



THE SHAPE OF WATER DISTRIBUTION SYSTEMS - DESCRIBING LOCAL STRUCTURES OF WATER NETWORKS VIA GRAPHLET ANALYSIS


Bulat Kerimov¹, Franz Tscheikner-Gratl², Riccardo Taormina³ and David B. Steffelbauer⁴

^{1,2}Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, (Norway)

³ Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, (The Netherlands)

⁴ KWB - Kompetenzzentrum Wasser Berlin, Berlin (Germany)

¹ bulat.kerimov@ntnu.no, ² franz.tscheikner-gratl@ntnu.no, ³ r.taormina@tudelft.nl,

⁴ david.steffelbauer@kompetenz-wasser.de

Abstract

The performance, vulnerability, and resilience of water distribution systems (WDS) depend, to varying degrees, on its underlying topological structure, herein referred to as its shape. Literature mostly differentiates between two main shapes of networks - branched or looped. However, the shape of real networks lies in between the two extremes of purely branched and looped systems. Although these networks are globally topologically different, they may show high similarity at the local scale of a borough or a neighbourhood. Recent studies focused on describing WDS via links and nodes by using graph theory. These first attempts at graph-theoretical applications showed promising results in describing the global structure and estimating the global resilience of WDS, but there are a limited number of measures that take the importance of local topology into consideration.

This research enters the new terrain of local WDS investigations using graphlet analysis to describe this local topology in more detail. Graphlets are small connected subgraphs of a large network which have recently gathered much attention as a useful concept to characterise local topology and uncover structural design principles of complex networks. Consequently, these novel analysis techniques can provide new insights into how local WDS structures influence their overall behaviour.

In this work, we first provide a framework to describe local and global topology with graphlets. We then employ the framework to assess the local criticality of two benchmark WDS, linking the results to topological metrics already adopted in the literature. Additionally, we analyse the potential gain of graphlet analysis in the prediction of local vulnerabilities by including graphlet features into a random forest regression setting. As a result, we observe a positive trend in performance in comparison to a similar model without graphlet features.

Keywords

Graphlet, graph theory, resilience, water distribution, topological structure.

1. INTRODUCTION

WDSs naturally relate to graphs, associating the pipes and junctions in the water system with links and nodes on a mathematical object. That is why graph theory can be utilised to describe and compare WDSs with different topologies in a unified mathematical language and serve as sound support in decision making in planning and optimization of networks. Water system research successfully adopted graph-theoretical measures in support of various tasks. The current set of instruments includes a large variety of graph-theoretical measures that are widely used in resilience and redundancy analysis [1], [2], [3], identifying critical components [4], [5], [6], and

sectorisation [7], [8]. The resilience of a network, defined as the ability of the system to recover after adversarial effects [3], partially depends on the topology of the WDS. Earlier works showed that more resilient networks can be designed by graph theoretical approaches, for example, by taking spectral and statistical graph measures into consideration. [2] systematically reviewed and evaluated the correlation between common topological measures and network resilience. Additionally, topological and spectral properties were used in [9] and [10] to assess global and local vulnerabilities. These works highlight the importance of such metrics as algebraic and spectral connectivities. The utility of centrality measures for the assessment of local vulnerability has been indicated in [11] and [12].

Nevertheless, traditional graph theoretical measures (e.g., average node degree, graph diameter, link density) do not provide detailed information on local neighbourhood structures of networks but rather a global description. However, this local information can be of high importance in automatic generation of WDS [13], and applications that require a more precise description of those networks (e.g., node and link criticality analysis) [4], [5]. There are two sides to the criticality of elements of WDSs: fault probability and the impact severity. In this work, we introspect the latter part of the equation by looking only at the local topological structure. With the terms of local *importance*, and *criticality* we refer to the same phenomena, particularly - the total impact of the fault in the element on the whole system. Utilities and engineers estimate the impact by simulating pipe breakages, abrupt increases in demands, and other faults with hydraulic engines such as EPANET. The main issue of a hydraulic-based critical element analysis is computational costs for larger WDSs. On the flip side of the coin, pure topological measures will fail to take into account energy dissipation and the physical nature of the water distribution process. Some works attempt to bridge energy-based measures of local criticality with topological indicators. For example, water-flow centrality [14] and node demand centrality [15] assign hydraulically informed weights to centrality metrics. Furthermore, energy-informed shortest routes gently introduce energy loss properties into the graph structure [16],

With this work, we open the way for a more detailed description of WDSs by introducing graphlet representation to address the aforementioned challenges. Graphlet analysis aims to extract small induced subgraphs that appear in the network. Graphlets have been successfully applied in network alignment, description of brain networks [17] and used as an early precursors for in social networks [18]. The ultimate goal is to provide a relevant and comparable numerical representation of local connectivities of water networks, i.e. a topological fingerprint of the WDS. This research additionally covers the assessment of the descriptive power of graphlets in the estimation of local (i.e. node) criticality. Analysis of global network vulnerability and various resilience metrics is outside the scope of this project.

2. METHODOLOGY

In the first step of our methodology we introduce the reader to graphlet decomposition, the products of graphlet analysis, and the relationship between graphlets and traditional topological metrics. We show how the coarseness of the network influences the results of the analysis by comparing graphlet representation of the network versus the representation of the skeletonized version of the network. Next, we evaluate the intrinsic relationship between the graphlet features with existing topological measures. In this step, we carry out a graphlet analysis on a set of generated networks and estimate the correlations with the topological metrics on a global (network-level) and local (node-level) scale. Specifically, we carry out a Principal Component Analysis (PCA) on extracted graphlet features and assess the correlations of principal components with those features.

