Modularity Index for Optimal Sensor Placement in WDNs

**Antonietta Simone, Daniele Laucelli, Luigi Berardi
and Orazio Giustolisi**

1 Introduction

The availability of a huge amount of data from infrastructure networks, e.g., water distribution networks (WDNs), continuously encourages researchers expanding the technical tasks that water companies could consider in order to increase service reliability and quality [1]. These complex networks present many problems related with their nonhomogeneous behavior and, simplifying these systems into smaller monitored modules/districts, could reduce many of management difficulties (e.g., managing pressures, demands, leakages, rehabilitation works, etc.). Several approaches to segmentation have been proposed [2–11] for identifying the optimal division of the network into districts with respect to WDN characteristics (e.g., diameter, leakages, elevations, etc.) and topology.

A way to segment networks, using a paradigm from complex network theory [12, 13], is to refer to community detection strategies. The most popular community detection strategy is based on modularity index [14], i.e., a descriptive measure of topology that relies strictly on the network structure. Newman [14] firstly proposed the modularity index as metric to measure the propensity of the network division into modules. High values of the modularity index indicate a better identification of communities and to the maximum value of the modularity corresponds the

A. Simone (✉) · D. Laucelli · L. Berardi · O. Giustolisi
Politecnico di Bari, 70125 Bari, Italy
e-mail: antonietta.simone@poliba.it

D. Laucelli
e-mail: danielebiagio.laucelli@poliba.it

L. Berardi
e-mail: luigi.berardi@poliba.it

O. Giustolisi
e-mail: orazio.giustolisi@poliba.it

maximum degree of segmentation [13, 15]. The original formulation of the modularity index for segmentation of immaterial networks has been proposed by Barthélemy [16], which stressed the need for paying attention to the uncritical application of complex network tools to infrastructure networks, i.e., networks strongly affected by a number of physical constraints (two-dimensionality, urban structure and planning, etc.). Afterwards, Giustolisi, and Ridolfi [17] tailored the original modularity index in order to obtain a WDN segmentation-oriented modularity index by means of the topological incidence matrix of WDNs (commonly used to describe the network topology of the hydraulic systems). They developed a cut position-sensitive index in order to account for the actual position along pipes (i.e., close to ending nodes instead of the middle of pipes) of devices generating districts, introducing the pipe weights (i.e., asset and/or hydraulic information at pipe level) in the formulation of the segmentation-oriented modularity index. The tailored modularity index allows dividing the network into districts similar to each other with respect to the assumed weight (e.g., pipe lengths, leakages, etc.). Giustolisi and Ridolfi [17] also proposed the extension of the tailored modularity to the division into districts having internal similar attribute (e.g., material, diameter, age, average elevation, average pressure, etc.). Since the proposed segmentation-oriented modularity index was affected by a resolution limit [18], i.e., a limit to the ability in identifying small districts whose size is related to the network dimension, Giustolisi, and Ridolfi [19] proposed the infrastructure modularity index. They demonstrated that the infrastructure index increases the resolution in identifying small districts also in large size networks and is unbiased with respect to the optimal planning considering already existing devices in the network.

Simone et al. [15] extended the concepts of network segmentation to pressure sampling design, introducing the sampling-oriented modularity index and the concept of "pressure" measurement districts extending the concept of "flow" measurement districts related to the "classic" district metering areas. This way the need of dividing a WDN considering pressure meters was introduced. Furthermore, the strategy has several analogies with the optimal segmentation design for "classic" district monitoring areas (DMAs) in order to allow the integration of the WDN division in districts to account for the mass and energy standpoints, which drives the hydraulic behavior of the system. Planning is a continuous and sequential process supporting management of WDNs, which is driven by budget and available information about the hydraulic system. Therefore, the budget and information uncertainties ask for flexibility of plans, i.e., a multi-objective optimal planning of sensor placement should account for the fact that a water company will start installing a lower number of devices which will be increased in the future also based on the information coming from the installed monitoring system itself. Therefore, the multi-objective strategy needs to provide solutions, which, increasing the number of devices, are one the starting point of the other. In other words, both optimal segmentation and optimal sampling design need to provide solutions which are nested, i.e., each scenario of districts corresponding to a greater resolution solution (i.e., greater number of districts) has to correspond to a nested

scenario with respect to lower resolution solutions. Another issue is the integration of the optimal segmentation with the optimal sampling design.

The present paper proposes a novel two-phase strategy for optimal flow and pressure sensors placement. The first step involves the optimal segmentation design in order to achieve scenarios of optimal positions of "conceptual cuts" dividing the network in "flow" measurement districts. Actually, the segmentation is the first phase to identify "classic" DMAs, which are formed deciding to install in the "conceptual cuts" flow meters or closed gate valves for example in order to reduce leakages [1].

