AQUA
Water Infrastructure, Ecosystems and Society

IWA PUBLISHING

 AQUA — Water Infrastructure, Ecosystems and Society Vol 70 No 8, 1257 doi: 10.2166/aqua.2021.107

Evaluation of the vulnerability in water distribution systems through targeted attacks

Aiman Albarakati [a,*], Asifa Tassaddiq [a] and Yogesh Kale[b]

[a] College of Computer and Information Sciences, Majmaah University, Al-Majma'ah, Saudi Arabia
[b] HPC Research Admin, CDSE, North Carolina A&T State University, Greensboro, NC 27401, USA
*Corresponding author. E-mail: a.albarakati@mu.edu.sa

AA, 0000-0002-9004-9718; AT, 0000-0002-6165-8055

ABSTRACT

This paper presents the results of a vulnerability analysis in different water distribution system (WDS) benchmarks, performed under a framework based on a graph model that integrates topological features and hydraulic characteristics, allowing the comparison between different attack strategies and centrality measures in terms of their ability to predict the shortage of water supply. This vulnerability framework has been previously applied to electric power systems and it employs a vulnerability prediction measure to quantify the amount of damage caused in terms of the physical damage measure. Different attack strategies and centrality measures were applied to four WDS benchmarks: the New York Tunnel, the Hanoi, the Modena, and the Balerma networks. It was determined that removing the most central element and recalculating the centrality for each stage are the most damaging attack strategies. Degree, eigenvector, and Katz centrality measures presented the best performance to predict the elements that are more relevant to the network and can have a larger impact on the water supply. It was demonstrated that the vulnerability framework can be applied to the WDS in the same way it was previously applied to electric power systems. Future work will be oriented to the design of the WDS using optimization techniques to minimize the vulnerability of the network under faults that can be generated by droughts and other extreme weather conditions.

Key words: centrality measures, targeted attacks, vulnerability, water distribution systems, water flow

HIGHLIGHT

- This paper presents the results of a vulnerability analysis in different WDS benchmarks, performed under a framework based on a graph model that integrates topological features and hydraulic characteristics, allowing the comparison between different attack strategies and centrality measures in terms of their ability to predict the shortage of water supply.

1. INTRODUCTION

Water is a scarce and indispensable resource for maintaining human life, and access to it has been undertaken as a sustainable development goal by the United Nations. To accomplish this goal, a water distribution system (WDS) must be extended and optimized to reliably deliver the required water quantity and quality to a larger percentage of the human population. Furthermore, these systems are fundamental to food security through crop irrigation and health through sanitary services (United Nations 2020). WDSs are interdependent and complex structures formed by water sources, treatment plants, pipes, junctions, tanks, reservoirs, pumps, and valves, among other elements, designed to satisfy the water demand at consumption points (Christie *et al.* 2011). Nevertheless, these structures might be vulnerable in front of natural disasters and other threats such as droughts and ill-intended attacks.

Furthermore, climate change has generated extreme weather conditions that affect water reservoirs and other elements of the WDS. Therefore, it is important to research how to guarantee the stability and continuity of the water supply by understanding the vulnerability of the WDS when they face 'attacks' which in this work are defined as the outages of a set of elements that might be caused by different natural or artificial conditions.

Mathematical models are fundamental in researching the vulnerability of the WDS, and typically a WDS model consists of establishing a relationship between the variables that intervene in the water flow problem (pressures and volume of water flow) and the parameters of each element (Cabrera & García-Serra 1999). The complex network theory provides an

invaluable framework for the representation and quantitative study of the WDS, through analysis of subsystem connections, quantifying topological redundancy, and identifying critical components whose failure may impact the overall performance of the system (Barabási & Albert 1999; Yazdani & Jeffrey 2012; Barabási 2013; Albarakati *et al.* 2017; Albarakati & Bikdash 2021; Albarakati 2021a, 2021b).