Second, we assess the descriptive power of graphlet analysis in the prediction of the importance of a node. We perform the analysis by incorporating the results of the graphlet into the random

forest regression model and estimate the performance gain by comparing it with the performance of the same model without graphlet features.

2.1 Graphlet analysis

Graphlets are small subgraphs that collectively comprise the structure of the network. One can think of graphlets as (atomic) building blocks of the graph. A plethora of research in graphlet analysis focuses on extracting the graphlet level description of the network with a relatively low computational effort. This is usually performed by some form of a counting algorithm. Since direct enumeration of the graphlets is rather computationally expensive, researchers attempt to speed up the calculations by exploiting the symmetry properties of graphs [19] or employing sampling strategies.

In this work, we employed the ORCA algorithm due to the availability of the source code and particular suitability for sparse graphs. Due to the computational constraints, the size of the graphlets included in ORCA is limited to 5. Theoretically, the number of potential distinct graphlets is infinite and grows exponentially with its size. However, this may reduce the applicability of the graphlet representation due to a sparse representation. Current rules to find certain graphlets are mostly designed by hand, although there are advancements towards the automatic generation of those rules [20].

For a finer representation of graphlets, counting algorithms operate on the level of *orbits*. An orbit is an automorphism of a graphlet, e.g. isomorphism of the graphlet to itself. In other words, the parts of the graphlet that are symmetrical to others will belong to the same orbit. Orbit counting algorithms thus identify how many times a node touches the corresponding orbit. In total, 29 graphlets of sizes up to 5 contain 73 distinct orbits.

As a result, ORCA produces vectors on 2 scales - a local of size 73 and assigned to each node level, and a global, i.e. a vector of 29 which is assigned to a whole graph. The values on the former vector correspond to the number of times an orbit occurs on each node. We further denote this vector as an orbit count vector (OCV). OCVs are comparable between the nodes of different water networks. The global vector is calculated by summing up OCVs over nodes of a graph and results in the total graphlet count in that network. We denote this vector as a graphlet count vector (GCV).

One important aspect of the graphlet representation is the influence of the coarseness of the graph layout. Commonly, the number of nodes and connections in the topology varies depending on the ultimate task. Skeletonization is a technique to simplify networks that simultaneously alters the granularity of the graph [21]. As a consequence, the graphlet analysis is sensitive to this granularity. In our work, we assess the influence after 2 steps of skeletonization. The first step is a unification of adjacent consecutive pipes. The second is trimming dead ends and parallel connections.

After retrieving graphlet and orbit counts from the network we perform a PCA. PCA is a widely used method of dimensionality reduction that decomposes the variables into a new set of uncorrelated features through a singular value decomposition. As the result, we obtain a low dimensional representation of the original data, while preserving most of the variance. These new features are called Principal Components (PC_i). Each of them is composed as a result of a linear combination of original features and explains a certain share of variance. In the case of GCV and OCV, we use PCA to find the inner correlations between the features and to compare the key components altogether rather than each graphlet or orbit individually. To extend the analysis, for each principal component we leverage the information about contributions of original features and explained variances.

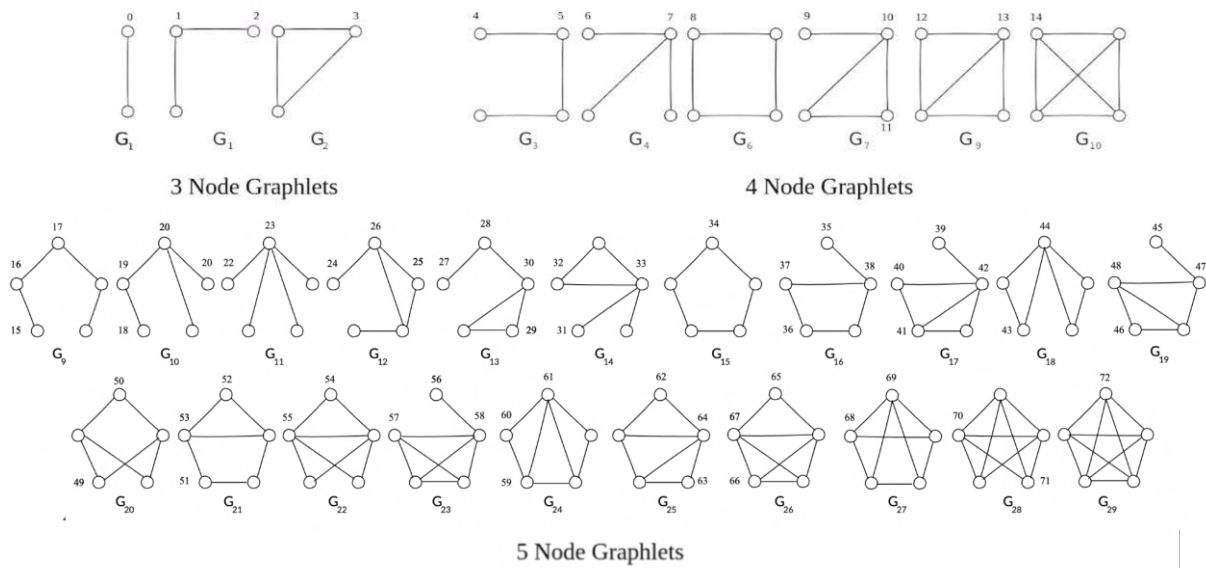


Figure 1. Existing graphlets (with sizes up to 5 nodes) and their orbits.

2.2 Graph-theory based measures

In this section we summarize topological measures that describe graph networks. As was mentioned above, graphlet analysis produces vectors on 2 scales - a local (node) level, and a global (network) level. The former representation can be viewed as a heatmap on the nodes network, while the latter assigns a vector to a whole graph. We thus compare both representations with the corresponding topological measures separately. These measures are widely used in WDS analysis, mostly in the context of vulnerability and resilience analysis.

