Hindawi
Mathematical Problems in Engineering
Volume 2017, Article ID 9053238, 12 pages
https://doi.org/10.1155/2017/9053238

Research Article

Social Network Community Detection for DMA Creation: Criteria Analysis through Multilevel Optimization

Bruno M. Brentan,[1] Enrique Campbell,[2] Gustavo L. Meirelles,[1] Edevar Luvizotto Jr.,[1] and Joaquín Izquierdo[3]

[1] *Laboratory of Computational Hydraulics (LHC), Unicamp, Av. Albert Einstein 951, Campinas, SP, Brazil*
[2] *Berliner Wasserbetriebe and Amtsgericht Charlottenburg, HRA 30951 B, 10864 Berlin, Germany*
[3] *FluIng, Universitat Politècnica de València, Camino de Vera, S/N (Edificio 5C-Bajo), 46022 Valencia, Spain*

Correspondence should be addressed to Bruno M. Brentan; brunocivil08@gmail.com

Received 18 October 2016; Revised 18 January 2017; Accepted 26 January 2017; Published 20 February 2017

Academic Editor: Tamas Kalmar-Nagy

Management of large water distribution systems can be improved by dividing their networks into so-called district metered areas (DMAs). However, such divisions must be based on appropriated technical criteria. Considering the importance of deeply understanding the relationship between DMA creation and these criteria, this work proposes a performance analysis of DMA generation that takes into account such indicators as resilience index, demand similarity, pressure uniformity, water age (and thus water quality), solution implantation costs, and electrical consumption. To cope with the complexity of the problem, suitable mathematical techniques are proposed in this paper. We use a social community detection technique to define the sectors, and then a multilevel particle swarm optimization approach is applied to find the optimal placement and operating point of the necessary devices. The results obtained by implementing the methodology in a real water supply network show its validity and the meaningful influence on the final result of, especially, elevation and pipe length.

1. Introduction

The frequent disorderly spatial expansion of cities, mainly in developing countries, compels water utilities to rethink their management practice, aiming at highly efficient systems. Optimal management of water distribution systems (WDSs) requires accurate decisions to reduce the waste of environmental resources and to supply consumers with high quality water. Usually, these decisions are made under uncertainty scenarios, mainly due to the size of the network. The segregation of large networks into nearly independent water supply zones can reduce the uncertainty of the problem, thus allowing smarter operation to get better service conditions [1, 2].

District metered area (DMA) design was introduced in the UK [3] and has been widely applied to pressure management and leakage control [4–8]. Also, the identification of entrance pipes and consequent measurement of inlet flow allow improving the water balance, helping to identify leakage and nonrevenue water. However, network segmentation can be a hard task due to various network characteristics: size, number of loops, topology changes, and necessary modifications of the hydraulic conditions during operation. DMA creation requires not only a perfect knowledge of the topological information of the network, but also a set of criteria able to generate a consistent network partition. Depending on the criteria adopted and the combination of them, different topologies can be found that can improve (or worsen!) the efficiency of the network.

The use of trial and error methodologies, which do not consider global perspectives of WDSs, can result in nonoptimal solutions, thus reducing the possibilities of performance improvement through DMA creation. Various developments of automatic tools to help the process of water network partition have been proposed. Among others, [9] presents a graph theory approach to water distribution network decomposition; [10] applies a method based on machine learning for DMA design; [11] proposes the use of a multiagent based approach to negotiate boundaries in DMA generation; [12] develops an automatic boundary generation to determine DMAs based on a social structure, a tool from the artificial intelligence field; and [13] presents an insightful comparison among global clustering, community structure, and graph partition methodologies applied to two big cities; in this last case, the authors argue positively about the high performance of community structures for DMA design from the computational and clustering viewpoints; however, hydraulic and quality analyses are left out.

Graph clustering and, more specifically, the use of an unknown number of subdivisions, as in the proposal based on social network theory [14] have proven to be a good approach for the sectorization problem. Physical and hydraulic features such as the lowest distance to water source, node elevation, and cumulated demand are easily used as criteria to divide the distribution network. However, considering the approach that the a priori number of DMAs is not generally defined, the performance of each scenario is affected by the DMA design criteria used, which enables the generation of very diverse partitions.

While the former approaches on DMA design had pressure management and leakage control as the main objectives, nowadays multiple criteria have been adopted to generate resilient and high-performance networks in terms of hydroenergetic issues, as observed in [15], where the authors present a graph-theoretical approach for the resilience assessment of large scale WDSs.

More than grouping the nodes inside their respective DMAs, a complete sectorization process must include the selection of an optimal entry choice, that is to say, an optimal Control Unit (CU) placement, with at least one flowmeter to permanently monitor the inlet flow [14]. If necessary, pressure reducing valves (PRVs) may also be installed at the entries (also in other points) of the DMAs for pressure regulation purposes. The maximization of a reliability indicator [6] and the minimization of costs are the most common approaches to find the optimal PRV placement. Also, [16] develops an approach for optimal DMA definition considering the possibility of generating energy by using turbines; graph theory and clustering techniques are used, jointly with Simulated Annealing, to identify the optimal entry and pipe replacement.

Furthermore, water distribution network segmentation implies the closure of some pipes. The consequent reduction of loops in the network can affect directly the reliability of the network [6]. In this way, a suitable balanced scenario among the benefits associated with DMA creation should be considered (or developed), where costs, reliability, and efficiency are jointly taken into account. The evaluation of DMA scenarios can be an important design tool and can help the decision-maker to choose an optimal, hopefully the best, option of DMA configuration.

Reference [17] presents a set of feasible solutions of DMA design and evaluates the solution in terms of the resilience index [18], the number of closed pipes, and the water age. The authors conclude that sectorization can produce a small decrease of some performance indicators, but it is insignificant when compared with the benefits of DMA implementation.

Taking into account the variability of feasible solutions for different nodal aggregation criteria for the problem of DMA design, social network community detection algorithms, grounded in graph theory, are used in this paper to define a set of scenarios of DMA configurations. This approach allows evaluating the effect of DMA creation considering the total or partial isolation of the network and even considering the presence of cascading DMAs. To evaluate the performance of a DMA configuration, various values for the criteria are applied. In a second-level, particle swarm optimization (PSO) is applied to determine the optimal placement and operating points of PRVs by considering all the boundary pipes as possible candidates to become DMAs' entrances. The results are evaluated in terms of resilience index, pressure uniformity, demand similarity, water age, electrical consumption, and implementation costs.

2. Materials and Methods

2.1. DMA Definition by a Community Detection Algorithm. Virtual social networks can be described as specialized graphs aimed at describing the interaction of elements that form part of a society, holding some degree of interdependence among them. In this context, an individual is an entity that generates a contribution to the society (network). One of the aspects of major interest in social network science corresponds to community detection, which allows understanding the organization and function of individuals in the network. Unlike traditional clustering in graphs, community detection is not solely focused on individuals, or nodes (and their features). It also takes into account the connection between them, so that the resulting communities are formed not only by individuals, but also by their interactions. One of the most widely known community detection algorithms is the Walktrap algorithm, proposed by Pons and Latapy [19]. This algorithm is based on random walks over graphs.

