Activity 3 : Leak Detection and Location in water distribution Networks (WDNs)

Objectives: Design and implement a leak detection and localization method that uses data from available sensors in WDN: Flow sensors in WDN inlets and pressure sensors in inner nodes. A well-known benchmark, the simplified model of the Hanoi WDN (Vietnam), will be used to assess the performance of the proposed method. Different sizes of leaks will be studied as well as different sensor placements. A new

Process description:

Simplified hydraulic model of Hanoi WDN was introduced in

O. Fujiwara and D. B. Khang, "A two-phase decomposition method foroptimal design of looped water distribution networks,"Water resourcesresearch, vol. 26, no. 4, pp. 539–549, 1990.

This network has 31 consumption nodes and 34 pipes organized in 3 loops. No pumping facilities are considered since only a single fixed head source at elevation of 100m is available.

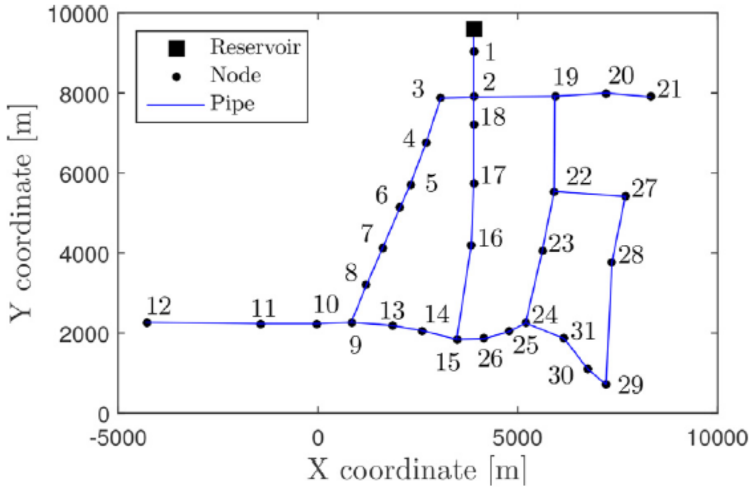

Fig 1: Hanoi WDN

Given the water consumption at every node, the pressure map in inner nodes of the WDN can be obtained by means EPANET. EPANET is a public domain, water distribution system modeling software package developed by the United States Environmental Protection Agency's (EPA) Water Supply and Water Resources Division.

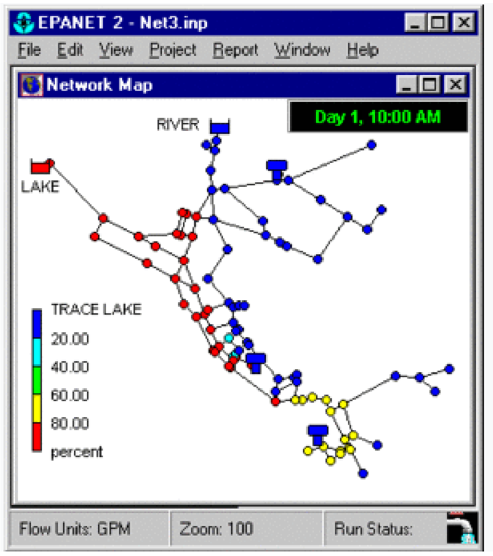

Fig. 2: EPANET Software

L. A. Rossman, "EPANET 2: Users manual," US EnvironmentalProtection Agency, Tech. Rep. EPA/600/R-00/057, September 2000.

<u>Work to be done:</u>

1. **Leak Detection (3 points):** Given a sequence of inlet flow data *"leak_free_data.mat"* where there is assumed that no leak was present in the WDN build a total demand model that assumes periodicity in the flow consumption of the users. Use the obtained model for leak detection in two different leak scenarios: *"leak_scenario1_data.mat"* and *"leak_scenario2_data.mat"*. Provide figures of the different results: nominal demand model, uncertainty bounds for every hour, leak detection results and discuss about the results obtained (leak detection times).
2. **Leak localization I (1 point):** From pressure residual data *"nominal_residuals.mat"* computed using Epanet models with and without leak

as the scheme of Fig. 3 and considering inner sensors installed in nodes 14 and 30 compute the leak sensitivity matrix. Remark: *"nominal_residuals.mat"* has computed for different leak magnitudes and for the 31 inner nodes, the plot of the data for the considered inner sensors installed in the network should be similar to Fig. 4. Discuss about the results obtained.