In recent years, water distribution networks have been studied as complex networks (Murthy & Murthy 2012; Alexiou & Tsouros 2017). In Giudicianni *et al.* (2018), the authors analyzed the topology of several water distribution networks, concluding that they tend to be sparse and resemble regular square grids with homogenous node degree values. Some vulnerability studies regarding the WDS have been carried out considering topological features of such networks, and additional metrics have been proposed to identify critical nodes (Yazdani & Jeffrey 2012). The relationship of the betweenness centrality of the network's nodes and the vulnerability of a WDS was demonstrated by Agathokleous *et al.* employing a study case under continuous and intermittent conditions (Agathokleous *et al.* 2017). In Gutiérrez-Pérez *et al.* (2013), a methodology based on spectral measurements was introduced to establish the importance of areas in the WDS, the authors employed two ranking algorithms: PageRank and HITS.

More recently, a methodology for calculating an index of vulnerability that represents the tendency of an injected contaminant to spread over the network was introduced, this index compares the behavior of the network to a scale-free network and allows different networks to be ranked in terms of their resistance to contamination (Nicolini 2020). Mortula *et al.* conducted an integrated network and hydraulic analysis for identifying critical nodes that can be impacted by water shortages and proposing managerial action to reduce network vulnerability, diversify water sources, and reduce possibilities of contamination (Mortula *et al.* 2020). The adjacency matrix from the WDS network has been employed to detect unintended isolation segments that can cause vulnerability, helping to determine the best location of valves for avoiding the interruption of the service for pipe failures (Jeong *et al.* 2021). Also, the nodal vulnerability of water distribution networks has been evaluated under cascading failures through a study case, where there was a high probability of the nodal failures triggering a cascading failure (Shuang *et al.* 2014).

Wéber *et al.* explored the behavior and topology of the segments resulting from an accidental pipe burst, especially the criticality of such segments from the point of view of the whole system. They defined the network vulnerability as a measure to quantify the expected value of relative consumption that cannot be served due to a single accidental pipe break (Wéber *et al.* 2020). In Giudicianni *et al.* (2021), a novel vulnerability index called the cut-vulnerability is introduced and several networks with different topological structures were analyzed under the proposed framework, while the vulnerability measure is based on the loss of connectivity.

Those research works represent important advances in the application of complex network techniques and centrality measures to the study of the WDS. However, such metrics need to be further investigated to understand their reliability by using benchmarks to establish appropriate ways to measure the vulnerability of water distribution networks. In this sense, this research conducts a vulnerability analysis in different WDS benchmarks, considering a graph model that integrates topological features and hydraulic characteristics to compare different centrality measures in terms of their ability to predict the shortage of water supply using the graph model. This framework is different from that proposed in the literature because it is based on the topology and the operating condition of the system, helping to predict the vulnerability under different operating conditions.

2. MODELING OF WATER DISTRIBUTION SYSTEM

The main components of the WDS are water sources, treatment plants, pipes, junctions, tanks, reservoirs, pumps, and valves. For modeling the network, these elements can be represented by nodes and links: nodes are the points of the system where there can be an input or output of water, and links convey water from one node to another.

Each node of the model is associated with an elevation and there are different types of nodes: for example, reservoir nodes where the head is forced and water is provided to the system, consumption nodes where there is a determined water demand, and the junction nodes that are mainly connections between pipes. Links in the network model represent pipes, which are the most abundant components of a WDS, allowing the water to move from one point to another of the system. In these links, it is important to consider the friction losses causing energy to dissipate across them.

In this paper, a simplified graph model is employed with the weighted graph $G(v, l)$, where $v = \{v_1, v_2, \ldots, v_n\}$ is the set of nodes and $l = \{l_1, l, \ldots, l_m\}$ is the set of links and each link is associated with a corresponding weight $w = \{w_1, w_2, \ldots, w_m\}$.