In the field of mathematics and probability theory, random walks (or diffusion processes) are defined as stochastic processes in which the position of a particle (walker) in a given instant depends on its previous position and a random variable that determines the direction taken from that previous position towards one of the neighboring nodes. Random paths in a (globally sparse but locally dense) graph tend to get trapped into densely connected parts, which correspond to communities [19]. A random walk in such a graph is a Markov chain that can be described by the information contained in the so-called transition matrix, p. Element p_{ij} of p, giving the transition probability from vertex i to vertex j, is calculated as the ratio $A_{ij}/d(i)$, where A_{ij} is the (i, j) element of the adjacency matrix, A, of the graph, and $d(i) = \sum_j A_{ij}$ is the degree of vertex i, that is to say, the number of its neighbors including itself. The tth power of this matrix, p^t, with elements noted as ${p_{ij}}^t$, gives the probabilities of moving from one node i to another node j through a path of length t. As shown in [19], these probabilities are enough to gather information on the topology of the network.

In this algorithm, distances between nodes and communities (see (1) and (2), resp.) based on matrix p^t are computed by

$$r_{ij}(t) = \sqrt{\sum_{k=1}^{n} \frac{\left(p_{ik}{}^{t} - p_{jk}{}^{t}\right)^2}{d(k)}}, \tag{1}$$

$$r_{c_1 c_2} = \sqrt{\sum_{k=1}^{n} \frac{\left(p_{c_1 k}{}^{t} - p_{c_2 k}{}^{t}\right)^2}{d(k)}}, \tag{2}$$

where i and j are two given nodes, $d(k)$ is the degree of node k, and n is the total number of nodes of the network. Equation (2) is actually a generalization of (1) to compute the distance between communities (c_1, c_2 represent two different communities). For this last calculation, random walks go between randomly chosen nodes in both communities. Note that, in particular, any of those communities may be reduced to just one node, what provides the distance between that node and a community.

These distances are compiled in the so-called dissimilarity matrix of the graph, which is used to feed a clustering process. The dissimilarity matrix of a graph is, thus, the square matrix with elements r_{ij} that gives the distance between every pair of nodes i and j of the graph. When using the hierarchical agglomerative approach based on Ward's method [20], one needs criteria to determine which communities to merge. In this method, the average of the squared distances between each node and its community is defined as an objective function, and the goal is its minimization:

$$\sigma_k = \frac{1}{nc} \sum_{C \in \mathcal{P}_k} \sum_{i \in C} r_{ic}{}^2, \tag{3}$$

where $\mathcal{P}_k$ corresponds to a given partition.

It is worth noting here two aspects: firstly, this function only depends on each community, and its minimization does not require information on other communities. Secondly, the method follows a "greedy" strategy; therefore, any time each pair of adjacent sectors are merged the variation of σ, $\Delta\sigma$, is calculated. The fusion that leads to the lower value of $\Delta\sigma$ is selected as the new partition. This produces a hierarchy of partitions at different levels (dendrogram levels). According to [19], from this set of partitions, the best adapted to some specified requirements can be selected. One of the outcomes of the method is the partition that generates the maximum value of the so-called modularity index, which allows measuring the quality of the subdivision of the network in communities. Such a partition is the one that best reflects the modular structure (blocks of lower/medium diameters separated by bigger diameter pipes) of a WDS.

Let us observe here that the best partition generated by the algorithm can produce extremely small communities, whose implementation could be economically unfeasible. This is the reason why a recursive merging process (see pseudocode) is proposed in this paper, to ensure that all the sectors comply with a series of preestablished constraints.

In the next pseudocode the following notation is used: BP (set of boundary pipes) and CP (set of candidate pipes) represent sets of pipes; index m represents a pipe; the end nodes of m are represented by IN_m, the initial node, and FN_m, the final node; C_i represents a community or sector; L refers to the characteristic used as a criterion (sector total length, sector total demand, sector maximum elevation, etc.); and ∇ represents an operation whose arguments are two values of L; the operation (sum, maximum, etc.) depends on the meaning of L.

Pseudocode

```
Input: a partition of the network with maximized modularity index
(1) A value of $L$ is calculated for every sector
    (community) in the partition
(2) For every $m$, if $\mathrm{IN}_m \in C_i$ and $\mathrm{FN}_m \in C_j$, $i \neq j \rightarrow m \in \mathrm{BP}$
(3) From BP, select the pipes whose end nodes
    belong to communities that meet a specified
    constraint for the considered characteristic
    $L$; build CP with those pipes.
(4) For every $m \in \mathrm{CP}$, let $i$ and $j$ values such that
    $\mathrm{IN}_m \in C_i$ and $\mathrm{FN}_m \in C_j$; if $L_i \nabla L_j$ meets $L_{\mathrm{bound}} \rightarrow$
    $C_i \cup C_j$ replace $C_i$ and $C_j$ in the partition.
(5) The characteristic of every $C$ in the new partition is recalculated.
(6) Steps (1)–(5) are repeated until there are not
    more pipes entering in CP.
Output: Sectors satisfying a series of constraints
```

The essence of the process is in Step (4), which states that two sectors C_i and C_j are merged only if their union $C_i \cup C_j$ produces a new sector that meets the constraint of interest (L_{bound}). Note that only one characteristic can be used as a merging criterion.

It is also important to set a lower limit for the characteristic used to define the size of the sectors. So if at the end of the process there are some sectors with a value lower than the limit, they are declared as *minisectors* (with no valves or CUs). This only applies to minisectors that cannot be merged with larger sectors. In the case that a minisector shares at least one connection with a sector that has reached its maximum feature value, the maximum limit is slightly relaxed to allow their fusion. For example, if the characteristic that is used as a criterion is the sector pipe length (e.g., 30 km maximum constraint), the maximum final length that a given sector eventually will have will equal that maximum length (30 km) plus a value that is smaller than the minimum value of any other sector pipe length (e.g., 4 km); in other words, a value between 30 km and 34 km may also be accepted.

2.2. Optimization Procedure

2.2.1. Optimization Problem Description. Graph theory is useful to describe various problems related to water distribution networks. Following the formulation in [21], let $G =$

(V, E), a graph, where the vertices in V represent the nodes of the network and the edges in E represent the pipes and other link elements. Once the number of DMAs is defined and all the nodes are classified, it is possible to identify a set of boundary pipes $E' \subseteq E$, whose elements (pipes or links) have different DMAs for their upstream and downstream nodes. Any of those pipes can be the entrance for a DMA or should be closed to generate an effective isolation. Temporarily, let us assume that a valve or control device is installed in each pipe $e \in E'$, with diameter D_e.

The choice of which pipe will be an entrance and which pipe will be closed can be translated into an optimization process, once the cost of the control device is linked to the diameter of the pipe e. The cost C_{cd} associated with the control devices can be written as a function of diameter as

$$C_{\mathrm{cd}} = \sum_{e=1}^{|E'|} c\left(D_e\right), \tag{4}$$

where D_e is the diameter of the control device installed in the boundary pipe e with a cost $c(D_e)$, for a DMA scenario with a number of $|E'|$ potential control devices.

The decision variables in this optimization level are the existence or not of control devices in each boundary pipe. This can be stated as a binary optimization problem, where the value 1 represents the existence of a control device, and 0 represents the closure of the boundary pipe.