2.2.1 Local measures

Node degree

Node degree is a basic metric that measures the number of incidental links (or pipes to the network).

$$d(v) = \sum_{\substack{j \in V \\ j \neq i}} A_{ij} \quad (1)$$

Here A_{ij} is the element of an adjacency matrix.

Betweenness centrality

Betweenness centrality is a statistical value that measures the relative position of a node on the network. A higher value of centrality is usually assigned to the “hubs” of the network. It indicates how the rest of the network “depends” on the node.

$$c_b(v) = \sum_{\substack{j \in V \\ j \neq i}} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \quad (2)$$

where σ_{ij} is the total number of shortest paths from node i to node j , $\sigma_{ij}(v)$ is a number of those paths that pass through the node v .

Closeness centrality

Closeness centrality is another centrality metric that assigns a value to a node depending on the distance from the node to the rest of the network. In contrast to betweenness, closeness centrality indicates the independence of the node from the rest of the network.

$$c_c(v) = \sum_{\substack{j \in V \\ j \neq i}} \frac{n-1}{d_{ij}} \quad (3)$$

Here d_{ij} denotes the distance from node i to node j , while n is the total number of nodes.

PageRank

PageRank is a ranking algorithm originally proposed by Google research and designed to assign relative importance to nodes on the graph based on number of incidental links and the “quality” of incidental links. This quality is likewise defined by the importance of the source nodes.

2.2.2 Global measures

For the comparison we chose the following set of descriptors, that are used in the estimation of redundancy and resilience of WDSs. [2] provides a systematic review over the vast majority of common topological metrics and their correlation with network resilience.

Distribution of node degrees

The distribution of node degrees is simply the normalized distribution of all node degrees over the entire graph G . Similarly to the average node degree, it indicates spatial organization of a network. In the context of WDS, where the node degree rarely exceeds 4, a right skew in the distribution corresponds to a highly connected structure. On the contrary, left-skewed distribution is related mostly to a tree-like structure.

Number of cycles

Number of cycles in the network is an indicator of redundancy. For example, clustering coefficient is defined by the total number of triangles in the loop. Intuitively, a higher number of 3-cycles suggests availability of alternative paths from supply to demand nodes.

Meshedness

Meshedness measures a fraction of the actual number of cycles of any order (or loops of any size) to the maximum possible number of cycles in the network. Meshedness is directly related to the number of alternative paths in a supply system and hence to the redundancy of the network [1].

$$c_m(G) = \frac{f}{2n-5} \quad (4)$$

Here, f denotes the number of total independent loops and n denotes the number of nodes.

Connectivity

Although connectivity metrics (e.g. spectral gap, algebraic connectivity) and modularity are important factors to the resilience of the network [2], we deliberately omit these measures from the comparison with GCV. We motivate it by the fact that connectivity and modularity measure properties of a larger scale rather than the one of graphlet analysis.

2.3 Local vulnerability analysis

A system is as resilient as its most vulnerable element, therefore there is a strong demand from utilities and planners to identify which elements have the highest chance of failing. One of the paradigms to approach this problem is based on simulating various types of faults in a simulator

such as EPANET or others. A drastic change in the operations of the WDS, such as a burst or a pipe break, will impact the overall distribution of water pressure. Consequently, low pressure on the consumer end results in the system's incapability to supply the required volumes of water. We simulate a pipe break by closing the pipe during a 72 hour simulation and measure the total amount of water that was not supplied to end consumers. The final indicator of the importance of the pipe is measured by the amount of unsupplied demand conditioned by the fault in this element in a pressure-driven simulation. Here \underline{q}_j and q_j denote demands during pressure driven simulations and base demand during normal conditions correspondingly. To facilitate comparability across WDSs, we evaluate this metric on a logarithmic scale.

$$c_{crit}(e) = \sum_{j \in V, j \neq i} \frac{(q_j - \underline{q}_j)}{\underline{q}_j} \quad (5)$$

We additionally constrain the lowest values at -10. As a result, the values of criticality are distributed in the negative range up to 0. Higher values correspond to higher criticality.

It is important to mention that criticality analysis yields pipe-level information, i.e. a value assigned to a link. In contrast, OCV is a junction-level vector. We thus transform the pipe criticality into a junction criticality by averaging the values over incidental links of the node.

$$C_{crit}^e = C_{crit}^v B D^{-1}, \quad (6)$$

where C_{crit} is vectors of criticality assigned to vertices or edges, B denotes an incidence matrix, and D is a degree matrix.

