin + k, nue + edge.ON.UEId) = -1$ $b(nin + k) = edge.Q * edge.R$
   - $k(H_i) = 0, \quad k(H_j) = 1, \quad k(Q_k) = 0$
     - Equation: $H_i - Q_k R_k = H_j$
     - Matrix form: $A(nin + k, nue + edge.IN.UEId) = 1$ $A(nin + k, edge.UEId) = -edge.R$ $b(nin + k) = edge.ON.H$
   - $k(H_i) = 0, \quad k(H_j) = 1, \quad k(Q_k) = 1$
     - Equation: $H_i = H_j + Q_k R_k$
     - Matrix form: $A(nin + k, nue + edge.IN.UEId) = 1$ $b(nin + k) = edge.ON.H + edge.Q * edge.R$
   - $k(H_i) = 1, \quad k(H_j) = 0, \quad k(Q_k) = 0$
     - Equation: $H_j + Q_k R_k = H_i$
     - Matrix form: $A(nin + k, nue + edge.ON.UEId) = 1$ $A(nin + k, edge.UEId) = edge.R$ $b(nin + k) = edge.IN.H$
   - $k(H_i) = 1, \quad k(H_j) = 0, \quad k(Q_k) = 1$
     - Equation: $H_j = H_i - Q_k R_k$
     - Matrix form: $A(nin + k, nue + edge.ON.UEId) = 1$ $b(nin + k) = edge.IN.H + edge.Q * edge.R$
   - $k(H_i) = 1, \quad k(H_j) = 1, \quad k(Q_k) = 0$
     - Equation: $Q_k R_k = H_i - H_j$
     - Matrix form: $A(nin + k, edge.UEId) = edge.R$ $b(nin + k) = edge.IN.H - edge.ON.H$
   - $k(H_i) = 1, \quad k(H_j) = 1, \quad k(Q_k) = 1$
     - Equation: $0 = -H_i + H_j + Q_k R_k$
     - Matrix form: $b(nin + k) = -edge.IN.H + edge.ON.H + edge.Q * edge.R$

The second problem is modeling the effects of leaks on measured parameters. We define the sensitivity matrix for pressure discrepancies (6) and flow discrepancies (7).

$$S_h = \begin{bmatrix} \frac{\partial h_1}{\partial q_{L1}} & \frac{\partial h_1}{\partial q_{L2}} & \cdots & \frac{\partial h_1}{\partial q_{Ln}} \\ \frac{\partial h_2}{\partial q_{L1}} & \frac{\partial h_2}{\partial q_{L2}} & \cdots & \frac{\partial h_2}{\partial q_{Ln}} \\ \vdots \\ \frac{\partial h_n}{\partial q_{L1}} & \frac{\partial h_n}{\partial q_{L2}} & \cdots & \frac{\partial h_n}{\partial q_{Ln}} \end{bmatrix} \tag{6}$$

$$S_q = \begin{bmatrix} \frac{\partial q_1}{\partial q_{L1}} & \frac{\partial q_1}{\partial q_{L2}} & \cdots & \frac{\partial q_1}{\partial q_{Ln}} \\ \frac{\partial q_2}{\partial q_{L1}} & \frac{\partial q_2}{\partial q_{L2}} & \cdots & \frac{\partial q_2}{\partial q_{Ln}} \\ \vdots \\ \frac{\partial q_n}{\partial q_{L1}} & \frac{\partial q_n}{\partial q_{L2}} & \cdots & \frac{\partial q_n}{\partial q_{Ln}} \end{bmatrix} \tag{7}$$

The partial derivatives can be interpreted as difference between the nominal state and the leak scenario. The sensitivity matrix is then normalized for each row (9) using the maximum value of pressure difference obtained from simulating each leak scenario (8).

