

The Performance of Spectral Clustering Algorithms on Water Distribution Networks: Further Evidence

Farid Belloum

Laboratory of Civil and Hydraulic Engineering
Civil and Hydraulic Engineering Department
University of Guelma
Guelma, Algeria
belloum.farid@univ-guelma.dz

Larbi Houichi

Department of Hydraulic Engineering
Faculty of Technology
University of Batna 2
Batna, Algeria
l.houichi@univ-batna2.dz

Mazouz Kherouf

Laboratory of Civil and Hydraulic Engineering
Civil and Hydraulic Engineering Department
University of Guelma
Guelma, Algeria
Kherouf_maz@yahoo.fr

Received: 2 June 2022 | Revised: 15 June 2022 | Accepted: 17 June 2022

Abstract-The aim of the current paper is to revisit the performance of spectral clustering algorithms for water distribution networks. In the literature, there have been attempts to introduce improved algorithms based on graph theory. We focus on a class of these algorithms that applies the concepts of the spectral clustering approach. We assess the performance of spectral clustering algorithms on a wider range of water network types (i.e. large, medium, and small sized networks) using a wider range of clustering methods (both partitioning and hierarchical) and performance indicators. Our findings suggest that partitioning methods, such as k-means are not consistently efficient in all types of networks. Nonetheless, the Partitioning Around Medoids (PAM) algorithm shows a relatively good performance according to modularity, while the internal indices of k-means and hierarchical clustering algorithms are more efficient. Stability indices show that PAM and CLARA algorithms are more efficient.

Keywords-sectorization; clustering methods; spectral clustering algorithms

I. INTRODUCTION

Water Distribution Networks (WDNs) are large and complex networks designed and built to serve the water supply needs of an urban area. Proper and efficient control and management of WDNs is one of the key aspects in the operational practices of water supply networks and a major challenge [1, 2]. This is particularly true for large urban areas and cities, where the network becomes larger and more complex. In order to manage and monitor unusual demands - caused generally by leaks - the common practice involves using District Metered Areas (DMAs) [3, 4], which allow water

network systems to monitor water inflows and outflows to reduce water and revenue losses. Thus, an optimal and efficient approach to optimally partition the water network into DMAs is of crucial importance. Although WDN is an efficient approach for the partitioning of the DMAs [5], achieving optimal partitioning remains challenging, due to the existence of many parameters that need to be accounted for, such as hundreds of valves, tanks, pipelines, and pumps [6-8].

Although many methods have been proposed, the literature does not offer a conclusive result on the best performing approach to cluster DMAs. For example, authors in [9] developed an algorithm based on graph theory, which divides the water distribution system into clusters according to the direction of the flow in the pipes. This algorithm is general enough to be applied for various purposes including the enhancement of water security by assigning sensors to clusters, and/ or the efficient isolation of a containment intrusion. Authors in [10] propose the use of pressure sensitivity matrix analysis combined with clustering techniques to optimally characterize and determine a sensor configuration that maximizes the degree of locatability subject to the budget constraint of sensor configuration. Furthermore, they show that their proposed method is successfully implemented to identify the optimal set of pressure sensors that need to be installed in a DMA in the Barcelona WDN. Authors in [36] propose a multistage approach that involves the simultaneous application of sectorization of a large WDN - into DMAs - and optimization of rehabilitation of the network with new components such as pipes, valves and tanks. Finally, authors in [34] conduct a comparative analysis of three commonly used

methods to create DMAs, Fast Greedy, Random Walk, and Metis. The performance of these methods is evaluated on two cases of a large and complex WDN EXNET with unweighted and weighted edges. They conclude that Fast Greedy has an overall better performance and was found to be more effective in weighted graph partitioning. In this context, water network partitioning has become one of the most popular methods and strategies adopted to improve the optimal WDN control and management, within which spectral graph theory is central tool for developing clustering algorithms. In particular, authors in [11] propose a holistic analysis framework to support water utilities for an efficient supply management. The proposed approach is based on graph spectral methods that employ the properties of eigenvalues and eigenvectors of matrices associated with graphs. The approach aims to reduce the challenges urban water supply may face in an increasingly complex city-based environment, which is characterized by various heterogeneous and interconnected infrastructures often affect the process of supplying households with safe water. The approach is tested on two WDNs, C-Town and an operational network. As a result, they show that their proposed approach is useful and optimal to efficiently manage WDNs.