The second phase involves the optimal sampling design [15] starting from the placement of pressure meters indicated by the optimal segmentation [17]. This way the "classic" DMAs (or "flow" measurement districts) are bounded by pressure meters, and the optimal sampling involves planning further pressure meters internal to "classic" DMAs. In other words, the optimal sampling starts assuming existing pressure meters on the boundary of the "flow" measurement districts of the optimal segmentation and the optimal planning of further pressure meters identifies internal "pressure" measurement districts. Therefore, the integrated optimal sensor placement allows planning "classic" DMAs coinciding with "pressure" DMAs once the pressure meters are assumed at their boundary. Then, other "pressure" DMAs [15] are designed internal to "classic" ones. The integrated strategy for optimal placement of flow and pressure meters is demonstrated and discussed using a real network of the Apulia region in Italy.

2 Modularity Index for WDN

The modularity index, Q, is a measure of the strength of a network division in modules. Newman and Girvan [10] proposed the first formulation of the modularity index:

$$Q = \frac{1}{2n_1} \sum_{ij} \left(A_{ij} - P_{ij}\right)\delta\left(\mathrm{M}_i, \mathrm{M}_j\right) = \frac{1}{2n_1} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2n_1}\right)\delta\left(\mathrm{M}_i, \mathrm{M}_j\right) \qquad (1)$$

where n_1 is the number of links/edges (pipes for WDN) in the network, A_{ij} are the elements of the adjacency matrix, P_{ij} is the expected fraction of links between nodes i and j in the random network, M_i is the identifier of network modules, δ is the function to apply the summation to the elements of the same module (i.e., $\delta = 1$ if $\mathrm{M}_j = \mathrm{M}_i$ and $\delta = 0$ otherwise), k_i (k_j) is the degree of the i-th (j-th) node and summation runs on all the possible node couples (i, j), with $i \neq j$.

The term $\Sigma_{ij}A_{ij}\delta(\mathrm{M}_i, \mathrm{M}_j)/2n_i$ represents the fraction of links connecting nodes that are in the same module and the term $\Sigma_{ij}k_i k_j\delta(\mathrm{M}_i, \mathrm{M}_j)/2n_i$ represents the expected fraction of links connecting nodes in the same module in a random graph having the same degree distribution of the original graph [20].

The classic formulation of the modularity index, tailored for WDNs using the general topological incidence matrix ($\mathbf{A}_{pn}$), is

$$Q = 1 - \frac{n_c}{n_p} - \sum_{m=1}^{n_m} \left[\sum_{i=1}^{n_n} \frac{\left(\left| \mathbf{A}_{pn}^T \right| \mathbf{u}_p \right)_i \delta(\mathbf{M}_m, \mathbf{M}_i)}{2n_p} \right]^2 \tag{2}$$

where $\mathbf{u}_p$ is the unit vector, n_p is the number of pipes, n_n is the number of nodes, n_m is the number of modules and n_c is the number of pipes linking modules, i.e., the number of cuts in the middle of the pipes. The modularity index formulation can be divided into two components:

$$Q = Q_1 + Q_2$$

$$Q_1 = 1 - \frac{n_c}{n_p} \qquad Q_2 = -\sum_{m=1}^{n_m} a_m^2 = -\sum_{m=1}^{n_m} \left[\sum_{i=1}^{n_n} \frac{\left(\left| \mathbf{A}_{pn}^T \right| \mathbf{u}_p \right)_i \delta(\mathbf{M}_m, \mathbf{M}_i)}{2n_p} \right]^2 \tag{3}$$

where Q_1 represents the fraction of the pipes with both the end nodes belonging to the m-th module and a_m is the fraction of pipes having at least one end node in the module m. The pipes dividing modules are counted ½ when computing a. Therefore, a_m is half the summation, divided by n_p, of the number of pipes connected to nodes (degree) falling in the module m.

Q_1 strictly decreases with the number of cuts and penalizes the excess of cuts for a given number of modules and Q_2 generally is an increasing function of the number of modules (and generally of n_c), driving the search to the set of most similar modules for a given number of cuts. Those cuts are virtual and relate to the division into modules ("flow" measurement districts), i.e., the metric Q can be explained as a measure of the module decomposition of the adjacency matrix of the network. The maximization of modularity index in Eq. (3) implies the minimization of the number of cuts in order to obtain the highest number of modules, which are similar to each other. The formulation of the modularity index was then tailored in order to develop a cut position-sensitive metric [17]. In fact, the segmentation for WDNs aims at designing DMAs through the installation of real devices (closing gates or flow meters), that are installed close to nodes where vaults or manholes are located [1]. Therefore, the "conceptual cuts" of the modularity index need to be close to the end nodes of pipes, whereas the original modularity assumes that they are in the middle. Then, the segmentation-oriented formulation of the modularity index [17] is:

$$Q(\mathbf{w}_\mathrm{p}) = \left\{ 1 - \frac{n_\mathrm{c}}{n_\mathrm{p}} \right\} + \left\{ -\sum_{m=1}^{n_\mathrm{m}} \left[\sum_{k=1}^{n_\mathrm{p}} \frac{(\mathbf{w}_\mathrm{p})_k \delta(\mathrm{M}_m, \mathrm{M}_k)}{W} \right]^2 \right\} = Q_1 + Q_2 \qquad (4)$$

where n_c is the number of pipes linking modules of the network, namely the number of "conceptual cuts" in the network (i.e., the decision variables of the WDN segmentation problem) and n_m is the number of network modules. The summation inside the square brackets is related to pipe weights stored in the vector $\mathbf{w}_\mathrm{p}$, whose sum is W, and Kronecker's δ function makes that the sum refers only to the weights of pipes belonging to the m-th module (i.e., $\delta = 1$ if $\mathrm{M}_m = \mathrm{M}_k$ and $\delta = 0$ otherwise).