$$\sigma_i = max\{s_{i1}, \cdots, s_{in}\} \tag{8}$$

$$\overline{S} = \begin{bmatrix} \frac{s_{11}}{\sigma_1} & \frac{s_{12}}{\sigma_1} & \cdots & \frac{s_{1n}}{\sigma_1} \\ \frac{s_{21}}{\sigma_2} & \frac{s_{22}}{\sigma_2} & \cdots & \frac{s_{2n}}{\sigma_2} \\ \vdots \\ \frac{s_{n1}}{\sigma_n} & \frac{s_{n2}}{\sigma_n} & \cdots & \frac{s_{nn}}{\sigma_n} \end{bmatrix} \tag{9}$$

IV. Results

We used Matlab for the simulation of various leak scenarios. The water network is defined in an Excel file. The nodes are defined by an ID, coordinates, value (pressure, if known) and type (real/simulated leak). The edges are defined in a separate sheet by input node ID, output node ID, resistance, value (flow, if known). The Matlab script first reads the Excel file and creates the node and edge object list which is used in the previously described algorithm. Then, the model equations are extracted and the optimal solution is calculated using the least squares method. The results are assigned according to their corresponding parameter (pressure, flow) and the graph representation is shown using the Matlab function *gplot*.

The results obtained under normal conditions (without leaks) are shown in figure 4 and those obtained when there is a leak are shown in figure 5. The simulated leak is represented by an additional demand from node 9. The calculated variables are shown in blue and leak nodes are shown in red. [11]

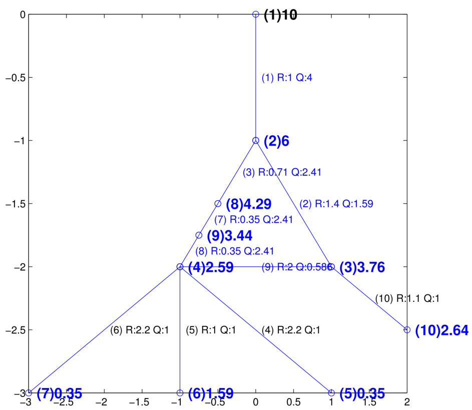

Fig. 4. Results without leaks

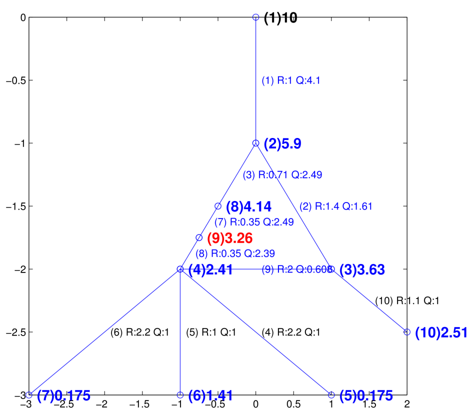

Fig. 5. Results with leaks

For calculating the effects of every possible leak on measured parameters, the sensitivity matrix is calculated by simulating every possible scenario. We used a 3D plot of the sensitivity matrix to highlight the most sensitive areas using a visual representation.

In figure 6 the effect of each leak on the pressure in each node is shown (the highest sensitivity is shown in red and the lowest sensitivity is shown in blue). The actual test cases are represented by blue dots, with results in between being interpolated. In figure 7, using the same convention, the effect of each leak on flow through each pipe is shown.

Using the normalized sensitivity matrix and a threshold, the location of the leak can be identified in one or more nodes in the network. When comparing the results obtained