Furthermore, the use of PRVs as control devices at DMA entrances can improve the pressure management and, consequently, reduce leakage inside the DMA. This is possible because the operating point of a PRV, also called set point, which corresponds to the pressure of the upstream node of the valve, can be defined according to the minimum pressure required in the DMA. In this sense, the choice of the PRV set point can also be written as an optimization problem, trying to operate the water distribution network with as lower pressure as possible. Pressure uniformity can be a useful indicator to find the optimal set point. Reference [22] uses this criterion to achieve network optimal designs. In this paper, as proposed in [23], this indicator is calculated by

$$\mathrm{PU} = \sum_{j=1}^{T} \left[\frac{1}{N_n} \cdot \sum_{i=1}^{N_n} \left(\frac{P_{i,j} - P_{\mathrm{req}_{i,j}}}{P_{\mathrm{req}_{i,j}}} \right) + \frac{\sqrt{\sum_{i=1}^{N_n} \left(P_{i,j} - P_{\mathrm{av}_j}\right)^2 / N_n}}{P_{\mathrm{av}_j}} \right], \tag{5}$$

where T is the simulation period, N_n is the number of demand nodes in the network, $P_{i,j}$ is the pressure at node i and time j; $P_{\mathrm{req}_{i,j}}$ is the pressure required at node i and time j, and P_{av_j} is the network average pressure at time j. It should be observed, to avoid confusion, that the PU indicator does not represent "pressure uniformity," as it appears. It is, in fact, a general measure of the lack of uniformity of the distribution network pressures relative to the desired pressure. So the minimization of (5) is the objective function for the set point optimization level, since this improves pressure uniformity, thus allowing the network to operate nearby the minimal required pressure and without large pressure differences in the network.

2.2.2. Optimization Algorithm and Constraints.

2.2.2. Optimization Algorithm and Constraints. Reference [21] presents a clear proof that valve placement in water distribution systems is an NP-hard problem when the network is not a line, a loop, or a tree. To solve NP-hard problems, [24] suggests the use of heuristic algorithms, which can help in the treatment of these problems.

The use of bioinspired algorithms to solve water distribution problems has been a common and successful approach [25–29], mainly because of the easy implementation and near-global optimal solution ability exhibited by those algorithms, which do not need the calculation of Jacobian or Hessian matrices.

Particle swarm optimization (PSO) [30] has been widely applied for hydraulic problems such as optimal design of water distribution networks [31, 32], optimal design of wastewater networks [33], calibration of water supply networks [34], optimal pump operation [35] and is applied in this work to the optimal PRV placement and set point definition. A PSO swarm consists of a set of particles, which have two associated vectors: position and velocity. Usually, each vector starts randomly inside a defined range. The position vector is interpreted as a solution to the problem and allows the objective function evaluation. Position and velocity are iteratively updated, according to the following equations:

$$x_s^{t+1} = x_s^t + v_s^{t+1}, \tag{6}$$

$$v_s^{t+1} = w^t \cdot v_s^t + c_1 \cdot r_1 \left| l_b^t - x_s^t \right| + c_2 \cdot r_2 \left| g_b^t - x_s^t \right|, \tag{7}$$

where x_s^{t+1} is the position of particle s at iteration $t+1$, updated using the velocity v_s^{t+1}.

Velocity updating is a combination of

(i) the last velocity value, weighted by the inertia parameter w^t, to avoid excessive particle roaming;

(ii) the difference between the best position of particle s, l_b^t, and its actual position x_s^t weighed by the cognitive parameter c_1;

(iii) finally, the difference between the best position of the particle and the position of the swarm leader, g_b^t, weighed by the social parameter c_2.

The random numbers r_1 and r_2, working as particle scatters, avoid premature convergence to local optimal points. The last two summing elements in (7) are responsible for the convergence of the method because they attract particles to the best individual point a particle has visited and the best global point found by the entire swarm.

However, most of these algorithms work with unconstrained problems and require penalty functions to treat real constraints. In this case, for valve placement, the minimal pressure is considered as a constraint, while for the valve set point, the constraints are both minimum pressure and tank levels. The penalty functions and the final objective functions are presented in

$$\mathrm{Pen}_p = \sum_{j=1}^{T}\sum_{i=1}^{N_n}\left(\left|P_{i,j} - P_{\mathrm{req}_{i,j}}\right|\right), \tag{8}$$

$$\mathrm{Pen}_{tk} = \sum_{i=1}^{N_{tk}}\left(\left|L_{i,0} - L_{i,T}\right|\right), \tag{9}$$

$$\mathrm{OF}_1 = C_{\mathrm{cd}} \cdot \mathrm{PU} \cdot \mathrm{Pen}_p, \tag{10}$$

$$\mathrm{OF}_2 = \mathrm{PU} \cdot \mathrm{Pen}_p \cdot \mathrm{Pen}_{tk}. \tag{11}$$

Here Pen_p penalizes the violation of the required pressure; Pen_{tk} is the penalty value when the final tank level $L_{i,T}$ is lower than the initial tank level $L_{i,0}$ for tank i in a network with N_{tk} tanks. OF_1 is the objective function corresponding to the control device placement, and OF_2 is the objective function corresponding to the set point adjustment.

2.2.3. Multilevel Optimization for Valve Placement. Complex, real optimization problems typically need to be described and solved by multiobjective methods. While single objective approaches for real problems must contemplate the presence of constraints, multiobjective problems can treat these constraints as new objectives to be fulfilled. Furthermore, many processes involve a hierarchical decision process, so that the solution of a single optimization can be used to solve, partially, another process or determine some constraints of another process [36].

If on the one hand, the use of multiobjective process can also solve multilevel problems, on the other hand the open final solution given by the Pareto front keeps the stakeholder at the core of the final solution. In this way, the use of multilevel optimization can be useful for real, complex problems, eventually producing a single solution.

A general multilevel optimization problem P with decision variables x_k, objective functions, f_k, and constraints, g_k can be described as follows.

Find the solution for the upper level problem

$$\begin{aligned} P(1): \min_{x_1} \quad & f_1(x_1, x_2, \dots, x_k), \\ \text{s.t.} \quad & g_1(x_1, x_2, \dots, x_k) \le 0, \end{aligned} \tag{12}$$

where x_1 is the solution for the second-level problem:

$$\begin{aligned} P(2): \min_{x_2} \quad & f_2(x_1, x_2, \dots, x_k), \\ \text{s.t.} \quad & g_2(x_1, x_2, \dots, x_k) \le 0, \end{aligned} \tag{13}$$

and so on, until the lower level problem k, with solution x_k.

In the case of optimal valve placement, an adaption of the multilevel optimization concept is applied to reduce the complexity of the optimization process. Reference [35] presents two groups of objectives that are minimized in the DMA creation process based on a multiobjective approach. The first group corresponds to structural costs, related to the installation of valves, while the second group is related to hydraulic aspects, such as minimum pressure and maximum resilience.

In this line, the first level corresponds to the optimal boundary pipe selection, by minimizing (10). The solution of this level reduces the number of valves for the set points determined in the second level of the optimization. In this way, the treatment of the problem in two stages can lead to a faster solution, mainly for large networks.

Figure 1 presents a flowchart to illustrate the entire process of DMA creation and optimal location of control devices, further than the optimal set point definition.