Despite the attractive features of this holistic approach, the extent to which this approach is robust is not known. For example, the authors in [11] offer limited evidence that spectral based algorithms are useful and lack robustness checks. Thus, the main aim of this paper is to extend the analysis offered by [11] by testing the sensitivity and robustness of the graph spectral methods in two ways. First, we test these methods on three different types of WDNs: A complex and large network (EXNET), a medium-sized network (C-Town), and a small-sized operational network (OUED EL MA). These networks represent different types of water networks in terms of structure and size. Second, we assess the performance of the spectral based algorithms using the PAM, CLARA, and DIANA clustering methods. We apply these algorithms on the three types of networks described above. Then, the performance of each algorithm is assessed based on three criteria, namely modularity, internal indices, and stability indices. Our findings suggest that the performance of algorithms differs depending on the performance criterion and the type of network. In this context, according to modularity, the PAM algorithm is found to perform better. K-Means and hierarchical clustering are found to be the most efficient according to the internal indices, while PAM, hierarchical clustering, and CLARA have better stability index. Thus, one cannot generalize the approach of [11] as a generic method applicable to all types of networks.

II. MATERIALS AND METHODS

A. Spectral Clustering

A water network can be represented as a graph whose edges and nodes represent network pipes and pipe junctions, respectively [12, 13]. This graphical structure makes the spectral clustering approach suitable for solving water network problems. The spectral clustering approach has many attractive features. It is simple to implement and can be solved efficiently by standard linear algebra software [14-17]. In addition, it outperforms standard clustering techniques such as k-means [14]. This method is implemented in three steps, namely: (i)

computing a similarity graph, (ii) projecting data onto a low-dimensional space, and (iii) identifying clusters. Figure 1 illustrates the workflow of the spectral clustering algorithm highlighting further the computations involved throughout all the steps. This includes associating both adjacency and Laplacian matrices to graphs. Next, their eigenvalues are computed, which are then related to the structural properties of graphs [18, 19].

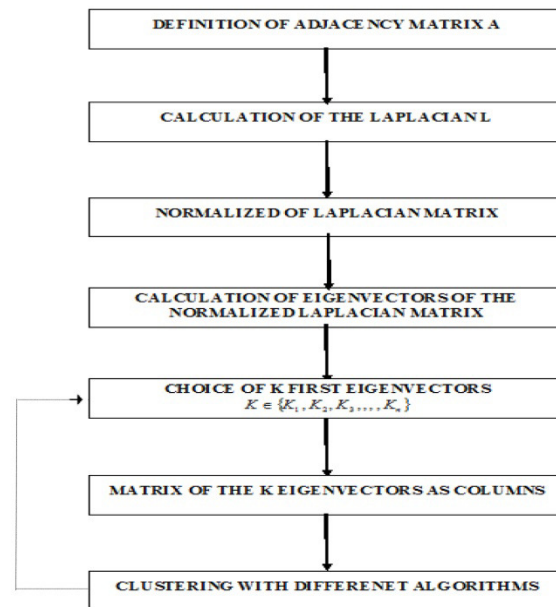


Fig. 1. Spectral clustering algorithm.

Finally, spectral k-means clustering algorithm is applied to extract the final clusters from the eigenvectors' matrix when all clusters are assumed to have equal size and homoscedastic. In most of the cases this assumption is not satisfied. Clusters will differ in their size, density, and variance [20]. Violating this assumption does not necessarily rule out the use of k-means algorithm. It simply implies one needs to correct the algorithm for the times this assumption is violated. In addition, and given the recent advances in computer processing power, using other alternative algorithms that relax the assumption of clusters' size and variance constancy is necessary to achieve robust results. Thus, in this paper, we conduct a comparative exercise using different clustering algorithms. This includes the following: Clustering Large Applications (CLARA), Partitioning Around Medoids (PAM), Agglomerative, Divisive Analysis (DIANA), and k-means. This setting will allow us to extract the best partitioning scheme according to internal quality and stability indices available in the literature.

B. Clustering Methods

There are two broader classes of clustering methods: (i) partitioning and (ii) hierarchical clustering methods. While partitioning methods divide the data into mutually exclusive partitions (i.e. clusters), hierarchical methods classify the data into multi-levels or a hierarchy of clusters. In general, these methods have key differences. Partitioning clustering methods are faster. They require, however, stronger assumptions such as

initial parameters and number of clusters. In contrast, hierarchical clustering methods require only a similarity measure and not initial parameters such as the number of clusters. Figure 2 illustrates the different methods used under each clustering class. In this paper, we discuss two types in each clustering method. This includes the k-means and k-medoids under partitioning methods and agglomerative and divisive under hierarchical.