It is worth noting that the term Q_1 of Eq. (4) decreases with the number of cuts, while Q_2 generally increases with the number of modules and, for a given number of modules, it increases with the similarity between modules. The resolution limit problem [18] for both original and segmentation-oriented modularity indexes was analyzed by Giustolisi and Ridolfi [19], whose proposed a new infrastructure segmentation-oriented modularity index to overcome such limit:

$$IQ(\mathbf{w}_\mathrm{p}) = 1 - \frac{n_\mathrm{c} - (n_\mathrm{act} - 1)}{n_\mathrm{p}} - \sum_{m=1}^{n_\mathrm{m}} \left[\sum_{k=1}^{n_\mathrm{p}} \frac{(\mathbf{w}_\mathrm{p})_k \delta(\mathrm{M}_m, \mathrm{M}_k)}{W} \right]^2 \qquad (5)$$

where n_act is the actual number of modules satisfying given constraints (e.g., the minimum length of the modules, the minimum number of pipes, etc.). Accordingly, the same authors demonstrated that the infrastructure modularity resolves the resolution limit, but might require the definition of technical constraints to avoid a resolution of the segmentation beyond the required by specific technical tasks [19]. The maximum value of the infrastructure modularity index IQ results:

$$IQ_\mathrm{max} = 1 - \frac{1}{n_\mathrm{m}} \qquad (6)$$

It is worth noting that the maximum value of IQ is asymptotically upper bounded to unit (for an infinite number of modules) as well as in the case of Q, while $IQ = 0$ for an unsegmented network, i.e., it strictly depends on the number of modules. A variety of purposes for designing "flow" monitoring districts exists, e.g., managing leakage management [1]. Consequently, various pipe weights can be defined considering the specific management task. For example, it is possible to define $w_k = L_k$ = pipe length; $w_k = \alpha L_k P_k^{\alpha}$ = background leakages along a pipe; etc. Giustolisi and Ridolfi [17] also proposed a segmentation-oriented metric, named attribute based, measuring the similarity into each module with respect to a specified attribute, which is not length based:

$$Q_a = 1 - \frac{n_c}{n_p} - \sum_{m=1}^{n_m} \left[\frac{\sum_{k=1}^{n_p} |\mathbf{a}_p - \bar{a}(\mathbf{M}_m)|_k \delta(\mathbf{M}_m, \mathbf{M}_k)}{\sum_{k=1}^{n_p} |\mathbf{a}_p - \bar{a}(\mathrm{N})|_k} \right]^2 \qquad (7)$$

where $\mathbf{a}_p$ is the vector of pipe attributes, $\bar{a}(\mathrm{N})$ is the mean value of the pipe attributes of the network N, i.e., of $\mathbf{a}_p$, and $\bar{a}(\mathbf{M}_m)$ is the mean value of the pipe attributes in $\mathbf{M}_m$. Function δ limits the summation of the pipe attributes to the elements belonging to the same module.

Giustolisi et al. [20] extended the infrastructure modularity index to attribute-based infrastructure segmentation index:

$$IQ_a = 1 - \frac{n_c - (n_{\mathrm{act}} - 1)}{n_p} - \sum_{m=1}^{n_m} \left[\frac{\sum_{k=1}^{n_p} |\mathbf{a}_p - \bar{a}(\mathbf{M}_m)|_k \delta(\mathbf{M}_m, \mathbf{M}_k)}{\sum_{k=1}^{n_p} |\mathbf{a}_p - \bar{a}(\mathrm{N})|_k} \right]^2 \qquad (8)$$

Afterwards, Simone et al. [15] proposed a novel modularity index named sampling-oriented modularity for an optimal sampling design approach. The segmentation-oriented and sampling-oriented modularity metrics differ for the approach of identifying monitoring districts in WDNs. The first approach segments the network considering pipes, i.e., by means of "conceptual cuts", the second considers "pressure nodes". This way the extension of the segmentation-oriented modularity to sampling can be performed by substituting the concept of "pressure nodes" to "conceptual cuts" as follows:

$$Q_{\mathrm{s}}(\mathbf{w}_p) = 1 - \frac{n_{\mathrm{obs}}}{n_{\mathrm{n}}} - \sum_{m=1}^{n_{\mathrm{m}}} \left[\sum_{k=1}^{n_{\mathrm{p}}} \frac{(\mathbf{w}_p)_k \delta(\mathbf{M}_m, \mathbf{M}_k)}{W} \right]^2$$