2.3. Performance Indicators

2.3.1. Resilience Index. According to [37], in a looped network, a change in water flow due to pipe failure or demand increase will increase the internal energy losses. If nodes are being supplied with exactly the necessary pressure, in a new scenario it will be impossible to deliver water with satisfactory pressure and flow. Therefore, it is desired to provide each node with an energy surplus to guarantee the supply in critical conditions. Reference [18] defined the resilience index (14), which represents the network capability to overcome sudden failures:

$$I_r = \frac{\sum_{i=1}^{N_n} q_i \cdot \left(h_i - h_{\mathrm{req}_i}\right)}{\left(\sum_{k=1}^{N_r} Q_k \cdot H_k + \sum_{j=1}^{N_p} Q_j \cdot H_j\right) - \sum_{i=1}^{N_n} q_i \cdot h_{\mathrm{req}_i}}. \tag{14}$$

Here I_r is the resilience index; N_n is the number of demand nodes in the network; q_i is the demand at node i; h_i is the available head at node i; h_{req_i} is the required head at node i; N_r is the number of reservoirs and tanks in the network; Q_k is the flow supplied by reservoir/tank k; H_k is the head of reservoir/tank k; N_p is the number of pumps in the network; Q_j is the flow supplied by pump j; and H_j is the head of pump j. Due to tank level variations, the resilience index was calculated for each time step, and the minimum value observed was retrieved, since it represents the critical condition observed.

2.3.2. Demand Similarity. The demand similarity is important to evaluate the DMAs' size and mutual affinity, since it indicates how close a DMA demand is to the mean demand. Close values for demand similarity reflect more uniform DMAs, which help the WDS operation. Reference [38] highlights the use of demand data for real-time management, which could be simplified with similar DMAs. This indicator is calculated by the standard deviation of the total demand of each DMA (15). High values indicate poor sectorization, since nodes with low consumption could be aggregated into bigger DMAs, thus reducing costs associated with PRVs. Demand similarity is defined by

$$\mathrm{DS} = \frac{\sqrt{\sum_{i=1}^{N_n} \left(q_i - q_{\mathrm{av}}\right)^2 / N_n}}{q_{\mathrm{av}}}, \tag{15}$$

where q_{av} is the average demand in the network.

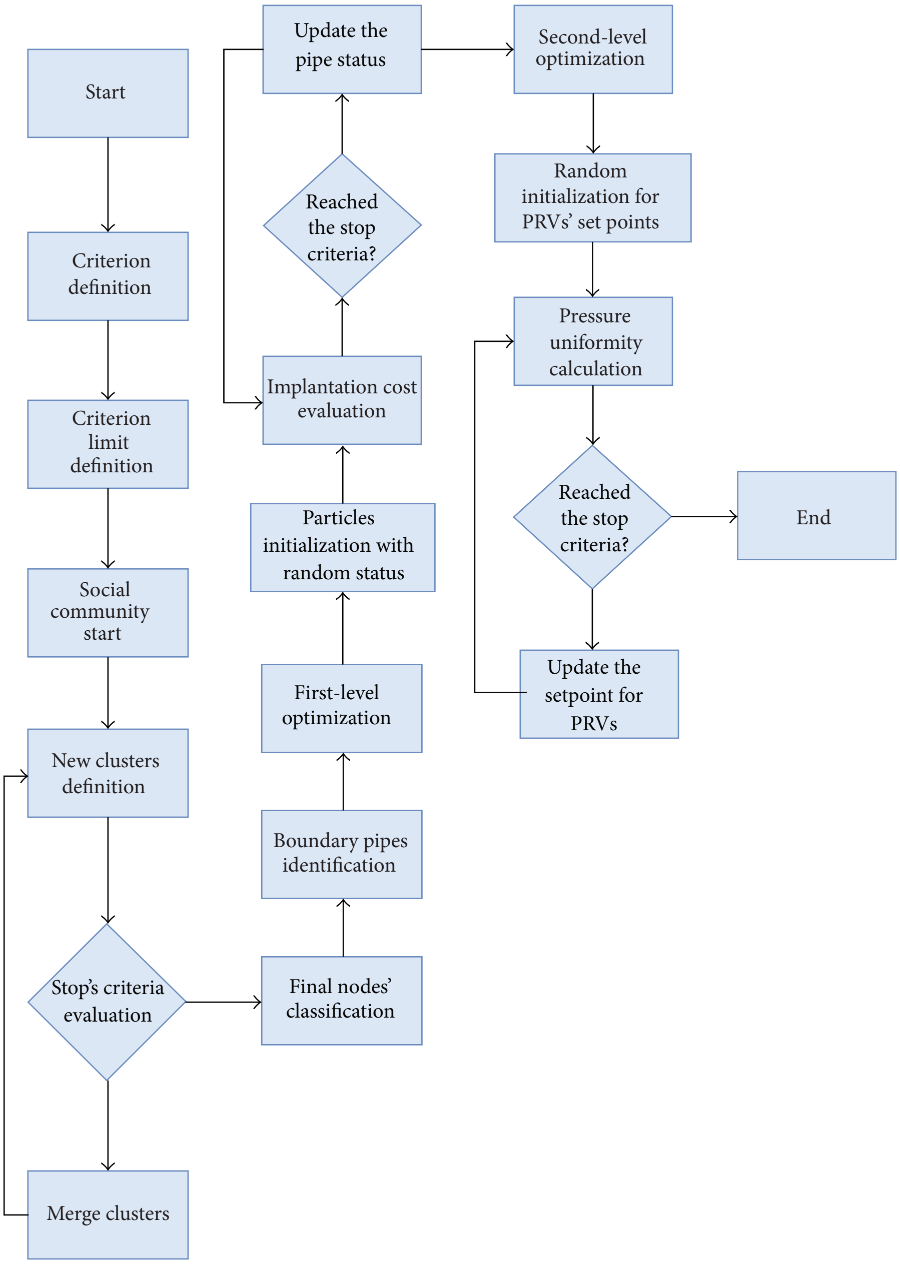

Figure 1: Flowchart of the automatic DMA creation with multilevel optimization.

2.3.3. Water Quality. Water quality is evaluated through the water age analysis. According to [39], chemical, biological, and physical problems can be caused or worsened by high detention times in the distribution system, which can cause health problems. The indicator proposed by [27] is the most used to evaluate the quality performance of the system:

$$I_{\mathrm{WA}} = \frac{\sum_{i=1}^{N_n} \sum_{j=1}^{T} k_{i,j} \cdot \left(\mathrm{WA}_{i,j} - \mathrm{WA}_{\mathrm{lim}_{i,j}}\right)}{\sum_{i=1}^{N_n} \sum_{j=1}^{T} q_{i,j}}, \qquad (16)$$

where I_{WA} is the water quality index, $\mathrm{WA}_{i,j}$ is the water age at node i, at time j, $\mathrm{WA}_{\mathrm{lim}_{i,j}}$ is the water age limit at node i, at time j, and $k_{i,j}$ is a coefficient equal to 1 if $\mathrm{WA}_{i,j} > \mathrm{WA}_{\mathrm{lim}_{i,j}}$ and to 0 if $\mathrm{WA}_{i,j} \leq \mathrm{WA}_{\mathrm{lim}_{i,j}}$.