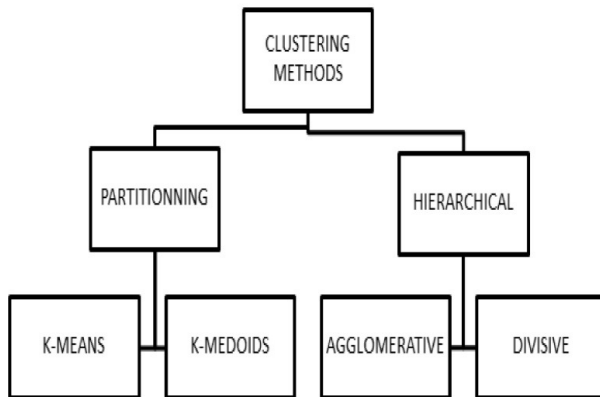


Fig. 2. Most popular clustering methods.

1) Partitioning Clustering

k-means algorithm is one of the oldest and most commonly used clustering algorithms. It attempts to partition the dataset into k predefined clusters where each observation belongs to the cluster with the nearest mean. Authors in [21] used k-means on large-scale networks to determine clusters in order to minimize the sum of squared distance error between nodes for developing DMAs configuration.

PAM is a variant of k-means in which instead of using the mean point as the center of a cluster. It uses an actual point in the cluster to represent it. Authors in [22, 23] combined PAM partitioning method with spectral clustering to introduce a greedy heuristic that reassigns the nodes belonging to minor connected components to major connected components of an electric power network.

CLARA is an extension to k-medoids (PAM). It is used to deal with data containing a large number of objects. Authors in [24] highlighted the advantages of using spectral clustering over CLARA and showed that spectral clustering has better performance to analytically localize leaks in a WDN. In addition, authors in [25] compared PSO algorithm to partitioning clustering performed by other algorithms that work with various attribute types (such as PAM and CLARA). The findings obtained by these algorithms were consistently similar. Authors in [25], however, show that PSO is more efficient.

2) Hierarchical Clustering

Agglomerative algorithms are based on graph theory concepts. They iteratively merge the nodes to build a bottom-up hierarchy. Authors in [26] presented an algorithm that uses a computationally efficient spectral and hierarchical clustering algorithm to reveal the internal connectivity structure of a network representing a power grid, subject to any choice of

electrical parameter associated to the transmission. Hierarchical spectral clustering has revealed a static internal structure of the network.

DIANA is a technique that constructs the hierarchy in the inverse order [27]. First, all the data points are considered as a single cluster. Then, using iterations, the dataset is divided into clusters until reaching the optimal n clusters in the last iteration.

3) Quality Measures of Graph Clustering

For robustness purposes, we apply various measures of quality. This includes the following:

Modularity (M): This measure is the most common quality measure for graph clustering [28]. It captures the trade-off between the number of edges bridging clusters and a measure accounting for the number and the similarity among each other [29]. Modularity can be defined as:

$$M = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (1)$$

where A_{ij} is the value of the adjacency matrix between vertices i and j , k_i is the sum of the weights of the edges adjacent to i , m is the number of edges in the graph, and δ is the Kronecker delta which takes value of 1 if its arguments are equal and 0 otherwise. Authors in [29] presented the classical modularity index and tailored and modified it for WDNs in order to optimize the framework based on the maximization of the WDN-oriented modularity-based index versus the minimization of the cost of newly installed devices to obtain network segments.

4) Internal Measures

Connectivity: It uses vertex distances to measure cohesion and separation of clusters. It is formally defined as [30, 31]:

$$\sum_{i=1}^N \sum_{j=1}^L \sum x_{i,nn_{i(j)}} \quad (2)$$

where L is a parameter that determines the number of neighbors that contribute to the connectivity measure, $nn_{i(j)}$ as the j is the nearest neighbor of observation i , $x_{i,nn_{i(j)}}$ is 0 if i and $nn_{i(j)}$ are in the same cluster and 1 otherwise. The connectivity value should be minimized.

Dunn (D): Dunn's index [32] shows the distance between objects in a cluster and its intercluster separation. It uses the minimum pairwise distance that each cluster has to reach to be compact. To calculate the Dunn index, we adopt [31]:

$$D = \frac{\min_{C_k, C_l \in C, C_k \neq C_l} (\min_{i \in C_k, j \in C_l} \text{dist}(i, j))}{\max_{C_m \in C} \text{diam}(C_m)} \quad (3)$$

where $\text{diam}(C_m)$ is the maximum distance between observations in cluster C_m . The Dunn index has a value between zero and 1 and should be maximized.

Silhouette (S): The silhouette value measures the degree of confidence in the clustering assignment of a particular

observation, with well-clustered observations having values near 1 and poorly clustered observations having values near 0 [31]. For observation i , S is formally defined as:

$$S(i) = \frac{b_i - a_i}{\max(b_i, a_i)} \quad (4)$$

where a_i is the average distance between i and all the other observations in the same cluster, and b_i is the average distance between I and the observation in the nearest neighboring cluster.