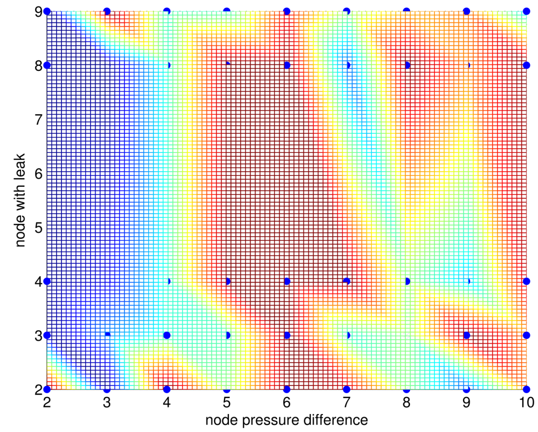

Fig. 6. The effects of leaks on pressure discrepancies

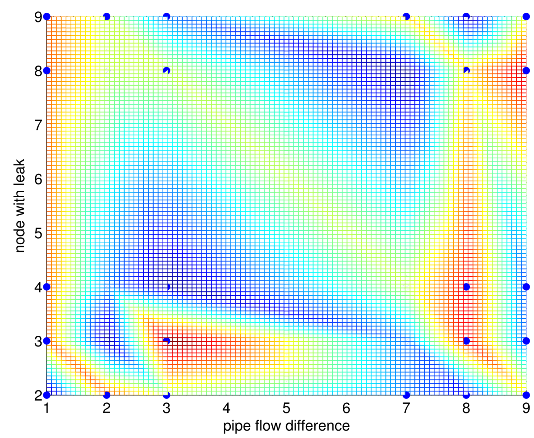

Fig. 7. The effects of leaks on flow discrepancies

from pressure and flow measurements, we find that pressure discrepancies are higher downstream and flow discrepancies are higher upstream. This can indicate that a leak in a consumer node (demand) can be more accurately identified by using flow sensors in adjacent nodes. This would allow for separation of problem areas and differentiation of distribution and consumer part. Also, the domain of uncertainty is lower in the case of flow sensors. For the previously shown scenario having a leak in node 9, the standard deviation for measured pressure in internal nodes is 0.0286 and the standard deviation for measured flow is 0.0464. This is in accordance to the higher sensitivity of flow measurements compared to pressure measurements, which is mentioned in this article.

This simulation is based on steady-state model equations, and it is used to estimate the location of a leak by comparing data from an ideal scenario to data from a leak scenario. In real applications, the location of a leak will be estimated using a statistical model with data from pressure and flow sensors. In both cases, a calibrated hydraulic model would validate the estimations and would provide accurate results.

V. Conclusion

In this article, we proposed an implementation of leak detection methods and sensitivity measurements based on a water distribution system which is defined in the context of the INSPIRE directive for spatial information. The proposed solution is described in terms of the conceptual data model and the algorithms which can be included as part of an autonomous system for better management of the water distribution system. Pressure measurements are compared to flow measurements in terms of sensitivity to leaks which can be useful for calculating the cost of the system, as flow sensors are usually more expensive while a higher number of pressure sensors would be necessary for obtaining a similar precision in leak detection algorithms. The results will be validated using an experimental setup and a web based implementation will be used in a small-scale water distribution system.

Acknowledgment

The paper presentation has been supported by the Data 4 Water - H2020 Twinning project no. 690900.

References

[1] C. Lupu, D. Chirita, S. Iftimie, and R. Miclaus, "Consideration on leak/fault detection system in mass transfer networks," *Energy Procedia. EENVIRO 2016*.

[2] European Commission. INSPIRE. [Online]. Available: http://inspire.ec.europa.eu/

[3] INSPIRE Thematic Working Group Utility and Government Services, "D2.8.iii.6 inspire data specification on utility and government services – technical guidelines," *INSPIRE Infrastructure for Spatial Information in Europe*.

[4] R. Pérez *et al.*, "Pressure sensor distribution for leak detection in barcelona water distribution network," *Water Science & Technology Water Supply*, 2009.

[5] M. Stachura and B. Fajdek, "Planning of a water distribution network sensors location for a leakage isolation," *Proceedings of the 28th EnviroInfo 2014 Conference, Oldenburg, Germany*.

[6] F. Nejjari, R. Sarrate, and J. Blesa, "Optimal pressure sensor placement in water distribution networks minimizing leak location uncertainty," in *Procedia Engineering 119*, 2015, pp. 953–962.

[7] R. Pérez, V. Puig *et al.*, "Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks."

[8] G. Sanz and R. Pérez, "Comparison of demand calibration in water distribution networks using pressure and flow sensors," in *13th Computer Control for Water Industry Conference*, 2015.

[9] ESRI. ArcGIS. [Online]. Available: https://www.arcgis.com/features/index.html

[10] Georgia State University. Hyperphysics. [Online]. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/ppois.html

[11] MathWorks. Matlab. [Online]. Available: https://www.mathworks.com/