3. CASE STUDY

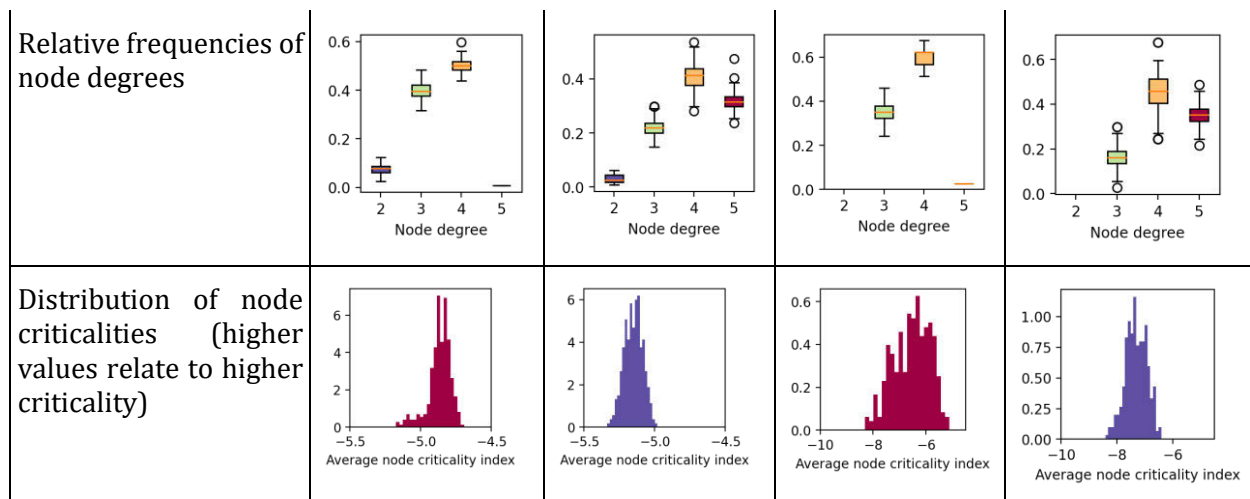
We selected ZJ and FOS networks due to their homogeneity in network geometry (e.g. elevations, diameters, and pipe length) and locations of the reservoirs, which limits the influence of network hydraulics on our graph-based analysis. In order to operate on a comparable dataset, we generated 1200 artificial variations of those networks. We introduced several surgical morphological transformations on the local connectivity and geometry during data generation.

3.1 Network generation

We first define the proximity radius r by calculating the median geographical distance between connected junctions in the network. Next, for every junction, we define unconnected candidates in the local neighbourhood located within the radius. We also define the upper limit to the degree of each junction as 4 and 5 and evaluate both cases separately. With a probability $p = 0.5$ a pipe is generated for each possible candidate node.

Table 1. Summary of network generation.

	ZJ		FOS	
Number of networks	300	300	300	300
Maximum node connectivity	4	5	4	5
Number of nodes	114	114	38	38



In Table 1, we observe a shift in the distribution of average node criticality towards less critical conditions (i.e., lower values) with the increase of the maximum connectivity (from 4 to 5) during data generation.

It is important to mention that some of the generated connections are likely to be unfeasible in real-life scenarios. Some of the connections intersect or double existing pipes. We leave the investigation of the feasibility and associated costs as a potential for future work.

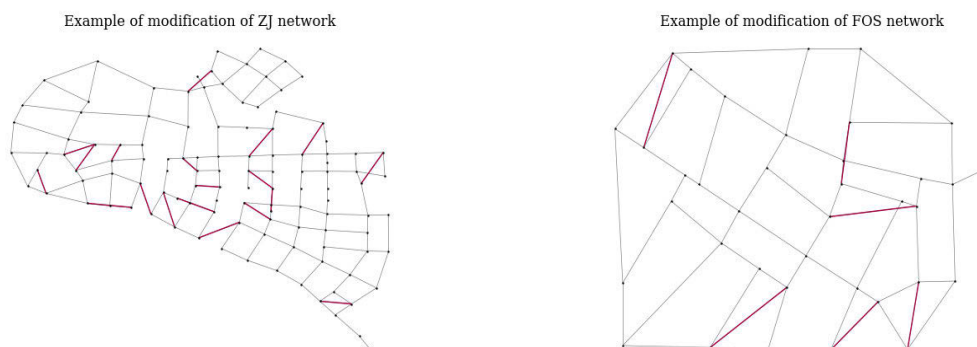


Figure 2. Examples of topology modification of generated network variants (red links denote generated pipes). Both have a limit of a maximum of 4 connections.

4. RESULTS AND DISCUSSION

In this section we provide the result of the analysis of the correlation of graphlet counts with the topological measures described above. Additionally, to inspect the relationship of the graphlet count to the local criticality we show the effect of including the graphlet representation in the random forest regression model.

4.1 The influence of skeletonization on graphlet representation

On the example of L-Town network, we observe that reducing the number of adjacent essentially reduces the order of cycles on the graph and increases the average degree. Consequently, this

increases the count of higher-level orbits related to the graphlets with cycles (e.g. G_2 , G_5 , G_{12} , etc.). We show in Figure 2 the example that skeletonization allows the graphlet counting algorithm to capture higher-order graphlets that wouldn't be captured otherwise. The second step continues the trend increasing the count of higher orbits, such as G_{37} and G_{38} . and results in more dense GCV. However, we argue that for criticality analysis, this level of coarseness might be detrimental.