3. Case Study

The methodology proposed in this work is applied to the D-town network, presented by [27] in the Battle of Water Networks II (BWN-II), with the final topology presented by [40]. The BWN-II presented the goal of identifying long-term improvements and operational strategies, given specific projected future demand and development of new areas. The final solution for this topology consists of 398 nodes, 458 pipes, 7 tanks, 1 reservoir, and 13 pumps, and the solution

Table 1: Performance indicators for the original topology of C-town proposed by [40].

Sectorization criteria	Resilience	Pressure uniformity	Demand similarity	Water quality	Energy	Costs	Number of DMAs
Original	0	468	78.8	0.280	76,495	—	5

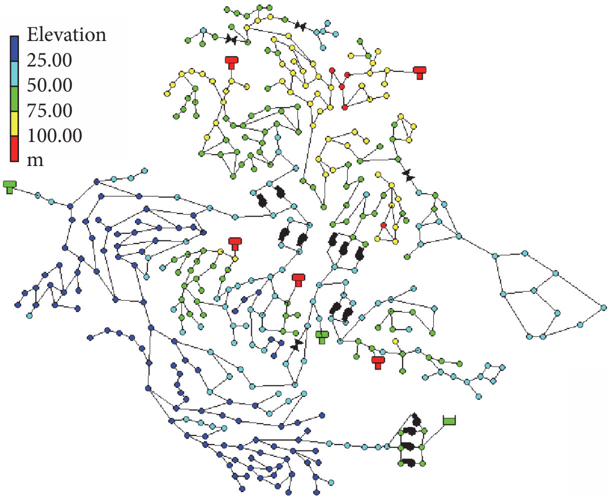

Figure 2: Topology of the D-town network with the solution proposed by [40].

presents 5 previously defined DMAs, segregated by 4 PRVs. Figure 1 shows the water distribution topology with a color scheme to represent the nodal elevation. The PRVs installed in the original case study do not have the objective to create DMAs, but to manage pressure zones for leakage control purposes. Comparisons among the new scenarios and the original network allow us to verify the hydraulic changes of the network under the presence of new DMA structures.

Following the operational constraints proposed by [27], the minimal pressure at the demand nodes should be above 25 m, and the water age limit for the nodes 48 hours. All simulations are done for the horizon of a week (168 hours), observing the energy price fluctuation during a typical week. To create the new DMAs, the existent valves are considered open, thus simulating a pipe in the network. Demand, length, and elevation upper bounds are used to create the set of DMA scenarios. Table 1 presents the original values for each indicator, considering the topology presented by [40]. When pressure at any demand node is under the minimum required, the resilience index calculated by (14) results in a negative value. However, taking into account the meaning of the resilience index, negative values do not make sense. To be coherent, the resilience index is taken as zero for every scenario in which the minimum pressure is not reached.

4. Results and Discussion

Taking into account the original DMAs of the network, the upper bound values for the criteria are selected to generate a similar number of DMAs. The map in Figure 2 shows the final DMA configuration considering the demand limitation criterion. Table 2 presents the relative performance indicators for demand criterion, using as reference the original value of the network. It is possible to observe, as expected, the merging of DMAs while the demand criterion increases (see also Figure 3).

The close relationship between the resilience index and pressure uniformity is observed in Table 2. The lower the pressure uniformity, the lower the value of the resilience index. This happens because the resilience index can be interpreted as the pressure surplus in the network. Moreover, pressure uniformity explains the deviation of the pressure from the minimal value. In this sense, the lower the pressure uniformity, the better the satisfaction of the minimal pressure in the network. However, the more adjusted the satisfaction of the minimal pressure in the network, the lower the resilience. We disregard here the original network, which presents a resilience index of zero. The increase of the number of DMAs allows better management of the pressure, as observed in the reduction of the pressure uniformity indicator. For this indicator, it is possible to observe an improvement for all DMA configurations, when compared with the original scenario, since the resilience index is greater than zero and pressure uniformity is reduced.

An improvement of the demand similarity index is observed for all scenarios with respect to the original topology. However, the demand criterion leads to an interesting behavior of the demand similarity. The increase of the DMA number improves the demand similarity until a certain limit. After this limit, in this case 8 DMAs, the values worsen. The decrease of the demand criterion leads to an increase of the number of DMAs. It is possible to obtain higher variability of the total demand per DMA, but harming the demand similarity indicator.

Since DMA creation requires the closure of boundary pipes, thus generating a preferred way to deliver the water, the age of water increases as observed in Table 2. In all the new scenarios with more DMAs than the original network, the water quality is harmed.

Even if the main objective of DMA creation in our study is to improve the pressure management, the energy consumption by pump stations can be an interesting indicator, since an extreme increase of energy consumption can turn a solution infeasible. With the demand criterion, only the scenario with lower number of DMAs has a slight increase of energy consumption, while in all the other scenarios this consumption decreases.

By comparing the original network with the scenarios of similar number of DMAs (demand limit 100 l/s and 120 l/s), it is possible to observe an important improvement of the demand similarity and pressure uniformity indicators with an energy consumption reduction in the scenario with 6 DMAs.

The second criterion evaluated in this work is the elevation. When nodes are grouped in the same DMA because of their similar elevation, a better pressure uniformity indicator

Table 2: Performance values for each DMA configuration with demand as criterion.

Sectorization criteria	I_r	PU	DS	I_{WA}	Energy	Costs	Number of DMAs	Number of devices
Demand								
40	0.272	0.679	0.670	2.869	0.944	42730	16	18
60	0.258	0.677	0.453	4.774	0.926	22859	10	13
80	0.282	0.777	0.257	3.878	0.937	27748	8	10
100	0.375	0.885	0.392	4.428	0.940	19653	6	7
120	0.316	0.899	0.355	1.319	1.016	19303	5	6

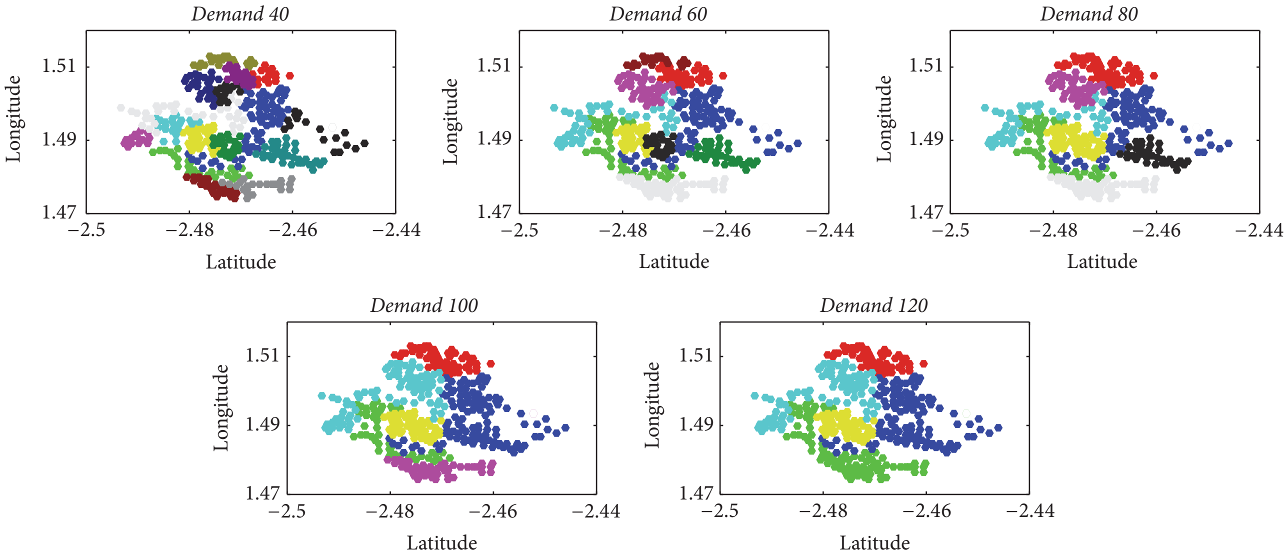

Figure 3: DMA scenarios using five different values for the demand criterion.

value is expected, since the pressure control can be more accurate, showing a DMA entry pressure nearby the minimal pressure required in the network. Table 3 shows the indicator parameters for the elevation criterion with five different scenarios of DMA configurations, and Figure 4 shows the maps with its respective final DMA configurations.