5) Stability Measures

Clustering is considered stable if the corresponding partitioning remains unchanged with its minimum change [33]. As a sub-category of internal indices, there are many stability measures that help validate cluster solutions available: Average Proportion of No overlap (APN), Average Distance (AD), Average Distance between Means (ADM), and Figure of Merit (FOM) [31].

III. WATER DISTRIBUTION NETWORKS

For robustness purposes, we assess the performance of spectral clustering algorithms on three different types of water networks. This is in contrast to [11] as it presents conclusions limited to medium sized networks. In this context, we consider 3 types of water networks, chosen according to their type (looped or mixed) as well as their size based on the number of nodes and edges (i.e. small, medium, and large).

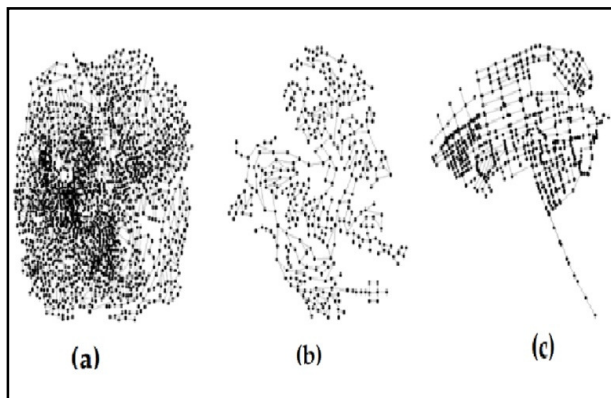


Fig. 3. WDNs: (a) EXNET, (b) C-Town, (c) Oued El Ma.

The first network is EXNET (Figure 3(a)), which is large-sized and complex. It serves a population of approximately 400,000 inhabitants. The network system consists of 2416 pipes and 1891 nodes. Elevated reservoirs provide the energy to the entire system [34]. The second network is the medium-sized and operational Oued El Ma WDN, as illustrated in Figure 3(c), which is located in an Algerian medium-sized city with a population of 14000. The water network consists of 621 nodes and 630 edges. C-Town network— as depicted in Figure 3(b) – is a small-sized WDN. C-Town is a benchmark water network model employed in the field of water distribution system analysis with 444 pipes and 396 nodes [35].

IV. CRITICAL DISCUSSION

Authors in [11] focused on the sectorization of the C-Town network for 5 sectors using the k-means algorithm. In this section, we assess the performance of the algorithm proposed in [11] and compare it to alternative algorithms including PAM, CLARA, hierarchical clustering, and DIANA using the same experimental settings. In other words, we use the same conditions including C-Town network with 5 sectors. Figure 4 illustrates the sectorization using the k-means and our proposed alternative clustering methods. In general, our replication is successful and consistent with the original work [11].

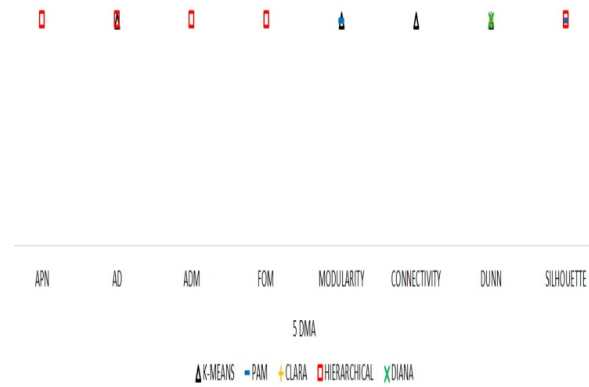


Fig. 4. Clusters of C-Town.