$$IQ_{\mathrm{s}}(\mathbf{w}_p) = 1 - \frac{n_{\mathrm{obs}} - (n_{\mathrm{act}} - 1)}{n_{\mathrm{n}}} - \sum_{m=1}^{n_{\mathrm{m}}} \left[\sum_{k=1}^{n_{\mathrm{p}}} \frac{(\mathbf{w}_p)_k \delta(\mathbf{M}_m, \mathbf{M}_k)}{W} \right]^2$$

$$Q_{\mathrm{a-s}}(\mathbf{a}_p) = 1 - \frac{n_{\mathrm{obs}}}{n_{\mathrm{n}}} - \sum_{m=1}^{n_{\mathrm{m}}} \left[\frac{\sum_{k=1}^{n_{\mathrm{p}}} |\mathbf{a}_p - \bar{a}(\mathbf{M}_m)|_k \delta(\mathbf{M}_m, \mathbf{M}_k)}{\sum_{k=1}^{n_p} |\mathbf{a}_p - \bar{a}(\mathrm{N})|_k} \right]^2$$

$$IQ_{\mathrm{a-s}}(\mathbf{a}_p) = 1 - \frac{n_{\mathrm{obs}} - (n_{\mathrm{act}} - 1)}{n_{\mathrm{n}}} - \sum_{m=1}^{n_{\mathrm{m}}} \left[\frac{\sum_{k=1}^{n_{\mathrm{p}}} |\mathbf{a}_p - \bar{a}(\mathbf{M}_m)|_k \delta(\mathbf{M}_m, \mathbf{M}_k)}{\sum_{k=1}^{n_p} |\mathbf{a}_p - \bar{a}(\mathrm{N})|_k} \right]^2$$

$$(9)$$

where n_{obs} are the number of "pressure nodes". The number of modules n_{m} has the same meaning of the segmentation case, as well as the actual number of modules matching the technical constraints to avoid excessive resolution of the segmentation, n_{act}, of the infrastructure version. Note that the term Q_2 is unchanged because it refers to the characteristics of modules, which are now created by removed nodes. Therefore, the Eq. (9) define, respectively, the sampling-oriented modularity Q_{s}, sampling-oriented infrastructure modularity IQ_{s}, sampling-oriented attribute-based

modularity Q_{a-s} and sampling-oriented infrastructure attribute-based modularity IQ_{a-s}. The first two divide the network in "pressure" monitoring districts similar to each other with respect to the assumed pipe weights, while the last two divide the network in "pressure" monitoring districts having similar internal characteristics with respect to the assumed pipe weights.

3   Monitoring Strategy

The division of WDNs into districts by means of "flow" and "pressure" monitoring is a useful practice for system management. In fact, a rationale system of flow and pressure observations allows monitoring and analyzing the hydraulic system behavior. For example, it is useful for assessing leakage level and designing optimal actions for leakage management activities [1].

   To this purpose, it is here proposed an optimal monitoring strategy, i.e., an optimal sensor placement strategy. Two main phases drive the strategy:

 (i) The optimal segmentation design [17] minimizes the number of "conceptual cuts" versus the maximization of the segmentation-oriented modularity index reported in Eqs. (5) or (8). It returns a Pareto set of solutions, which are optimal with respect to the number of "conceptual cuts" (i.e., the candidate positions for flow meters, but also for pressure reduction valves and closed gates once the DMA are built). The multi-objective optimization is constrained to search for nested solutions providing flexibility to the entire procedure. This means that each segmentation solution can be generated starting from the one having the lowest number of districts increasing number of districts, which are nested in the previous segmentation. Such feature allows dynamical planning of the segmentation increasing over time its resolution (i.e., the number of districts) considering budget uncertainty and growing of the system knowledge.

(ii) The optimal sampling design starts from the selected optimal segmentation solution assuming pressure meters at the boundary nodes of each district forming the "classic DMAs". This way, the optimal sampling design returns the placement of pressure meters inside the "classic" DMAs forming internal "pressure" DMAs. The optimal sampling design [15] minimizes the number of "pressure nodes" versus the maximization of one of the sampling-oriented modularity index reported in Eq. (9). Clearly, the procedure is biased by the assumed pressure meters at the boundary of the already designed "flow" districts.

   For the sake of clarity, the strategy is here applied to a very small network, named Apulian [15].

   The first phase of the strategy, i.e., the optimal segmentation design, returns a Pareto set of optimal segmentation solutions by solving the following problem:

$$
\begin{cases}
[M, n_{\mathrm{c}}, n_{\mathrm{act}}] = \mathrm{Connectivity}\left(I_{\mathrm{c}}, \left|\mathbf{A}_{\mathrm{np}}\right|\right) \\
f_1 = \max\left\{IQ(\mathbf{w}_{\mathrm{p}})\right\} = \max\left\{IQ(\mathbf{w}_{\mathrm{p}}) = 1 - \frac{n_{\mathrm{c}} - (n_{\mathrm{act}}-1)}{n_{\mathrm{p}}} - \sum_{m=1}^{n_m}\left[\sum_{k=1}^{n_p}\frac{(\mathbf{w}_{\mathrm{p}})_k \delta(M_m, M_k)}{W}\right]^2\right\} \\
f_2 = \min\{n_{\mathrm{c}}\}
\end{cases}
$$

$$(10)$$

where $\mathbf{A}_{\mathrm{np}}$ is the incidence matrix, I_{c} is the set of n_{c} cuts in the network, the Connectivity$(I_{\mathrm{c}}, \left|\mathbf{A}_{\mathrm{np}}\right|)$ stands for component analysis of the undirected graph for the given cuts, n_{act} indicate the cuts that are used to separate modules and W is the sum of pipe weights stored in the vector $\mathbf{w}_{\mathrm{p}}$ [17].