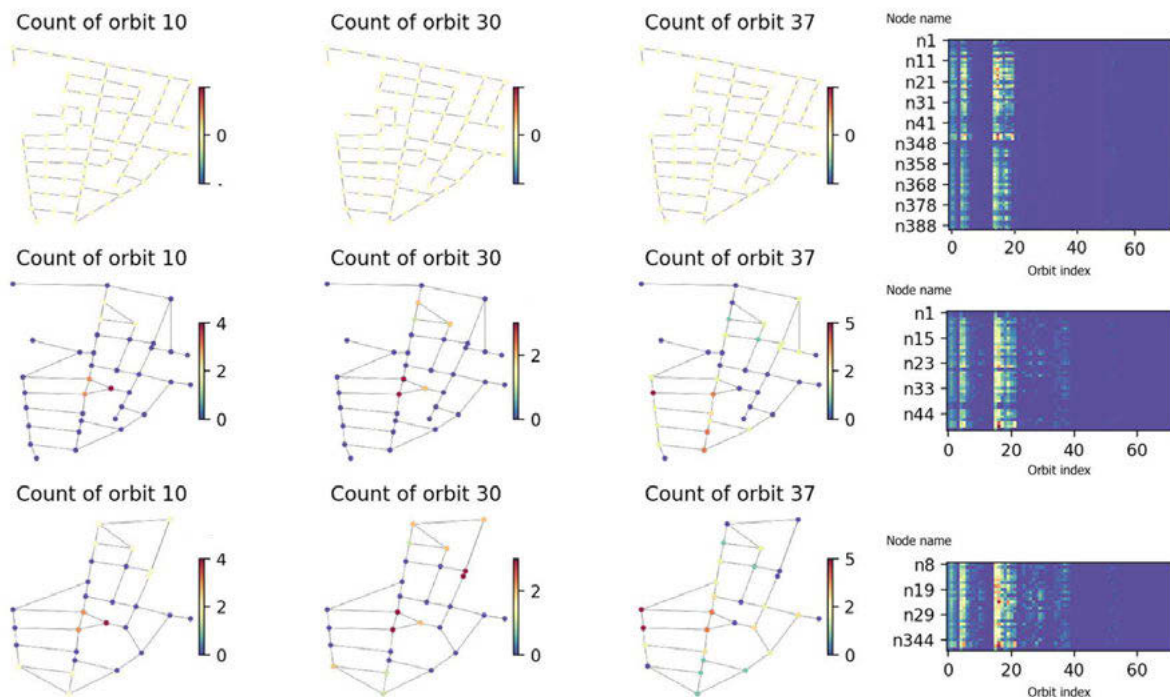


Figure 3. Example of influence of skeletonization on the graphlet count on the example of area C in L-Town water network (top - no skeletonization, center - unification of adjacent pipes, bottom - removal of parallel connections and dead-ends)

4.2 Principal component analysis

High granularity of the orbit representation comes at the cost of interpretability. Specifically, some of the graphlets can be present as components of the others, which can lead to certain autocorrelation in the orbit representation vector. We hypothesize that it is possible to translate OCV into a lower dimensionality representation via principal component decomposition.

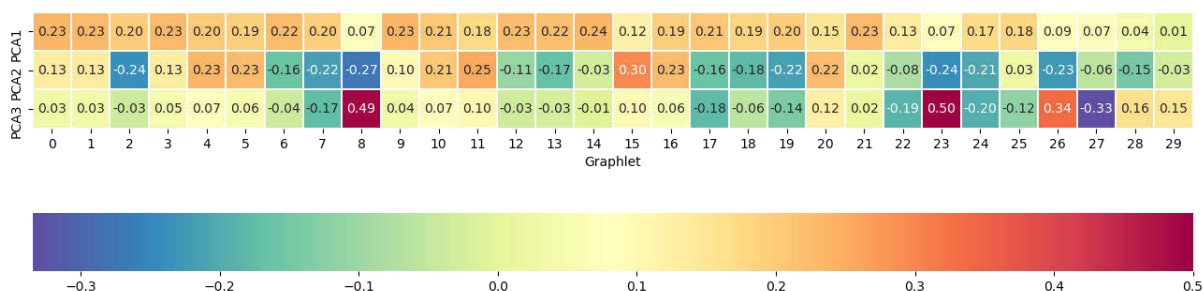


Figure 4 (A). Contribution of each graphlet to the principal components of GCV of the overall dataset. A higher magnitude indicates a higher contribution of that graphlet with the principal components.