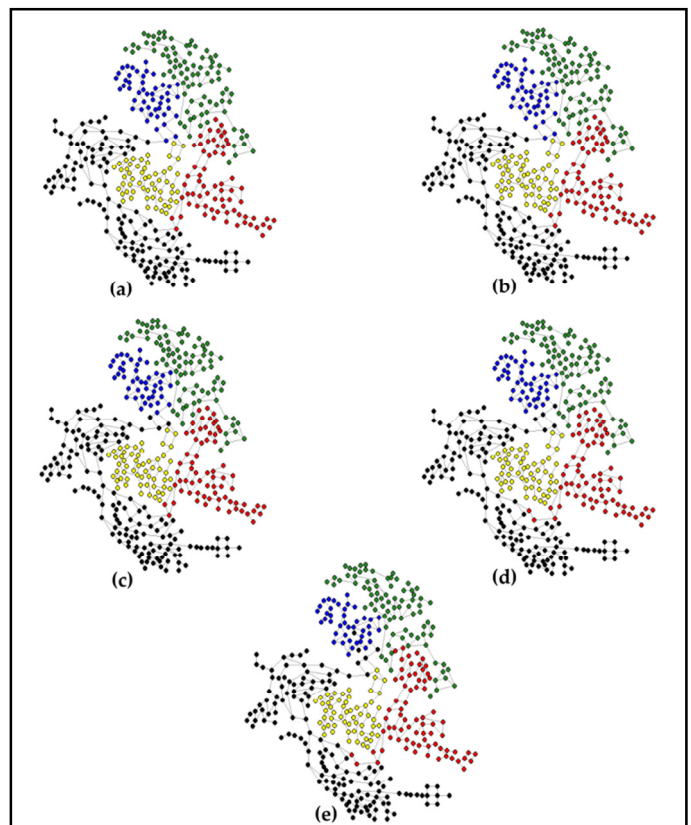


Fig. 5. C-Town network sectorized: (a) k-means (a), (b) PAM, (c) CLARA, (d) hierarchical, (e) DIANA.

Figure 5 reports the internal, stability, and modularity indices for the 5 algorithms on a C-Town WDN with 5 sectors. According to the findings, k-means clustering algorithm is reported to be stable by only 1 criterion (AD). It performs, however, better based on modularity and internal criteria including Connectivity and Dunn. Our findings suggest, that hierarchical clustering performs relatively well in terms of stability and internal criteria.

However, our evidence show that the k-means performs better only in a very special setting, i.e. for small-sized WDNs with a small number of sectors. For larger number of sectors and medium to large-sized networks k-means performs poorly compared to PAM and hierarchical clustering in general.

V. RESULTS AND DISCUSSION

In this section, we report and discuss the findings of our simulations. In this context, spectral partitioning was carried out for the three considered WDNs using the techniques discussed above. For each WDN, we report the identified clusters using the k-means, PAM, CLARA, general hierarchical clustering, and DIANA. Furthermore, for each network, the performances of clustering methods are evaluated on different number of clusters. Our experimental setting allows 6 different partitions of clusters as follows:

- Large water network allowing 10, 20, 25, 30, 40, and 50 partitions.
- Medium sized network allowing for 4, 6, 8, 10, and 12 partitions.
- Small-sized network: allowing for 3, 4, 5, 6, 8, and 10 partitions.

A. Large and Complex Network (EXNET)

Figure 6 illustrates the performance of clustering algorithms across 6 different partition combinations.

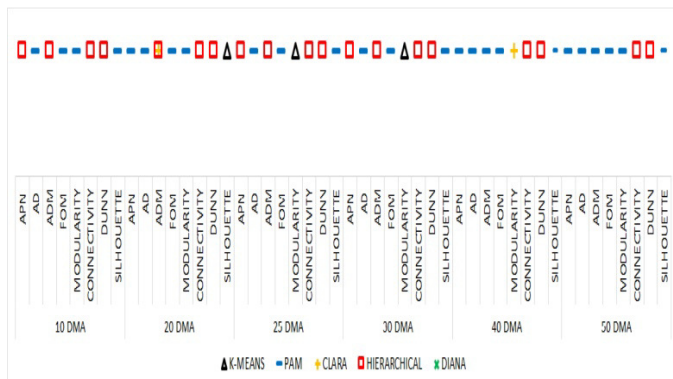


Fig. 6. EXNET WDN spectral clustering.

The number of clusters across these partitions ranges between 10 and 50. There are 8 performance criteria reported for each of the 5 clustering algorithms. According to these criteria, PAM has been found to perform consistently well across all 6 combinations of partitions. In particular, PAM is dominant as the number of clusters increases to 40 and 50. Hierarchical is found to perform relatively well, especially

when the number of clusters is lower than 40. Furthermore, the performance of hierarchical clustering methods is as good as that of PAM when the number of clusters is 10, 25 and 30. Based on the overall performance, PAM seems to meet all the criteria when the number of clusters is 50. In other words, PAM is found to be better in terms of internal, stability, and modularity performance.

B. Oued El Ma Water Distribution Network

Figure 7 reports the performance of the algorithms for a medium-sized WDN. According to the findings, the hierarchical clustering algorithms perform better when the number of clusters is smaller (3 to 6). In contrast, the overall performance of partitioning methods is better when the number of clusters increases as shown in the cases of 8, 10, and 12 clusters.