It is important to note that the existing flow meters (e.g., of pumps, tanks, reservoirs, etc.) are considered as constraints corresponding to existing "conceptual cuts". It represents an initial configuration biasing the optimization; therefore, the proposed procedure is flexible with respect to the applications in real hydraulic systems [17, 19]. Figure 1 reports a "flow" districts scenario corresponding to the optimal solution having the maximum value of the segmentation-oriented modularity. The "conceptual cuts", equal to nine, divide the network into five "flow" measurement districts, i.e., "classic" DMAs once the closed gates are installed [1].

As introduced above, the "conceptual cut" close to the unique reservoir of the network is assumed because the correct practice asks for a flow measurement. Therefore, this existing device is not counted during the optimization procedure,
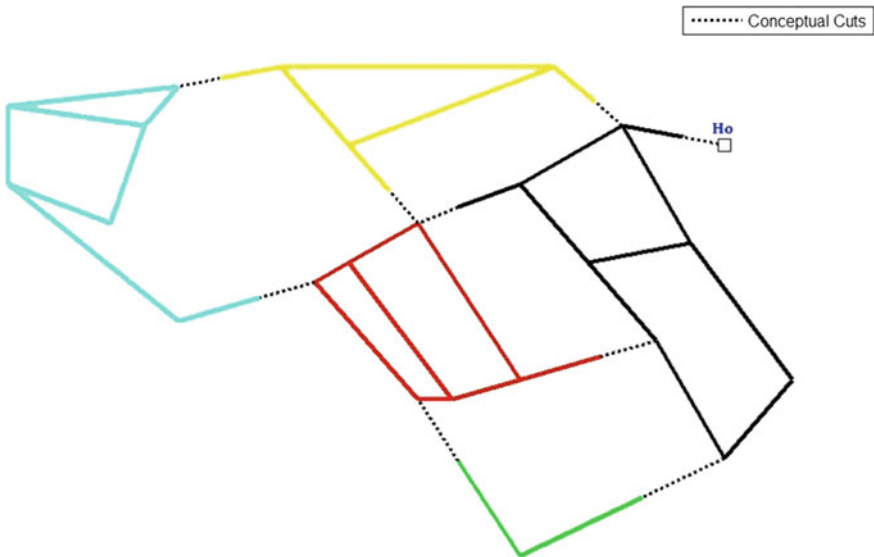
Fig. 1 Apulian network: a scenario of "flow" measurement districts. Each district corresponds to a color
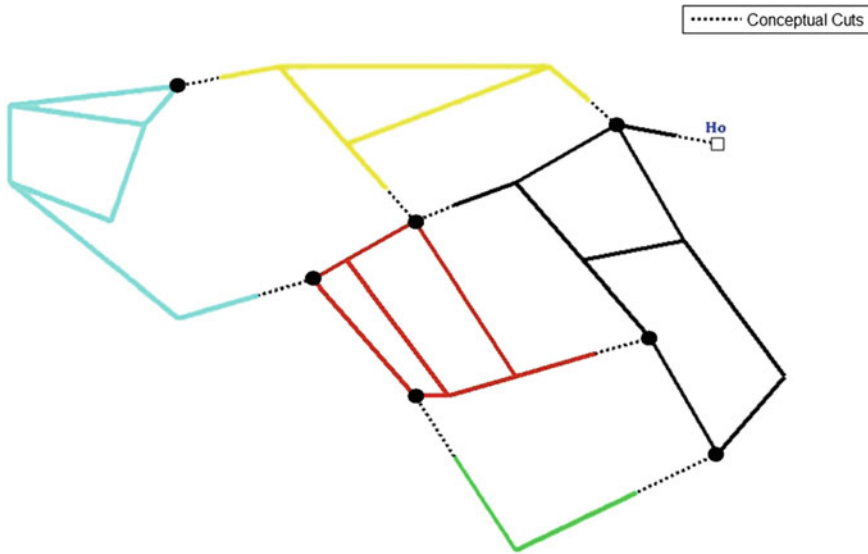
Fig. 2 Apulian network: pressure meters on the boundary of districts of Fig. 1

which starts from an initial configuration that already contains the "conceptual cut" related to the flow observation at the reservoir.

The second phase involves the optimal sampling design starting from the placement of pressure meters indicated by the optimal segmentation. In fact, "conceptual cuts" identified in the first phase suggests the placement of pressure meters at the boundary of the "flow" districts as reported in Fig. 2.