The weights represent the water flow Q_m calculated using the quasi-steady flow (Equations (1) and (2)):

$$\sum_{m=1}^{N_i} Q_m - Q_{d,i} = 0 \tag{1}$$

$$H_{i,u} - H_{i,d} - \Delta H_i = 0 \tag{2}$$

where Q_m is the unknown flow in every N_i link/pipe connected to the ith network node, $Q_{d,i}$ is the known demand at the node i, $H_{i,u}$ is the unknown total head at the upstream node of the ith pipe, $H_{i,d}$ is the unknown total head at the downstream node of the ith pipe, and ΔH_i is the calculated difference between the ith pipe's total headloss and pumping head.

The change in total energy across each link is determined by the following equation:

$$\Delta H_i = R_i Q_i |Q_i|^{n-1} + K Q_i |Q_i| - H_{p,i}(Q_i) \tag{3}$$

where R_i is the pipe's resistance coefficient, whose value depends on the friction head loss model, the value of n also depends on such a model, K is the local head loss coefficient, and $H_{p,i}(Q_i)$ is the head delivered by a pump installed at link i, which is a function of the flow delivered (Christie *et al.* 2011).

Since these equations are nonlinear, a numeric solver is required to obtain a fast solution. In this research, the water flow is determined with the EPANET-MATLAB Toolkit under quasi-static conditions (Eliades *et al.* 2016).

3. CENTRALITY MEASURES

Centrality metrics are the most widely employed indicators for identifying important nodes in a network, and depending on the type of network and the study case, the importance might have a different meaning (Agathokleous *et al.* 2017). In this research, the centrality measures the importance in terms of the water flow that runs through a node or a pipe.

In previous works, different centrality measures have been evaluated for similar applications in electric power systems, and the most consistent were degree, eigenvector, and Katz centralities based on power traffic, which can be considered an equivalent of water flow. Furthermore, betweenness, closeness, and PageRank centralities have been considered relevant in current computational applications. Betweenness, closeness, and degree centralities have been cataloged as suitable metrics for studying centrality in spatial networks (Giustolisi *et al.* 2019).

The complexity of calculating centrality metrics depends on the network size, connectivity, and the weights associated. In this work, the mentioned centralities are compared in terms of their efficiency for assessing the vulnerability of the WDS. Most of the centralities are calculated using the weighted graph model previously described as detailed in the following subsections. For a better comparison, the centralities are normalized in such a way that the sum of the measures for all the nodes or links is equal to one.

3.1. Degree centrality

Degree centrality ($D_i(G)$) is a popular centrality measure that gives some insight into the connectivity of a node i; however, it can miss potentially important aspects of the network's architecture and the position of the node within it (Bloch & Jackson 2021).

In general, it measures the number of edges connected to a node i, and for the weighted graph as in the current research, it must consider the weights of such edges; therefore, it can be calculated as follows:

$$D_i(G) = \frac{\sum_j Q_{ij}}{\sum_i \sum_j Q_{ij}} \tag{4}$$

where Q_{ij} is the water flow between junctions i and j.

3.2. Eigenvector centrality

The eigenvector centrality ($E_i(G)$) was proposed by Bonacich (2007) as a measure of relevance relying on the idea that the prominence of a node is related to the relevance of its neighbors. It is calculated assuming the centrality of node i is

proportional to the sum of the centralities of its neighbors (Bloch & Jackson 2021). This metric is calculated as follows:

$$E_i(G) = \frac{1}{\lambda_{max}} \sum_{j=1}^{n} Q_{ij} u_j \tag{5}$$

where λ_{max} is the largest eigenvalue of the weighted adjacency matrix based on water flows, and u_j is the jth element of the eigenvector corresponding to λ_{max}.

3.3. Betweenness centrality

The Freeman's betweenness centrality ($B_i(G)$) of node i measures the relevance of such a node in terms of its role connecting the rest of the nodes in the network. Betweenness is meant to capture the importance of the node as an intermediary in the association between other nodes, in the case of the WDS, in the transportation of water among the rest of the nodes.