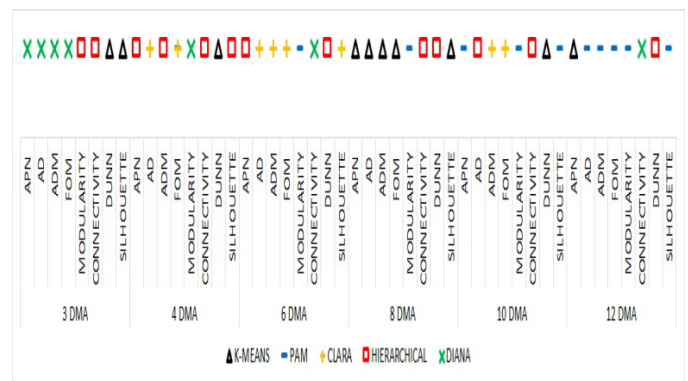


Fig. 7. Oued El Ma WDN spectral clustering.

C. C-Town Water Distribution Network

Figure 8 reports the performance of clustering algorithms across the 6 different partition combinations for a small-sized WDN. Similar to the medium-sized network, the findings suggest that partitioning-based clustering algorithms perform better than the hierarchical based methods when the number of clusters increases. For example, when the number of clusters is 10, PAM performs better than any other algorithm in terms of stability and modularity performance. Hierarchical based methods, however, perform, better when the number of clusters is 3 and 5 in terms of stability and internal performance.

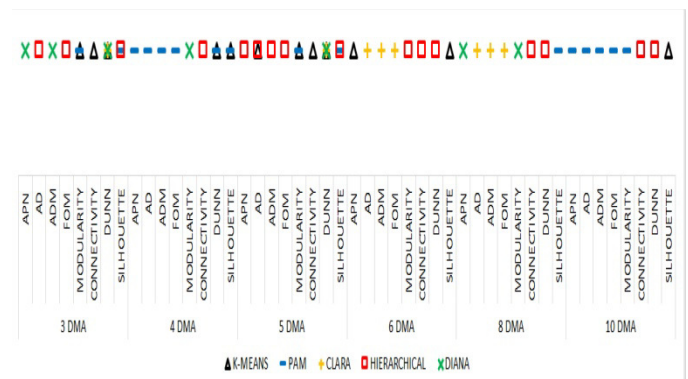


Fig. 8. C-town WDN spectral clustering.

VI. CONCLUSIONS

Improving the management of WDNs can be achieved by dividing the network into sectors (DMAs). Among the most used methods for network partitioning is spectral partitioning, which consists of generating a spectral space from the eigenvectors of the Laplacian matrix associated with the network. The final phase of this partitioning is often done by the k-means algorithm, which is not always adequate. It is, therefore, useful to consider the performance of other existing algorithms in order to highlight the most efficient in terms of clustering quality indices such as modularity, internal indices (Connectivity, Silhouette, and DUNN) as well as stability indices such as APN, AD, ADM, and FOM.

A simulation exercise was conducted to assess and evaluate the performance of 5 clustering algorithms on 3 types of WDNs. The studied networks were chosen according to their type and size. Our findings show that k-means is not consistently effective either in terms of type of network or number of sectors. Our critical study of [11], shows further that k-means performs well only under special conditions including small-sized networks with a very limited number of sectors. Thus, one cannot confirm the validity of using k-means as a tool to sectorize DMAs in larger and medium-sized networks.

ACKNOWLEDGEMENTS

This research paper would not have been possible without the exceptional support of Dr. Issam Malki from the University of Westminster, Department of Economics and Quantitative Methods (EQM).