This way, the "flow" districts of Fig. 1 becomes "pressure" districts [15], i.e., the "classic" DMAs coincide with the "pressure" DMAs. The sampling design phase have to be completed, for large size networks, by planning further internal pressure meters identifying "pressure" districts internal to "flow" ones. The problem to solve is similar to that of Eq. (10):

$$
\begin{cases}
[M, n_{\mathrm{obs}}, n_{\mathrm{act}}] = \text{connectivity}(I_{\mathrm{c}}, \mathbf{L}) \\
f_1 = \max\{IQ_{\mathrm{s}}(\mathbf{w}_{\mathrm{p}})\} = \max\left\{ 1 - \frac{n_{\mathrm{obs}} - (n_{\mathrm{act}}-1)}{n_{\mathrm{n}}} - \sum_{m=1}^{n_{\mathrm{m}}} \left[\sum_{k=1}^{n_{\mathrm{p}}} \frac{(\mathbf{w}_{\mathrm{p}})_k \delta(\mathrm{M}_m, \mathrm{M}_k)}{W} \right]^2 \right\} \\
f_2 = \underset{I_{\mathrm{c}}}{\min}\{n_{\mathrm{obs}}\}
\end{cases}
$$

(11)

where $\mathbf{L}$ is the edge adjacency matrix, I_{c} is the set of n_{obs} (i.e., pressure nodes) and connectivity(I_{c}, $\mathbf{A}$) stands for component analysis of the graph with respect to edge matrix. Note that being the decision variables related to new pressure meters to be

installed, in this case, the pressure measurements from segmentation and, e.g., of control valves, pumps, tanks, reservoirs, etc., are considered as constraints. The case study will show this last part of the strategy.

4   Case Study

The optimal monitoring strategy is here presented using a real WDN serving a town in Southern Italy with about 24,000 inhabitants. The network is composed of 2099 pipes and 1762 nodes and its layout is reported in Fig. 3. The node elevation ranges between 302 and 218 m a.s.l. and the highest point of the hill where the town is located is about 261 m a.s.l. The network is served by gravity from the unique reservoir, which is about 4 km far from the distribution network, by two feeding lines having, respectively, nominal diameters of 500 and 300 mm.

The optimal segmentation returned a Pareto set of optimal segmentation solutions (Fig. 4), where the black circles represent the optimal tradeoffs between the number of "conceptual cuts" (x-axis) on the border of each "flow" measurement district versus the maximization of the segmentation-oriented modularity (y-axis). It is worth noting that the number of "conceptual cuts" (x-axis) starts from a value equal to one, to indicate that the strategy starts from a basic configuration that already contains a cut, a flow meter, close to the reservoir.

The solution with the maximum number of "conceptual cuts" corresponds to the maximum value of the segmentation-oriented modularity index. A number of "conceptual cuts" equal to 76 and a number of "flow" districts equal to 19 characterize the last solution (Fig. 5). It is here selected as the basis for the second phase also considering flexibility of the segmentation solutions.

Fig. 3 Network layout
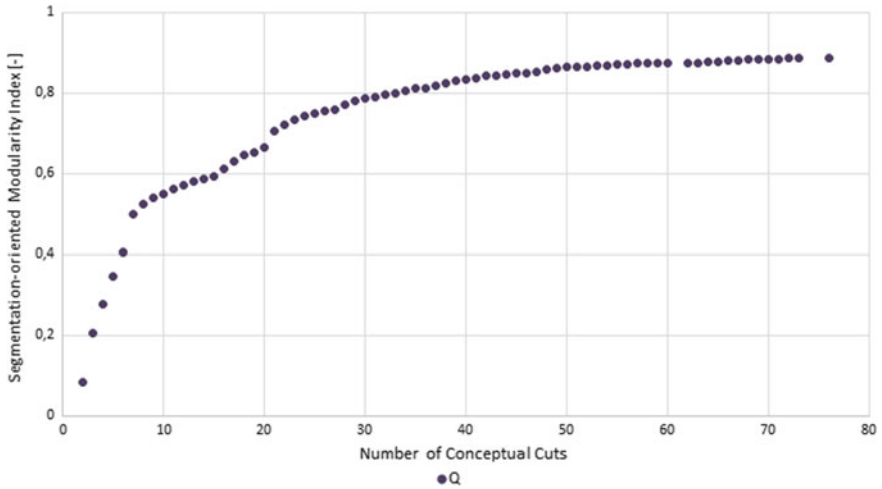
Fig. 4 Pareto set of optimal segmentation solutions
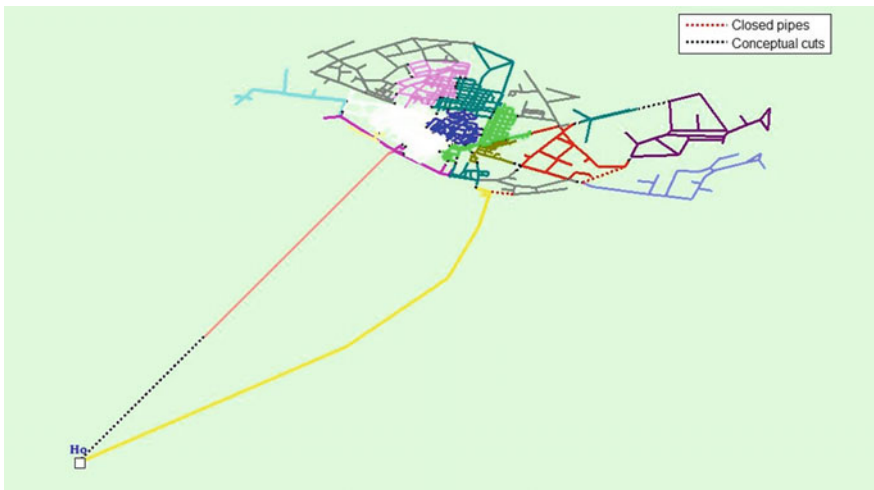
Fig. 5 Colored scenario of flow measurement districts corresponding to the maximum value of the segmentation-oriented modularity