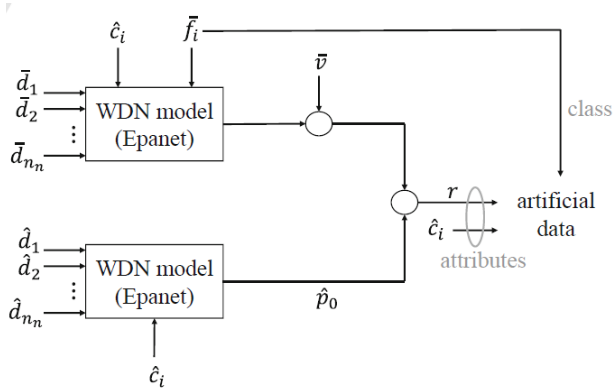

Fig. 3. Residual generation

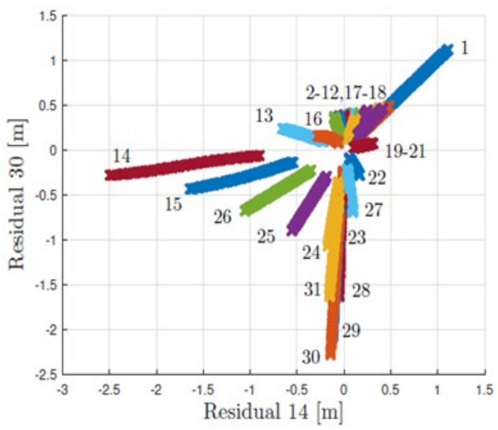

Fig. 4. Nominal Residuals considering inner sensors 14 and 30

3. **Leak localization II (4 points):** Considering the sensitivity matrix computed in the previous Section apply a leak location procedure by means correlation analysis to the leaky data contained in the files: *"hanoi_residuals_f_20.mat"* and *"hanoi_residuals_f_50.mat"* That contain data from small to medium size leaks. In order to evaluate the performance of the leak localization method, you will have to compute the confusion matrix Γ.

	$\hat{f}_1$	$\cdots$	$\hat{f}_i$	$\cdots$	$\hat{f}_n$
f_1	$\Gamma_{1,1}$	$\cdots$	$\Gamma_{1,i}$	$\cdots$	$\Gamma_{1,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_i	$\Gamma_{i,1}$	$\cdots$	$\Gamma_{i,i}$	$\cdots$	$\Gamma_{i,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_n	$\Gamma_{n,1}$	$\cdots$	$\Gamma_{n,i}$	$\cdots$	$\Gamma_{n,n}$

Table 1. Components of confusion matrix Γ

The confusion matrix is a square matrix with as many rows and columns as nodes in the network (potential leak locations), where each coefficient Γ_{ij} indicates how many times a leak in node *i* is recognized as a leak in node *j*. In case of a perfect classification, the confusion matrix should be diagonal, with $\Gamma_{ii} = m_i$ with m_i the number of residuals generated considering leak *i*. In practice, non-zero coefficients will appear outside the main diagonal. For a leak in node *i*, the coefficient Γ_{ii} indicates the number of times that the leak is correctly identified as f_i, while $\sum_{j=1}^{n_n} \Gamma_{ij} - \Gamma_{ii}$ indicates the number of times that is wrongly classified. In order to weight the different errors, the Average Topological Distance (ATD) is defined as the average value of the minimum distance in nodes between the node provided by the leak localization algorithm as most probable leak candidate and the node with the actual leak. Given confusion matrix Γ and a distance matrix $\mathbf{D} \in \mathbb{N}^{\mathrm{l}\times\mathrm{l}}$ whose components D_{ij} contain the minimum distance in nodes from node *i* to node *j* the ATD can be computed as

$$\mathrm{ATD} = \frac{\sum_{i=1}^{n_n} \sum_{j=1}^{n_n} \Gamma_{ij} \mathrm{D}_{ij}}{\sum_{i=1}^{n_n} \sum_{j=1}^{n_n} \Gamma_{ij}}$$

Remark 1: The diagonal elements of matrix **D** are zero, i.e. $D_{ii} = 0 \quad \forall i = 1,...,n_n$

Remark 2: The term Γ_{ij} contributes to the ATD weighted by D_{ij} i.e. as the bigger is the distance between nodes *i* and *j* the more contribution to the ATD. e.g Γ_{ii} as $D_{ii} = 0$ contributes with 0 (no error)

Remark 3: The term considering m_i the number of residuals generated considering leak *i* equal a N for all the leaks, i.e. $m_i = N \quad i{=}1,...,n_n$ then

$$\sum_{i=1}^{n_n}\sum_{j=1}^{n_n}\Gamma_{ij} = n_n N$$

So as a result of this Section you should compute ATD_{20} and ATD_{50} from confusion matrices Γ_{20} and Γ_{50} computed with respective leak residual data files. Discuss about the results obtained, what are the obtained ATD_{20} and ATD_{50}? What are the leaks that are more difficult to separate?

4. **Leak localization III (2 points):** Considering inner pressure sensors installed in nodes 10 and 12 instead of in 14 and 30 as was considered in previous sensors, compute ATD_{20} and ATD_{50} and compare with the ones obtained in the previous Section. Discuss about the results obtained. What are the leaks that are more difficult to separate?