The pressure uniformity indicator is improved when compared with the original network. The high number of DMAs implies a greater number of PRVs installed, leading to better management of pressure. As a consequence, the resilience index is reduced, since the operation of the system is close to the minimal operating pressure.

Since the demand is not taken into account to generate the DMAs, the demand similarity index is worse than in the scenarios generated by the demand criterion. However, a similar behavior as when considering the demand criterion is observed. With the increase of the DMA number, the demand similarity is impaired.

In terms of water quality, as observed under the demand criterion, DMA scenarios generated by the elevation criterion worsen water quality, mainly because of the generation of preferred paths, which lead water to spend more time in the pipes before reaching the consumers.

Finally, in terms of energy and implementation costs, the elevation criterion resulted in scenarios with energy consumption close to the original network, but with better resilience and pressure uniformity index values. In comparison to the demand criterion scenarios, implementation costs are lower, which is related to the fact that the diameters of the pipes selected as entrances are lower than in the case that demand is used as a criterion.

Taking similar DMA scenarios (with 6 and 7 DMAs), the obtained configurations improve the hydraulic indicators (pressure uniformity, resilience, and demand similarity). The configuration with 6 DMAs even reduces slightly the energy consumption.

The last criterion applied in this work is the sum of pipe's length within the DMA. This criterion allows managing the size of DMAs, generating not only a similarity of elevation but also a similarity of demand, depending on the uniformity of consumers and the topography. Figure 5 shows the maps generated by the length criterion, and Table 4 presents the performance indicators for these scenarios.

Pressure uniformity and resilience indexes are improved when compared with the original scenario. Also the demand similarity is improved for all scenarios using the length criterion.

An intermediate improvement of demand similarity when the length criterion is used is observed, when comparing with the other two criteria. However, demand similarity is further improved by the demand criterion, as expected, due to the nature of the limits for DMA creation.

TABLE 3: Performance values for five DMA configurations using elevation as a criterion.

Sectorization criteria	I_r	PU	DS	I_{WA}	Energy	Costs	Number of DMAs	Number of devices
Elevation								
5	0.311	0.611	1.095	3.387	1.009	50535	19	25
10	0.340	0.708	1.092	4.349	0.977	14846	11	14
15	0.377	0.701	1.131	4.063	0.982	14846	9	12
20	0.435	0.944	0.923	2.278	1.032	11464	7	9
25	0.405	0.761	0.939	3.974	0.995	10454	6	7

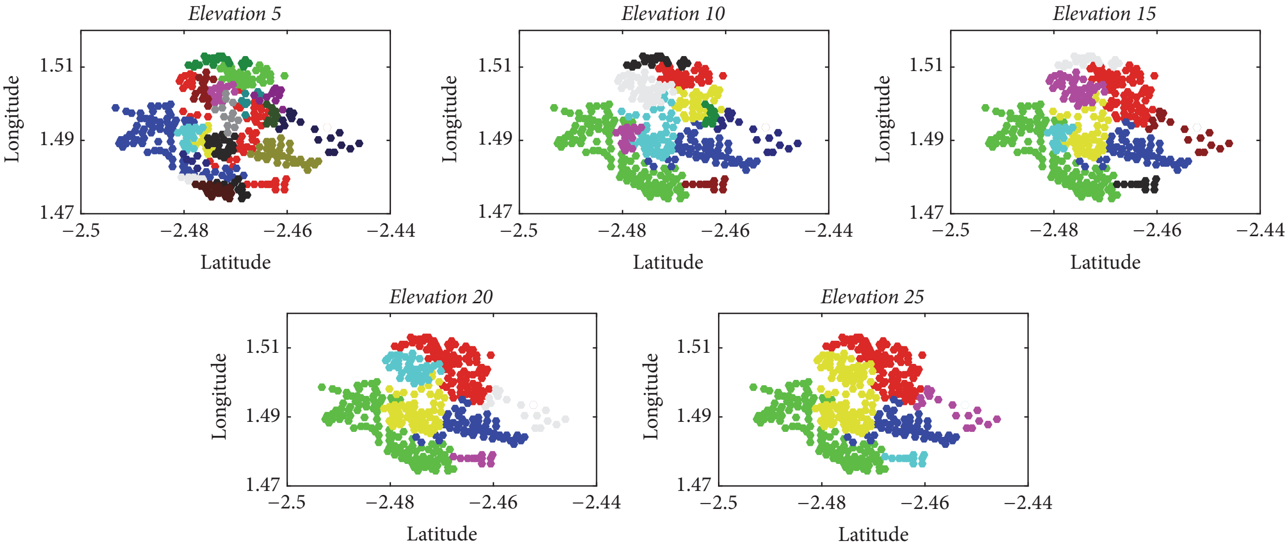

FIGURE 4: DMA scenarios using five different values for the elevation criterion.

Energy consumption is improved in all scenarios for the length criterion, when compared with the original network. Water quality is impaired when compared with the original scenario. Comparing with the elevation criterion, the implementation costs for the length criterion are larger, and a similar behavior is observed for the demand criterion.

5. Conclusions

DMA design for water distribution networks can help the management of WDSs, allowing key aspects in specifics areas to be addressed, for example, better pressure regulation. With small DMAs, pressure control is more accurate and pressure can achieve for many nodes the minimum required pressure. Considering the reality of many WDSs where high values of leakage are observed, mainly in developing countries, the operation of the systems under minimal pressure can reduce leakage, since leakage is strongly linked to operational pressure. Furthermore, the hydraulic balance into the individual DMAs, comparing the volume of water measured by flow meters at the entrances of the DMAs and the volumes of consumed water, can help water utilities to know the real state of the hydraulic efficiency of the system and, consequently, to develop better strategies to improve their quality. However, the task of dividing a large network into DMAs is not simple, since a set of performance indicators can greatly affect such a division.

Different criteria lead to different DMA scenarios, including worse scenarios than that of the original network with no segregation. The evaluation of hydraulic indicators must be considered a key analysis to guarantee the improvement of the water network after DMA generation.

Length and elevation criteria are presented as more adequate for pressure management purposes, generating a final scenario with lower implementation costs, energy consumption reduction, and good resilience index.

DMA creation using the demand as a criterion shows the best behavior in terms of demand similarity, but the worst incidence in terms of pressure uniformity. The elevation criterion shows an intermediate performance with good values for pressure uniformity and the resilience index.