This metric takes into consideration the number of shortest paths that cross the selected node and the total amount of shortest paths. It is calculated as follows:

$$B_i(G) = \sum_{j \neq k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \tag{6}$$

where $\sigma_{jk}(i)$ is the number of shortest paths between junctions j and k that pass through junction i, and σ_{jk} is the total number of shortest paths between junctions j and k. In this research, the calculation of the shortest paths does not take into account the weights of the paths.

3.4. Closeness centrality

The closeness centrality of node i($C_i(G)$) is based on the network distance between such an element and the rest of the nodes. The most relevant nodes using this metric are the ones that need fewer steps to communicate with other nodes in the network. It can be calculated as follows:

$$C_i(G) = \frac{1}{\sum_j d_{ij}} \tag{7}$$

where d_{ij} is the distance between junctions i and j, which is calculated based on the weights of the links that connect them. The harmonic closeness is an aggregation of the inverse distances between junctions (nodes), and this is an alternative way of calculating the closeness (Bloch & Jackson 2021), which can be expressed as follows:

$$C_i(G) = \sum_{j \neq i} \frac{1}{d_{ij}} \tag{8}$$

This measure can also be normalized so that it spans from 0 to 1, obtaining:

$$C_i(G) = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}} \tag{9}$$

It is important to note that the closeness works for fully connected networks; otherwise, the distance between two nodes belonging to different components is infinite, and then, the metric of Equation (7) is null, and harmonic closeness (Equation (8)) is employed to avoid this problem (Giustolisi *et al.* 2019).

3.5. PageRank centrality

The PageRank centrality was proposed as a metric to classify websites according to their relative importance. Intuitively, the PageRank centrality is high when the sum of the backlinks, i.e., the links that point to the website, is high or these backlinks are highly rated (Page *et al.* 1999; Riquelme *et al.* 2018). In the case of the WDS, if a junction has a high PageRank centrality,

it means the number of other junctions pointing to it is high, or if few junctions are pointing, then they have a large PageRank. A simplified formula for the PageRank centrality is as follows:

$$PR_i(G) = \alpha \sum_{j \neq i} \frac{PR_j(G)}{\delta + (j)} \tag{10}$$

where $\delta + (j)$ is the out-degree of node j, and the damping factor α is a number between 0 and 1.

In this work, the PageRank centrality $PR_i(G)$ calculation is based on the MATLAB© algorithm, which is the result of a random walk of the network. At each node in the graph, the next node is chosen with a probability (by default 0.85) from the set of successors of the current node: if the node has no successors, the next node is chosen from all the nodes. The centrality score is the average time spent at each node during the random walk (The Mathworks Inc. 2020).

3.6. Katz centrality

The Katz centrality was introduced as a method for calculating the popularity of an individual considering not only the number of direct 'votes' received by each individual but the status of each one of the individuals that emit such 'votes' (Katz 1953). In the case of the WDS, Katz centrality represents the importance of a node considering the water that passes through it and the importance of the junctions from which they come. This centrality is calculated as follows:

$$K_i(G) = ((I - \alpha \boldsymbol{A}^T)^{-1} - I)e_n \tag{11}$$

where $\boldsymbol{A}$ is the adjacency matrix of the graph G based on the water flow weights, I is the $n \times n$ identity matrix, e_n is a column vector with n ones, and α is the attenuation factor, which must be smaller than the reciprocal of the absolute value of the largest eigenvalue of A, that is:

$$\alpha < \frac{1}{|\lambda_{max}|} \tag{12}$$

This measure has been previously applied to the WDS (Giudicianni *et al.* 2020) and in social networks (Riquelme *et al.* 2018). In some large-size network cases, it did not converge.