REFERENCES

- [1] K. B. Adedeji, Y. Hamam, B. T. Abe, and A. M. Abu-Mahfouz, "Pressure Management Strategies for Water Loss Reduction in Large-Scale Water Piping Networks: A Review," in *Advances in Hydroinformatics*, P. Gourbesville, J. Cunge, and G. Caignaert, Eds. New York, NY, USA: Springer, 2018, pp. 465–480.
- [2] S. Ates, "Hydraulic modelling of closed pipes in loop equations of water distribution networks," *Applied Mathematical Modelling*, vol. 40, no. 2, pp. 966–983, Jan. 2016, <https://doi.org/10.1016/j.apm.2015.06.017>.
- [3] A. Di Nardo and M. Di Natale, "A Design Support Methodology for District Metering of Water Supply Networks," in *12th Annual Conference on Water Distribution Systems Analysis*, Arizona, USA, Sep. 2010, pp. 870–887, [https://doi.org/10.1061/41203\(425\)80](https://doi.org/10.1061/41203(425)80).
- [4] H. E. Mutikanga, S. K. Sharma, and K. Vairavamoorthy, "Methods and Tools for Managing Losses in Water Distribution Systems," *Journal of Water Resources Planning and Management*, vol. 139, no. 2, pp. 166–174, Mar. 2013, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000245](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000245).
- [5] A. Di Nardo, M. di Natale, G. Santonastaso, and S. Venticinqu, "Graph partitioning for automatic sectorization of a water distribution system," in *Proceedings of 11th International Conference on Computing and Control for Water Industry (CCWI). Urban Water Management: Challenges and Opportunities*, Jan. 2011, pp. 841–846.
- [6] X. Khoa Bui, M. S. Marlim, and D. Kang, "Water Network Partitioning into District Metered Areas: A State-Of-The-Art Review," *Water*, vol. 12, no. 4, Apr. 2020, Art. no. 1002, <https://doi.org/10.3390/w12041002>.
- [7] J. Saldarriaga et al., "Battle of the Water Networks District Metered Areas," *Journal of Water Resources Planning and Management*, vol. 145, no. 4, Apr. 2019, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001035](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001035).
- [8] P. Korkkana, V. Kanakoudis, M. Patelis, and K. Gonelas, "Forming District Metered Areas in a Water Distribution Network Using Genetic Algorithms," *Procedia Engineering*, vol. 162, pp. 511–520, Jan. 2016, <https://doi.org/10.1016/j.proeng.2016.11.095>.
- [9] L. Perelman and A. Ostfeld, "Topological clustering for water distribution systems analysis," *Environmental Modelling & Software*, vol. 26, no. 7, pp. 969–972, Jul. 2011, <https://doi.org/10.1016/j.envsoft.2011.01.006>.
- [10] J.-P. Attal, "Nouveaux algorithmes pour la detection de communautés disjointes et chevauchantes bases sur la propagation de labels et adaptes aux grands graphes.," Ph.D. dissertation, Université de Cergy Pontoise, Cergy-Pontoise, France, 2017.
- [11] A. Di Nardo, C. Giudicianni, R. Greco, M. Herrera, and G. F. Santonastaso, "Applications of Graph Spectral Techniques to Water Distribution Network Management," *Water*, vol. 10, no. 1, Jan. 2018, Art. no. 45, <https://doi.org/10.3390/w10010045>.
- [12] V. G. Tzatchkov, V. H. Alcocer-Yamanaka, V. G. Tzatchkov, and V. H. Alcocer-Yamanaka, "Graph theory based single and multiple source water distribution network partitioning," *Tecnología y ciencias del agua*, vol. 10, no. 6, pp. 197–221, Dec. 2019, <https://doi.org/10.24850/j-tyca-2019-06-08>.
- [13] E. Price and A. Ostfeld, "Optimal Water System Operation Using Graph Theory Algorithms," *Procedia Engineering*, vol. 89, pp. 502–508, Jan. 2014, <https://doi.org/10.1016/j.proeng.2014.11.245>.
- [14] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, Dec. 2007, <https://doi.org/10.1007/s11222-007-9033-z>.
- [15] B. Gharnali and S. Alipour, "MRI Image Segmentation Using Conditional Spatial FCM Based on Kernel-Induced Distance Measure," *Engineering, Technology & Applied Science Research*, vol. 8, no. 3, pp. 2985–2990, Jun. 2018, <https://doi.org/10.48084/etasr.1999>.
- [16] H. Alizadeh and B. M. Bidgoli, "Introducing A Hybrid Data Mining Model to Evaluate Customer Loyalty," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1235–1240, Dec. 2016, <https://doi.org/10.48084/etasr.741>.
- [17] S. P. Singh and S. C. Sharma, "A Novel Energy Efficient Clustering Algorithm for Wireless Sensor Networks," *Engineering, Technology & Applied Science Research*, vol. 7, no. 4, pp. 1775–1780, Aug. 2017, <https://doi.org/10.48084/etasr.1277>.
- [18] B. Nica, *A Brief Introduction to Spectral Graph Theory*. Zürich, Switzerland: European Mathematical Society, 2018.
- [19] R. Horaud, "A Short Tutorial on Graph Laplacians, Laplacian Embedding, and Spectral Clustering," 2009.
- [20] Y. P. Raykov, A. Boukouvalas, F. Baig, and M. A. Little, "What to Do When K-Means Clustering Fails: A Simple yet Principled Alternative Algorithm," *PLOS ONE*, vol. 11, no. 9, Aug. 2016, Art. no. e0162259, <https://doi.org/10.1371/journal.pone.0162259>.
- [21] R. Han and J. Liu, "Spectral Clustering and Genetic Algorithm for Design of District Metered Areas in Water Distribution Systems," *Procedia Engineering*, vol. 186, pp. 152–159, Jan. 2017, <https://doi.org/10.1016/j.proeng.2017.03.221>.
- [22] X. Jin and J. Han, "K-Medoids Clustering," in *Encyclopedia of Machine Learning*, C. Sammut and G. I. Webb, Eds. Boston, MA, USA: Springer, 2010, pp. 564–565.
- [23] I. Tyuryukanov, J. Quiros-Tortos, M. Naglic, M. Popov, M. A. M. M. van der Meijden, and V. Terzija, "A post-processing methodology for robust spectral embedded clustering of power networks," in *17th International Conference on Smart Technologies*, Ohrid, Macedonia, Jul. 2017, pp. 805–809, <https://doi.org/10.1109/EUROCON.2017.8011221>.
- [24] A. Candelieri, D. Conti, and F. Archetti, "Improving Analytics in Urban Water Management: A Spectral Clustering-based Approach for Leakage Localization," *Procedia - Social and Behavioral Sciences*, vol. 108, pp. 235–248, Jan. 2014, <https://doi.org/10.1016/j.sbspro.2013.12.834>.
- [25] J. L. Diaz, M. Herrera, J. Izquierdo, I. Montalvo, and R. Perez, "A Particle Swarm Optimization derivative applied to cluster analysis," in *4th international congress on environmental modelling and software*, Barcelona, Spain, Jul. 2008.