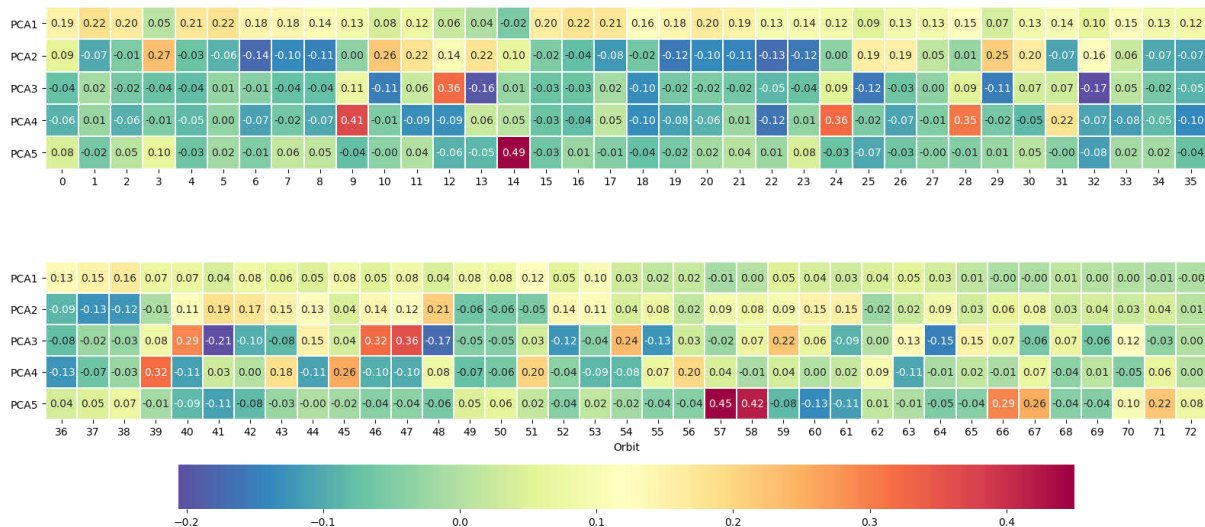


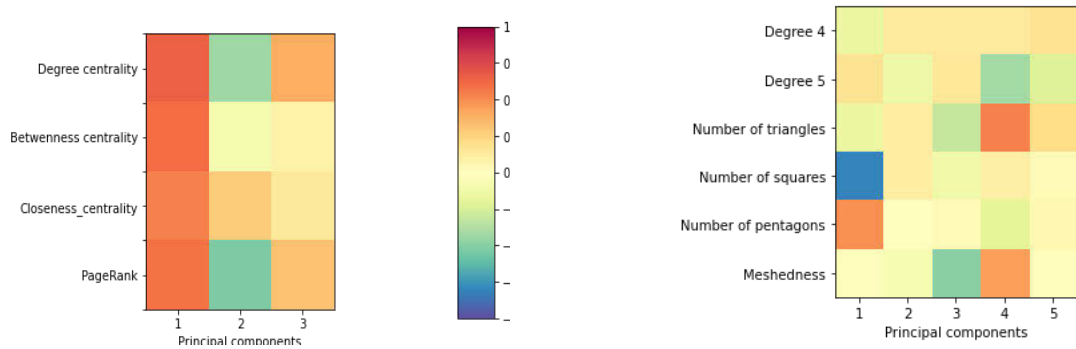
Figure 4 (B) Contribution of each orbit to the principal components of OCVs of the overall dataset. A higher magnitude indicates a higher contribution of that orbit with the principal components. Explained variances are 24%, 16%, 5.8%, 5.1, and 4.5% correspondingly.

According to weighting in Figure 5 (A), most of the graphlets collectively contribute to the first principal component (PC₁). This might be a result of intercorrelation between the counts of most graphlets. PC₂ is correlated with the graphlets that contain cycles in the structure and can be translated to the “loopedness” of networks. Lastly, PC₃ is predominantly described as graphlets that contain 4 fully connected nodes in the structure. Another observation is that orbit 14 (which belongs to G₈) has comparably low contribution to PC₁ in OCV as well. Since 4 fully connected nodes is a rare structure, these features explain the smaller share of variance (PC₃ in GCV and PC₅ in OCV).

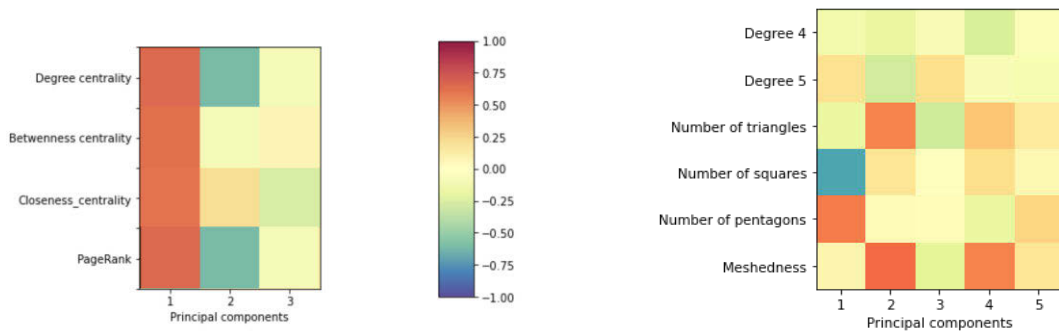
Likewise we observe that a large proportion of graphlets contribute to the first principal component of OCV. We can also note that orbits 1, 4, 5, 15, 16, 17, 20, and 21 have slightly higher weighting than the others. All these orbits belong to 4 graphlets of the same shape - a sequence of nodes (2, 3, 4, and 5 nodes) and additionally to G₁₉. We extend the discussion in the section. Likewise, orbits that correspond to “corners” of the triangles are included in PC₂. At the same time, orbits 19-23 have a negative weight, which emphasises the focus on the triangular structures. PC₄ is highly correlated with orbits 9, 24, and 28. All these orbits are located in the neighbourhood of triangular structures.

Below we summarize the correlation of the principal components of the entire OCV and GCV datasets for two networks separately. In each case, we carry out PCA separately in order to assess the difference in the main PCs.

4.2.1 Relationship of PCs to graph-theoretical metrics

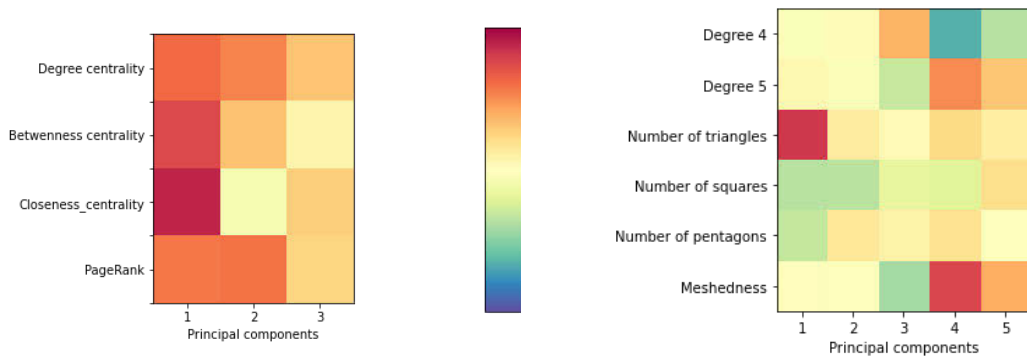


(a) Generated networks with maximum connectivity 4

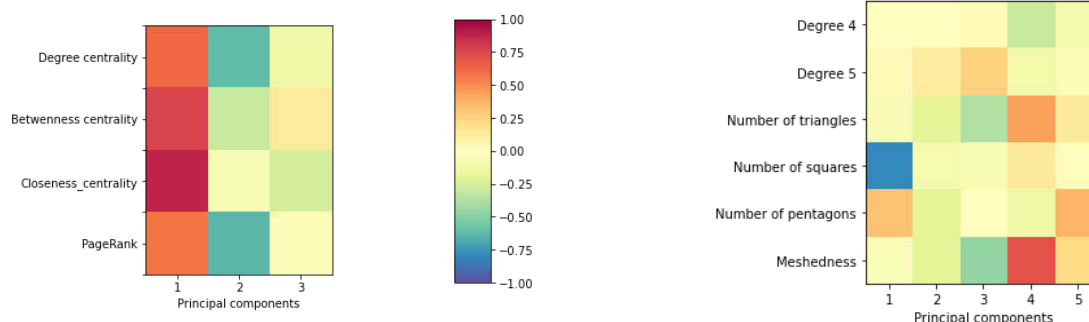


(b) Generated networks with maximum connectivity 5

Figure 5. Correlations of local graph-theoretical metrics with the principal components of OCV (left) and global graph-theoretical metric with GCV (right) on the example of modifications of the ZJ network.