To exemplify the calculation of the centrality measures, Figure 1 shows an example network with five nodes and its weighted adjacency matrix $\boldsymbol{A}$, and this graph models a WDS example obtained from the EPANET-MATLAB Toolkit (EPANET-MSX Example Network) (Eliades *et al.* 2016). Table 1 contains all the centrality measures calculated as described

$$\boldsymbol{A} = \begin{bmatrix} 0.000 & 4.069 & 7.131 & 0.000 & 15.300 \\ 4.069 & 0.000 & 0.669 & 0.000 & 0.000 \\ 7.131 & 0.669 & 0.000 & 2.300 & 0.000 \\ 0.000 & 0.000 & 2.300 & 0.000 & 0.000 \\ 15.300 & 0.000 & 0.000 & 0.000 & 0.000 \end{bmatrix}$$

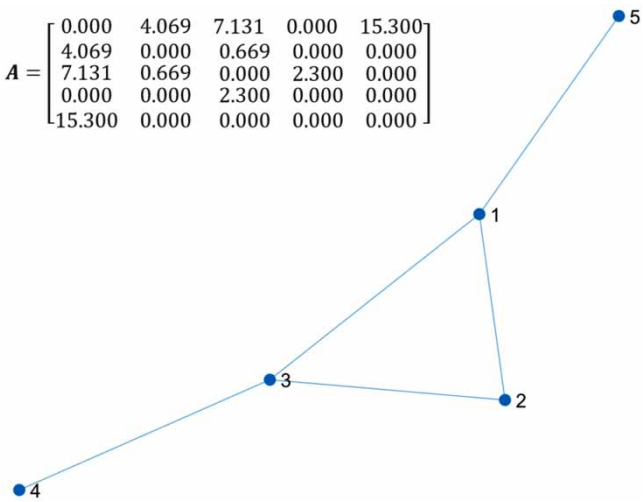

Figure 1 | Example network graph.

Table 1 | Centrality measures calculated for the example network graph

Junction	Degree	Eigenvector	Betweenness	Closeness	PageRank	Katz
1	0.449	0.385	0.500	0.244	0.290	0.402
2	0.080	0.095	0.000	0.204	0.194	0.090
3	0.171	0.163	0.500	0.244	0.290	0.171
4	0.039	0.021	0.000	0.153	0.112	0.031
5	0.259	0.336	0.000	0.153	0.112	0.305

in the previous subsections, and the same functions are applied through all the research work to determine the centrality measures.

4. CONVERSION TO LINE GRAPHS

The previous section was based on node (junction) centralities; however in this work, the link (pipe) centralities are also evaluated. To obtain the pipe centralities, the weighted adjacency matrix $\boldsymbol{A}$ of graph $G(v, e)$ is converted to an equivalent line graph adjacency matrix $\boldsymbol{M}$ by applying the transformation to the line graph (Yoshida 2013). $\boldsymbol{M}$ is a $e \times e$ matrix with e equal to the number of pipes in the grid, the elements m_{ij} are the line weights, which represent a water flow from pipe i to pipe j.

After converting the graph to a line graph, the centrality measures are calculated in the same way that they are calculated for a node graph, with $\boldsymbol{M}$ as the adjacency matrix.

Figure 2 shows the results of transforming the example network graph from Figure 1 to a line graph where the five pipelines are mapped into nodes, and $\boldsymbol{M}$ is the new adjacency matrix of the line graph.

Table 2 shows the results of the centrality measures calculated for the equivalent line graph.

4.1. Framework for the evaluation of attacks

The vulnerability assessment framework is based on previous works from the authors focused on electric power systems' vulnerability implementing individually targeted attacks and sequential attack profiles on nodes and links of the corresponding graph model. An attack profile is a sequence of elements (nodes or links) selected from the network for their consecutive removal.

The framework proposes using the vulnerability curve, which represents a physical damage measure in the horizontal axis and a functional damage measure in the vertical axis. This graph is helpful to determine the amount of functional damage caused by a fault profile.