The use of social community detection associated with a multilevel PSO optimization is a powerful tool to generate several DMA scenarios, allowing an insightful comparison among them.

The breakup of a multiobjective problem into levels can ease the decision making process, although the computational time in the multilevel approach may be longer than in the multiobjective approach.

As future works, we point towards the implementation of fully isolated DMAs by trunk network identification.

TABLE 4: Performance values for each DMA configuration using DMA length as a criterion.

Sectorization criteria	I_r	PU	DS	I_{WA}	Energy	Costs	Number of DMAs	Number of devices
Longitude								
5000	0.342	0.733	0.581	4.978	0.921	50928	17	20
7500	0.312	0.676	0.659	5.185	0.910	38837	14	16
10000	0.262	0.822	0.568	2.513	0.989	21849	9	12
12500	0.256	0.817	0.568	3.373	0.957	21849	9	11
15000	0.330	0.839	0.512	4.418	0.949	19653	7	9

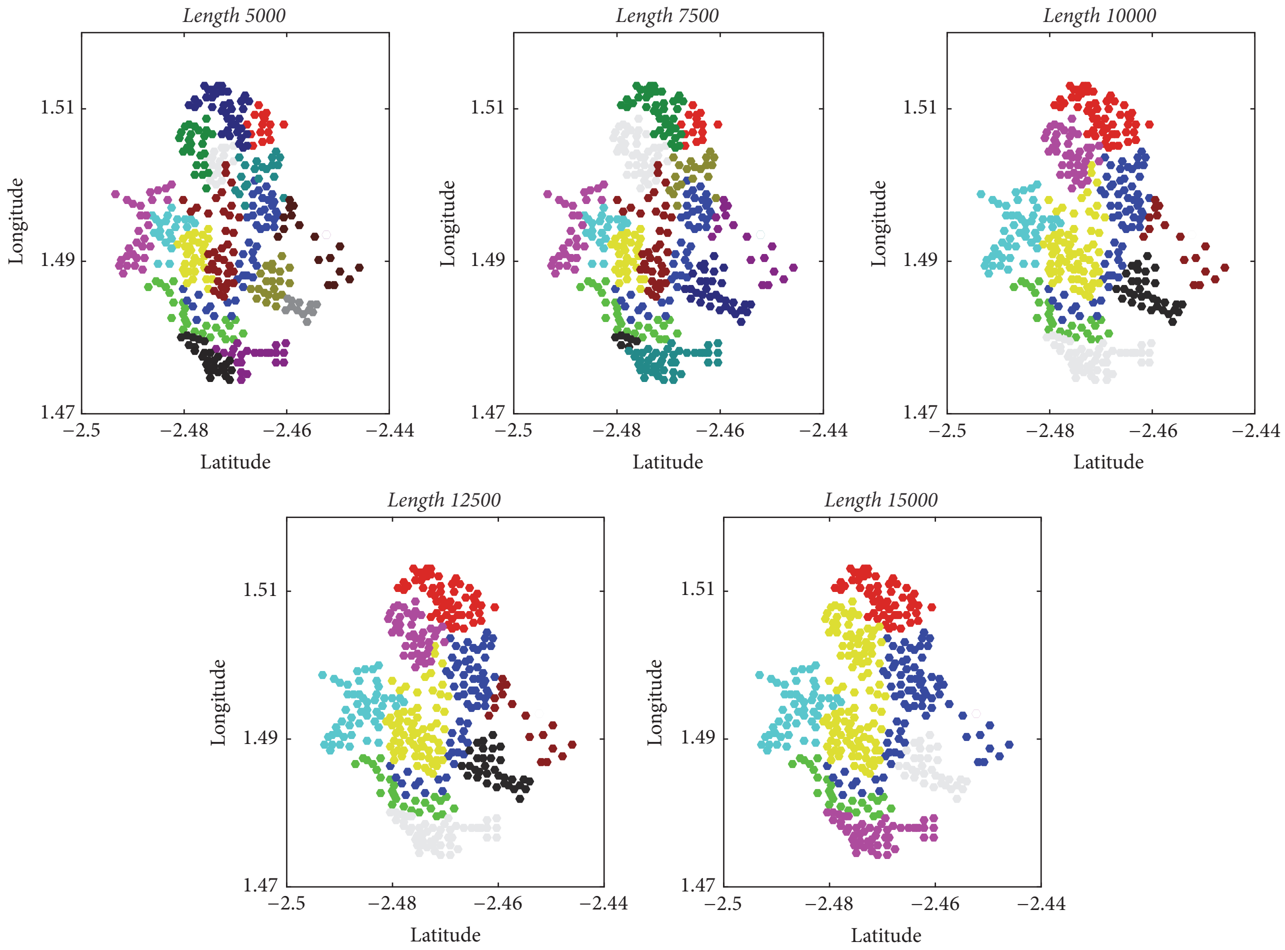

FIGURE 5: DMA scenarios using five different values for the length criterion.

Additionally, a deeper criteria analysis, mainly in terms of the computation effort in the optimization process, is also deemed of interest.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors wish to express a special gratitude and remembrance towards Professor Rafael Pérez García, who died recently.

References

[1] A. Di Nardo, M. Di Natale, G. F. Santonastaso, and S. Venticinque, “An automated tool for smart water network partitioning,” *Water Resources Management*, vol. 27, no. 13, pp. 4493–4508, 2013.

[2] W. M. Grayman, R. Murray, and D. A. Savic, “Effects of redesign of water systems for security and water quality factors,” in *Proceedings of the World Environmental and Water Resources Congress*, vol. 49, pp. 1–11, Kansas City, Mo, USA, 2009.

[3] WRC/WSA/WCA Engineering and Operations Committee, “Managing Leakage: UK Water Industry Managing Leakage,” Report A-J. London WRC/WSA/WCA, 1994.

[4] E. Campbell, J. Izquierdo, I. Montalvo, and R. Pérez-García, “A novel water supply network sectorization methodology based

on a complete economic analysis, including uncertainties," *Water*, vol. 8, no. 5, article no. 179, 2016.

[5] A. di Nardo, M. D. Natale, G. F. Santonastaso, V. G. Tzatchkov, and V. H. Alcocer-Yamanaka, "Water network sectorization based on graph theory and energy performance indices," *Journal of Water Resources Planning and Management*, vol. 140, no. 5, pp. 620–629, 2014.

[6] S. Alvisi and M. Franchini, "A heuristic procedure for the automatic creation of district metered areas in water distribution systems," *Urban Water Journal*, vol. 11, no. 2, pp. 137–159, 2014.

[7] R. Gomes, A. Sá Marques, and J. Sousa, "Identification of the optimal entry points at District Metered Areas and implementation of pressure management," *Urban Water Journal*, vol. 9, no. 6, pp. 365–384, 2012.

[8] A. Di Nardo and M. Di Natale, "A heuristic design support methodology based on graph theory for district metering of water supply networks," *Engineering Optimization*, vol. 43, no. 2, pp. 193–211, 2011.

[9] V. G. Tzatchkov, V. H. Alcocer-Yamanaka, and V. Bourguett Ortíz, "Graph theory based algorithms for water distribution network sectorization projects," in *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium (WDSA '06)*, p. 172, ASCE, Cincinnati, Ohio, USA, August 2006.