(a) Generated networks with maximum connectivity 4



(b) Generated networks with maximum connectivity 5

Figure 6. Correlations of local graph-theoretical metrics with the principal components of OCV (left) and global graph-theoretical metric with GCV (right) on the example of modifications of FOS network

Particular orbital indices are inherently related to graph-theoretical measures that characterize the structure and topology of a graph network. For example, an orbital count of the orbit 0 is equivalent to node degree. In other words, the number of times a node “touches” orbit 0 equals to a number of incidental links. Graph orbits of graphlets sequence of nodes (e.g. orbits 4, 5, 15, 16, and 17) are likely to appear in the parts of the network that are located in between larger neighbourhoods as a hub. One can imagine this as a hyperbolic increase in the number of nodes in the k -hop neighbourhood of a node as k increases. In other words, the number of combinations a graphlet G_9 can be placed in that neighbourhood will be high, which increases the count in its turn. In all of the cases, these orbits showed the highest contribution to the first principal component of OCVs (left pictures in Figures 5 and 6) and a strong correlation with centrality metrics.

The number of loops (or n -cycles) such as triangles, squares, or pentagons, in the network, is natively related to the number of identified graphlets that have loops in them. Since the graphlets in the ORCA counting algorithm are limited to the size of 5 nodes, higher-order cycles are omitted. Some of the graphlets, e.g. G_9 , G_7 , G_{14} , and G_{22} contain triangular structures in themselves, which means that their orbit counts might be correlated. As we show in the principal component analysis, some of the principal components are highly correlated with the number of identified squares and triangles. We further observe in the case of ZJ a stronger correlation of its PC_1 with the number of squares and pentagons than in the case of FOS. We argue that the original structure of ZJ contains higher-order cycles which contribute to the variability in GCVs, hence higher correlation with the first principal component.

4.3 Random Forest Regression model

To assess the informational gain from incorporating the graphlet analysis we introduce the feature vector from orbit count in a random forest model. Here we hypothesize that incorporating graphlet features into the regression model that estimates the criticality of the node will increase the performance of a regression model.

We then set up the task as a regression model with the importance of a node as the target variable of the i . We evaluate 3 cases of used features in the model: solely traditional graph measures (a), solely OCV (b), and traditional features and OCV altogether (c). Further on, we compare the performance measured by coefficient determination. Traditional metrics include degree, betweenness and closeness centrality, and distance to reservoirs.

Table 2. Summary of network generation

	ZJ		FOS	
Maximum node connectivity	4	5	4	5
Topological features, R²	0.36	0.25	0.30	0.18
OCV, R²	0.36	0.21	0.32	0.15
Topological features + OCV, R²	0.47	0.34	0.39	0.22

The random forest model has been set up with the same set of hyperparameters and evaluated with a coefficient of determination on 10% of the dataset.

We can see in the example above the improvement in the predictive power of the regression model as we include GCV in the analysis. This hints at the conclusion that graphlet representation contains local information that assists in the prediction.

5. CONCLUSIONS & FUTURE WORK

Consequently, this work shows that graphlet representation is beneficial as a descriptor of WDS. GCV captures local patterns and implicitly contains information about the global structure. PCA showed that GCV captures well the number of loops of different sizes and meshedness of a network. Simultaneously, OCV aids in the identification of small hubs and contains information about the degrees of the network. It is important to note that although the graphlet analysis enriches the set of descriptors of the graph structure and provides higher resolution to a local neighbourhood, it is not an exhaustive descriptor of the network structure. A key factor that limits the expressive power of graphlet analysis appears from the influence of global connectivity. Although graphlet decomposition is capable of identifying hubs in the influence of connectivity and network modularity on a global scale might not be captured.

Furthermore, this research introduces graphlets as an assisting factor in assessing local criticality. As it was mentioned, solely topological representation is not a piece of sufficient but necessary information. This problem implies the need for a holistic approach that includes both hydraulics and geometry. For example, the location of the reservoir and the main artery of the WDS is definitive for the importance of the elements of the network. Nevertheless, incorporating graphlet features into the analysis shows promising first results.

For future work, we suggest extending the dataset with more hydraulically “independent” and less modular networks in order to further mitigate the influence of these factors. For example, an artificial WDS with many homogeneously placed reservoirs can aid in the analysis.

Adding hydraulically informed edge weights to the graphlet count could be a potential direction for the analysis. Orbit count might be weighted by hydraulic costs, e.g. by multiplication of node counts by a weight factor. This can serve as a bridge between hydraulics and topological description of the local neighbourhood and increase the expressiveness of graphlets.