The optimal sampling is biased by the decision to place pressure meters on the boundary of the "flow" districts, i.e., in nodes adjacent to "conceptual cuts", see Fig. 6.

The task of the optimal sampling is then to plan internal "pressure" measurement districts. The two-objective optimization returns a Pareto set of optimal sampling solutions (Fig. 7), where the black circles represent the optimal tradeoffs between

**Fig. 6** Pressure meters at the boundary of the "flow" districts of Fig. 5
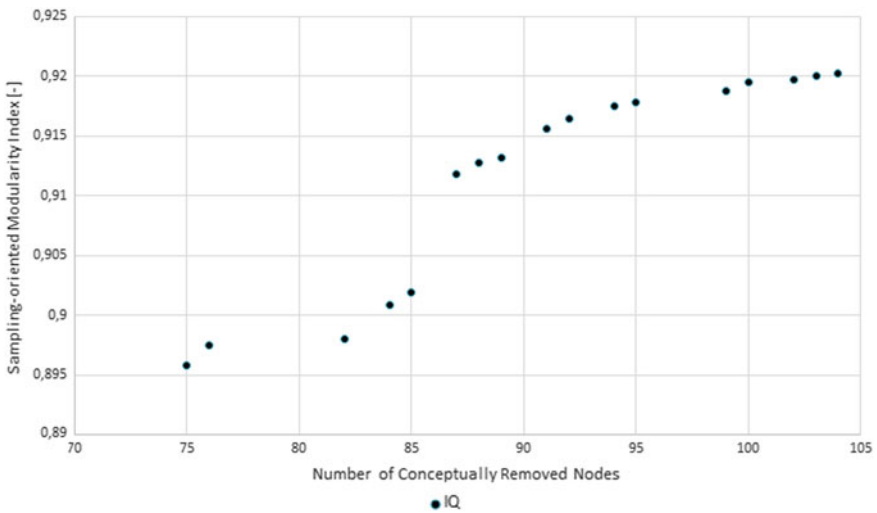
Fig. 7 Pareto set of optimal sampling solutions

the number of pressure nodes (*x*-axis) on the border of each "pressure" district versus the maximization of the sampling-oriented modularity (*y*-axis). It is worth noting that the number of "pressure nodes" (*x*-axis) starts from a value equal to 75, to indicate that the strategy starts from the pressure meters conceived at the boundary of "flow" measurement districts.
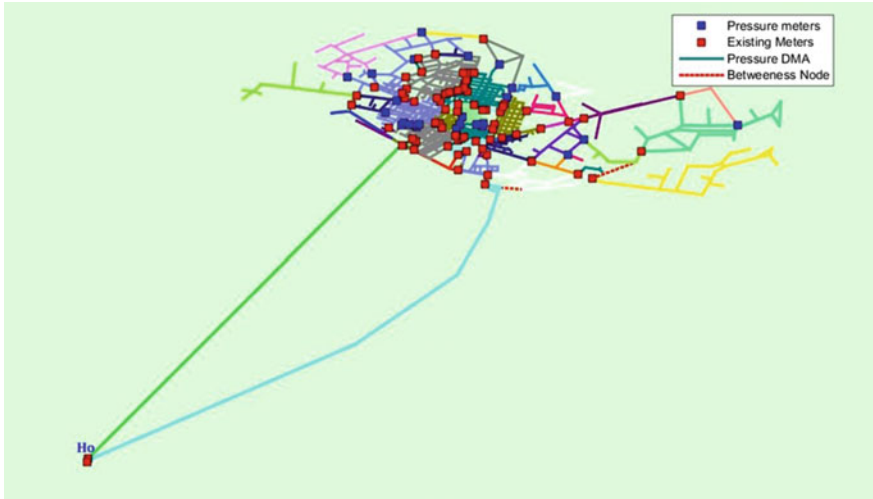
Fig. 8 Optimal monitoring solution

The solution with the maximum number of "pressure nodes", equal to 104, and a number of "pressure" districts equal to 78 characterizes the scenario having the maximum value of the sampling-oriented index.

The optimal monitoring solution for this network, with flow and pressure observations, is reported in Fig. 8, where the red squares represent the location of existing pressure measurements close to flow ones and the blue squares represent the location of pressure measurements from the optimal sampling.