[10] M. Herrera, S. Canu, A. Karatzoglou, R. Perez-García, and J. Izquierdo, "An approach to water supply clusters by semi-supervised learning," in *Proceedings of the 5th Biennial Conference of the International Environmental Modelling and Software Society: Modelling for Environment's Sake (iEMSs '10)*, pp. 1925–1932, Ottawa, Canada, July 2010.

[11] J. Izquierdo, M. Herrera, I. Montalvo, and R. Pérez-García, "Division of water supply systems into district metered areas using a multi-agent based approach," in *Software and Data Technologies (Communications in Computer and Information Science)*, chapter 13, pp. 167–180, Springer, Berlin, Germany, 2011.

[12] K. Diao, Y. Zhou, and W. auch, "Automated creation of district metered area boundaries in water distribution systems," *Journal of Water Resources Planning and Management*, vol. 139, no. 2, pp. 184–190, 2013.

[13] L. S. Perelman, M. Allen, A. Preis, M. Iqbal, and A. J. Whittle, "Automated sub-zoning of water distribution systems," *Environmental Modelling & Software*, vol. 65, pp. 1–14, 2015.

[14] E. Campbell, D. Ayala-Cabrera, J. Izquierdo, R. Pérez-García, and M. Tavera, "Water supply network sectorization based on social networks community detection algorithms," *Procedia Engineering*, vol. 89, pp. 1208–1215, 2014.

[15] M. Herrera, E. Abraham, and I. Stoianov, "A graph-theoretic framework for assessing the resilience of sectorised water distribution networks," *Water Resources Management*, vol. 30, no. 5, pp. 1685–1699, 2016.

[16] F. De Paola, N. Fontana, E. Galdiero, M. Giugni, D. Savic, and G. S. Degli Uberti, "Automatic multi-objective sectorization of a water distribution network," *Procedia Engineering*, vol. 89, pp. 1200–1207, 2014.

[17] D. A. Savic and G. Ferrari, "Design and performance of district metering areas in water distribution systems," *Procedia Engineering*, vol. 89, pp. 1136–1143, 2014.

[18] E. Todini, "Looped water distribution networks design using a resilience index based heuristic approach," *Urban Water*, vol. 2, no. 2, pp. 115–122, 2000.

[19] P. Pons and M. Latapy, "Computing communities in large networks using random walks," *Journal of Graph Algorithms and Applications*, vol. 10, no. 2, pp. 191–218, 2006.

[20] J. H. Ward Jr, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, pp. 236–244, 1963.

[21] H. L. Bodlaender, A. Hendriks, A. Grigoriev, and N. V. Grigorieva, "The valve location problem in simple network topologies," *INFORMS Journal on Computing*, vol. 22, no. 3, pp. 433–442, 2010.

[22] D. Araque and J. G. Saldarriaga, "Water distribution network operational optimization by maximizing the pressure uniformity at service nodes," in *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium (WDSA '06)*, pp. 1–14, Cincinnati, Ohio, USA, 2006.

[23] H. A. Alhimiary and R. H. Alsuhaily, "Minimizing leakage rates in water distribution networks through optimal valves settings," in *Proceedings of the World Environmental and Water Resources Congress*, pp. 1–13, Tampa, Fla, USA, May 2007.

[24] K. W. Tindell, A. Burns, and A. J. Wellings, "Allocating hard real-time tasks: an NP-Hard problem made easy," *Real-Time Systems*, vol. 4, no. 2, pp. 145–165, 1992.

[25] I. Montalvo, J. Izquierdo, R. Pérez-García, and M. Herrera, "Water distribution system computer-aided design by agent swarm optimization," *Computer-Aided Civil and Infrastructure Engineering*, vol. 29, no. 6, pp. 433–448, 2014.

[26] A. Vassiljev and R. Puust, "Decreasing leakage and operational cost for BBLAWN," *Journal of Water Resources Planning and Management*, vol. 142, no. 5, pp. 1–7, 2016.

[27] A. Marchi, E. Salomons, A. Ostfeld, and Z. Kapelan, "Battle of the water networks II," *Journal of Water Resources Planning and Management*, vol. 140, no. 7, pp. 1–14, 2014.

[28] M. Nicolini and L. Zovatto, "Optimal location and control of pressure reducing valves in water networks," *Journal of Water Resources Planning and Management*, vol. 135, no. 3, pp. 178–187, 2009.

[29] B. A. Tolson, H. R. Maier, A. R. Simpson, and B. J. Lence, "Genetic algorithms for reliability-based optimization of water distribution systems," *Journal of Water Resources Planning and Management*, vol. 130, no. 1, pp. 63–72, 2004.

[30] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the 6th International Symposium on Micro Machine and Human Science (MHS '95)*, vol. 1, pp. 39–43, 1995.

[31] C. R. Suribabu and T. R. Neelakantan, "Design of water distribution networks using particle swarm optimization," *Urban Water Journal*, vol. 3, no. 2, pp. 111–120, 2006.

[32] I. Montalvo, J. Izquierdo, R. Pérez, and M. M. Tung, "Particle swarm optimization applied to the design of water supply systems," *Computers and Mathematics with Applications*, vol. 56, no. 3, pp. 769–776, 2008.

[33] J. Izquierdo, I. Montalvo, R. Pérez, and M. Tavera, "Optimization in water systems: a PSO approach," in *Proceedings of the 2008 Spring Simulation Multiconference (SpringSim '08)*, pp. 239–246, ACM/SIGSIM, Ottawa, Canada, April 2008.

[34] J. Izquierdo, I. Montalvo, R. Pérez, and V. S. Fuertes, "Design optimization of wastewater collection networks by PSO," *Computers and Mathematics with Applications*, vol. 56, no. 3, pp. 777–784, 2008.

[35] B. Brentan, E. Luvizotto Jr., I. Montalvo, J. Izquierdo, and R. Pérez-García, "Water pump scheduling optimization using

agent swarm optimization," in *Proceedings of the International Conference on Mathematical Modelling in Engineering & Human Behaviour 2015*, pp. 50–55, Polytechnic University of Valencia, Valencia, Spain, 2015.

[36] A. Migdalas, P. M. Pardalos, and P. Värbrand, *Multilevel Optimization: Algorithms and Applications*, vol. 20, Springer Science & Business Media, Berlin, Germany, 1998.

[37] E. Creaco, A. Fortunato, M. Franchini, and M. R. Mazzola, "Comparison between entropy and resilience as indirect measures of reliability in the framework of water distribution network design," in *Proceedings of the 12th International Conference on Computing and Control for the Water Industry (CCWI '13)*, vol. 70, pp. 379–388, Perugia, Italy, 2013.

[38] S. Alvisi and M. Franchini, "A procedure for the design of district metered areas in water distribution systems," in *Proceedings of the 12th International Conference on Computing and Control for the Water Industry (CCWI '13)*, vol. 70, pp. 41–50, Perugia, Italy, September 2013.

[39] US Environmental Protection Agency—EPA, "Effects of water age on distribution system water quality," Tech. Rep., US Environmental Protection Agency—EPA, 2002.

[40] C. Stokes, W. Wu, and G. Dandy, "Battle of the water networks II: combining engineering judgement with genetic algorithm optimization," in *Proceedings of the 14th Water Distribution Systems Analysis Conference*, pp. 77–89, Adelaide, Australia, September 2012.