6. REFERENCES

- [1] Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(1), 016111. <https://doi.org/10.1063/1.3540339>
- [2] Meng, F., Fu, G., Farmani, R., Sweetapple, C., & Butler, D. (2018). Topological attributes of network resilience: A study in water distribution systems. *Water Research*, 143, 376–386. <https://doi.org/10.1016/J.WATRES.2018.06.048>
- [3] Herrera, M., Abraham, · Edo, Stoianov, I., & Abraham, E. (2016). A Graph-Theoretic Framework for Assessing the Resilience of Sectorised Water Distribution Networks. *Water Resour Manage*, 30, 1685–1699. <https://doi.org/10.1007/s11269-016-1245-6>
- [4] Torneyeviadzi, H. M., Neba, F. A., Mohammed, H., & Seidu, R. (2021). Nodal vulnerability assessment of water distribution networks: An integrated Fuzzy AHP-TOPSIS approach. *International Journal of Critical Infrastructure Protection*, 34, 100434. <https://doi.org/10.1016/J.IJCIP.2021.100434>
- [5] Giudicianni, C., di Nardo, A., di Natale, M., Greco, R., Francesco Santonastaso, G., & Scala, A. (n.d.). Topological Taxonomy of Water Distribution Networks. <https://doi.org/10.3390/w10040444>
- [6] Pinto, J., Varum, H., Bentes, I., Agarwal, J., Pinto, J., Bentes, · I, Bentes, I., Varum, H., & Agarwal, J. (2010). A Theory of Vulnerability of Water Pipe Network (TVWPN). *Water Resources Management* 24:15, 24(15), 4237–4254. <https://doi.org/10.1007/S11269-010-9655-3>
- [7] Perelman, L., & Ostfeld, A. (2011). Topological clustering for water distribution systems analysis. *Environmental Modelling & Software*, 26(7), 969–972. <https://doi.org/10.1016/J.ENVSOFT.2011.01.006>
- [8] Hajebi, S., Roshani, E., Cardozo, N., Barrett, S., Clarke, A., & Clarke, S. (2016). Water distribution network sectorisation using graph theory and many-objective optimisation. *Journal of Hydroinformatics*, 18(1), 77–95. <https://doi.org/10.2166/HYDRO.2015.144>
- [9] Gutiérrez-Pérez, J. A., Herrera, M., Pérez-García, R., & Ramos-Martínez, E. (2013). Application of graph-spectral methods in the vulnerability assessment of water supply networks. *Mathematical and Computer Modelling*, 57(7–8), 1853–1859. <https://doi.org/10.1016/J.MCM.2011.12.008>
- [10] Phan, H. C., Dhar, A. S., & Bui, N. D. (2021). Accounting for Source Location on the Vulnerability Assessment of Water Distribution Network. *Journal of Infrastructure Systems*, 27(3), 04021024. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000620](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000620)
- [11] Giustolisi, O., Ridolfi, L., & Simone, A. (2019). Tailoring Centrality Metrics for Water Distribution Networks. *Water Resources Research*, 55(3), 2348–2369. <https://doi.org/10.1029/2018WR023966>
- [12] Agathokleous, A., Christodoulou, C., Christodoulou, S. E., & Christodoulou schristo, S. E. (2017). Topological Robustness and Vulnerability Assessment of Water Distribution Networks. *Water Resour Manage*, 31, 4007–4021. <https://doi.org/10.1007/s11269-017-1721-7>
- [13] Sitzenfrey, R., Möderl, M., & Rauch, W. (2013). Automatic generation of water distribution systems based on GIS data. *Environmental Modelling and Software*, 47, 138–147. <https://doi.org/10.1016/J.ENVSOFT.2013.05.006>
- [14] Herrera, M., Abraham, E., & Stoianov, I. (2015). Graph-theoretic surrogate measures for analysing the resilience of water distribution networks. *Procedia Engineering*, 119(1), 1241–1248. <https://doi.org/10.1016/J.PROENG.2015.08.985>
- [15] Zarghami, S. A., & Gunawan, I. (2020). A domain-specific measure of centrality for water distribution networks. *Engineering, Construction and Architectural Management*, 27(2), 341–355. <https://doi.org/10.1108/ECAM-03-2019-0176/FULL/PDF>
- [16] Ulusoy, A.-J., Stoianov, I., & Chazerain, A. (n.d.). Hydraulically informed graph theoretic measure of link criticality for the resilience analysis of water distribution networks. <https://doi.org/10.1007/s41109-018-0079-y>
- [17] Finotelli, P., Piccardi, C., Miglio, E., & Dulio, P. (2021). A Graphlet-Based Topological Characterization of the Resting-State Network in Healthy People. *Frontiers in Neuroscience*, 15, 376. <https://doi.org/10.3389/FNINS.2021.665544/BIBTEX>
- [18] Jamra, H., Savonnet, M., & Leclercq, E. (2021). Detection of Event Precursors in Social Networks: A Graphlet-Based Method. *Research Challenges in Information Science*, 415, 10. <https://doi.org/10.1007/978-3-030>

- [19] Hočevar, T., & Demšar, J. (2014). A combinatorial approach to graphlet counting. *Bioinformatics*, 30(4), 559–565. <https://doi.org/10.1093/BIOINFORMATICS/BTT717>
- [20] Melckenbeeck, I., Audenaert, P., Michoel, T., Colle, D., & Pickavet, M. (2016). An Algorithm to Automatically Generate the Combinatorial Orbit Counting Equations. *PLoS ONE*, 11(1). <https://doi.org/10.1371/JOURNAL.PONE.0147078>
- [21] Saldarriaga, J. G., Ochoa, S., Rodríguez, D., & Arbeláez, J. (2009). Water distribution network skeletonization using the resilience concept. *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference, WDSA 2008*, 852–864. [https://doi.org/10.1061/41024\(340\)74](https://doi.org/10.1061/41024(340)74)
- [22] The PageRank Citation Ranking: Bringing Order to the Web. (1998). www.yahoo.com