Figure 9 expands a small portion of the network and shows the configurations from the first phase to the second one. The left panel of Fig. 9 shows the "pressure" districts based on the segmentation procedure and the right panel shows the additional "pressure" districts based on the optimal sampling procedure, which divides internally the existing ones.
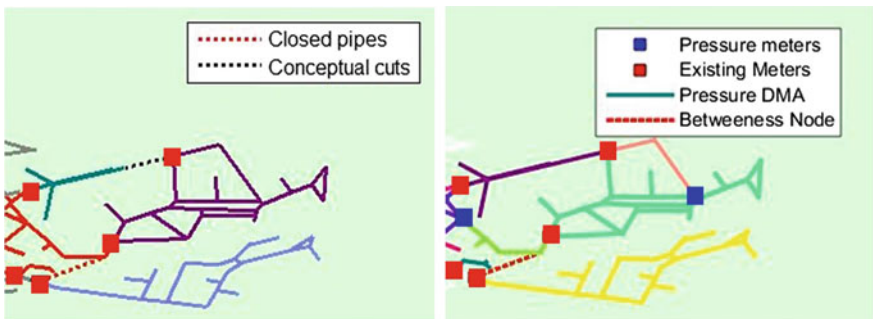
Fig. 9 "Pressure" districts based on segmentation (left panel) and internal "pressure" districts based on optimal sampling (right panel)

5   Conclusions

The integrated planning of the monitoring system strategy is conceived to support water utilities for different management targets, choosing the optimal solution considering, for example, the available budget for flow/pressure observations.

The present work proposes a novel methodology, with a two-phase strategy for optimal placement of flow and pressure meters. The first phase deals with the optimal segmentation using the segmentation-oriented modularity index in order to plan flow and pressure meters at the boundary of "classic" DMAs.

The second phase, starting from the scenario of pressure meters of the first phase, deals with the optimal sampling design using the sampling-oriented modularity. The second phase allows designing "pressure" DMAs, which are internal to "classic" DMAs. They are also "pressure" DMAs because of the assumption to install pressure meters at their boundary nodes.

References

 1. Laucelli, D., Simone, A., Berardi, L., & Giustolisi, O. (2017). Optimal design of district metering areas for the reduction of leakages. *Journal of Water Resources Planning and Management, 143,* 04017017-1–0401701712.
 2. Yang, S.-L., Hsu, N.-S., Loule, P. W. F., & Yeh, W. W.-G. (1996). Water distribution network reliability: Connectivity analysis. *Journal of Infrastructure Systems, 2,* 54–64.
 3. Walski, T. M. (1983). Technique for calibrating network models. *Journal of Water Resources Planning and Management, 109,* 360–372.
 4. Davidson, J., Bouchart, F., Cavill, S., & Jowitt, P. (2005). Real-time connectivity modelling of water distribution networks to predict contamination spread. *Journal of Computing in Civil Engineering, 19,* 377–386.
 5. Deuerlein, J. W. (2008). Decomposition model of a general water supply network graph. *Journal of Hydraulic Engineering, 134,* 822–832.
 6. Perelman, L., & Ostfeld, A. (2011). Topological clustering for water distribution systems analysis. *Environmental Modelling & Software, 26,* 969–972.
 7. Alvisi, S., & Franchini, M. (2014). A heuristic procedure for the automatic creation of district metered areas in water distribution systems. *Urban Water Journal, 11,* 137–159.
 8. Scibetta, M., Boano, F., Revelli, R., & Ridolfi, L. (2013). Community detection as a tool for complex pipe network clustering. *EPL, 103,* 48001.
 9. Diao, K., Zhou, Y., & Rauch, W. (2013). Automated Creation of District Metered Area Boundaries in Water Distribution Systems. *Journal of Water Resources Planning and Management, 139,* 184–190.
10. Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69,* 026113.

11. Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E, 69,* 066133.
12. Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74,* 47–97.
13. Fortunato, S. (2010). Community detection in graphs. *Physics Reports, 486,* 75–174.
14. Newman M. E. J. (2010). *Networks: An introduction*. UK: Oxford University Press.
15. Simone, A., Giustolisi, O., & Laucelli, D. B. (2016). A proposal of optimal sampling design using a modularity strategy. *Water Resources Research, 52,* 6171–6185.
16. Barthélemy, M. (2011). Spatial networks. *Physics Reports, 499,* 1–101.
17. Giustolisi, O., & Ridolfi, L. (2014). A new modularity-based approach to segmentation of water distribution network. *Journal of Hydraulic Engineering, 140,* 1–14.
18. Fortunato, S., & Barthélemy, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 36–41.
19. Giustolisi, O., & Ridolfi, L. (2014). A novel infrastructure modularity index for the segmentation of water distribution networks. *Water Resources Research, 50,* 7648–7661.
20. Giustolisi, O., Ridolfi, L., & Berardi, L. (2015). General metrics for segmenting infrastructure networks. *Journal of Hydroinformatics, 17,* 505–517.