- [26] R. J. Sanchez-Garcia *et al.*, "Hierarchical Spectral Clustering of Power Grids," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2229–2237, Sep. 2014, <https://doi.org/10.1109/TPWRS.2014.2306756>.
- [27] A. K. Patnaik, P. K. Bhuyan, and K. V. Krishna Rao, "Divisive Analysis (DIANA) of hierarchical clustering and GPS data for level of service criteria of urban streets," *Alexandria Engineering Journal*, vol. 55, no. 1, pp. 407–418, Mar. 2016, <https://doi.org/10.1016/j.aej.2015.11.003>.
- [28] S. Emmons, S. Kobourov, M. Gallant, and K. Borner, "Analysis of Network Clustering Algorithms and Cluster Quality Metrics at Scale," *PLOS ONE*, vol. 11, no. 7, Jun. 2016, Art. no. e0159161, <https://doi.org/10.1371/journal.pone.0159161>.
- [29] O. Giustolisi and L. Ridolfi, "New Modularity-Based Approach to Segmentation of Water Distribution Networks," *Journal of Hydraulic Engineering*, vol. 140, no. 10, Oct. 2014, Art. no. 04014049, [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000916](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000916).
- [30] J. Handl and J. Knowles, "Exploiting the Trade-off — The Benefits of Multiple Objectives in Data Clustering," in *Evolutionary Multi-Criterion Optimization*, Guanajuato, Mexico, Mar. 2005, pp. 547–560, https://doi.org/10.1007/978-3-540-31880-4_38.
- [31] G. Brock, V. Pihur, S. Datta, and S. Datta, "clValid: An R Package for Cluster Validation," *Journal of Statistical Software*, vol. 25, no. 4, pp. 1–22, Mar. 2008, <https://doi.org/10.18637/jss.v025.i04>.
- [32] J. C. Dunn, "Well-Separated Clusters and Optimal Fuzzy Partitions," *Journal of Cybernetics*, vol. 4, no. 1, pp. 95–104, Jan. 1974, <https://doi.org/10.1080/01969727408546059>.
- [33] V. V. Ryazanov, "Collective Solutions on Sets of Stable Clusterings," in *Recent Applications in Data Clustering*, H. Pirim, Ed. London, UK: IntechOpen, 2018, pp. 221–236.
- [34] H. Liu, M. Zhao, C. Zhang, and G. Fu, "Comparing Topological Partitioning Methods for District Metered Areas in the Water Distribution Network," *Water*, vol. 10, no. 4, Apr. 2018, Art. no. 368, <https://doi.org/10.3390/w10040368>.
- [35] E. Pournaras, R. Taormina, M. Thapa, S. Galelli, V. Palleti, and R. Kooij, "Cascading Failures in Interconnected Power-to-Water Networks," *ACM SIGMETRICS Performance Evaluation Review*, vol. 47, no. 4, pp. 16–20, Dec. 2020, <https://doi.org/10.1145/3397776.3397781>.
- [36] D. Gilbert, E. Abraham, I. Montalvo, and O. Piller, "Iterative Multi-Stage Method for a Large Water Network Sectorization into DMAs under Multiple Design Objectives," *Journal of Water Resources Planning and Management*, vol. 143, no. 11, Nov. 2017, Art. no. 04017067, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000835](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000835).