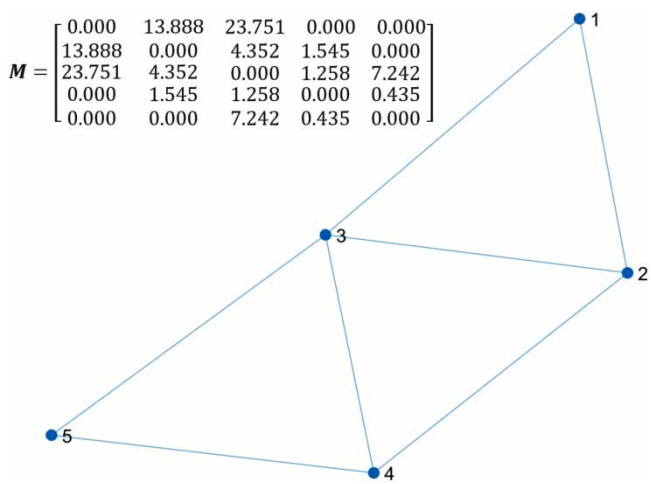

Figure 2 | Example network line graph.

Table 2 | Centrality measures calculated for the example network line graph

Pipes	Degree	Eigenvector	Betweenness	Closeness	PageRank	Katz
1	0.358	0.355	0.000	0.169	0.149	0.356
2	0.188	0.211	0.167	0.203	0.212	0.201
3	0.348	0.328	0.167	0.254	0.277	0.337
4	0.031	0.025	0.167	0.203	0.212	0.028
5	0.073	0.079	0.000	0.169	0.149	0.077

For the WDS, the functional damage is measured by the fraction of unsatisfied demand (UD) determined after removing the faulty elements. This is calculated under the removal of the first k elements of the failure sequence:

$$UD(k) = 1 - \frac{wd_s(k)}{wd_{total}} \tag{13}$$

where $wd_s(k)$ is the water demand satisfied by the available sources in the current stage, and wd_{total} is the initial total water demand of the system.

The physical damage is measured by the fraction of removed elements on the mth stage of the failure sequence, that is

$$foe(k) = \frac{k}{m} \tag{14}$$

The vulnerability prediction measure (VPM) describes the amount of damage caused in terms of the physical damage measure, and it is calculated with the area under the vulnerability curve. In this sense, the higher the VPM, it is said that the attack sequence causes more severe electrical damage to the power system; thus, the power system is more vulnerable to such attack.

For calculating the unsatisfied demand in each stage, the WDS modeled in EPANET is converted to a MATLAB graph object (The Mathworks Inc. 2020), and then $UD(k)$ is calculated and recorded for each stage depending on the type of attack, as it is described in the following subsection.

4.2. Types of attacks

In this research, three types of fault sequences are evaluated: remove most central element first (RMCEF), iterated most central element first (IMCEF), and iterated most damaging element first (IMDEF). The RMCEF profile is represented by the vector $\boldsymbol{V}$ containing the ordered elements according to centrality, which is as follows:

$$\boldsymbol{V} = \{v_1, v_2, \ldots, v_n\} \tag{15}$$

$$C(v_1) > C(v_2) > \cdots > C(v_n) \tag{16}$$

Figure 3(a) shows the simplified flow diagram for the RMCEF algorithm, and in this type of attack, the elements are removed from the network starting from the highest to the lowest centrality one element per stage.

The IMCEF is an attack profile that also relies on centrality measures, but in this type of attack, the centrality measures are recalculated after the removal of an element. The idea under this attack profile is that the centrality measures change after the removal of the most central element; thus, the second most central element in the initial ranking may not be the most central, once the most central element is removed. Then, the vector of centralities must be recomputed to obtain the new ranking of elements according to centrality (Figure 3(b)).

Furthermore, the IMDEF fault profile is defined in this work as an attack based on removing the element that generates the largest raise in the UD of the WDS in the current stage. The simplified algorithm for obtaining the vulnerability curve using the IMDEF fault profile is shown in Figure 3(c). This algorithm has an internal cycle to determine the unsatisfied demand of each element of the network under current conditions; afterward, it selects the most damaging element (MDE) to remove it from the network.

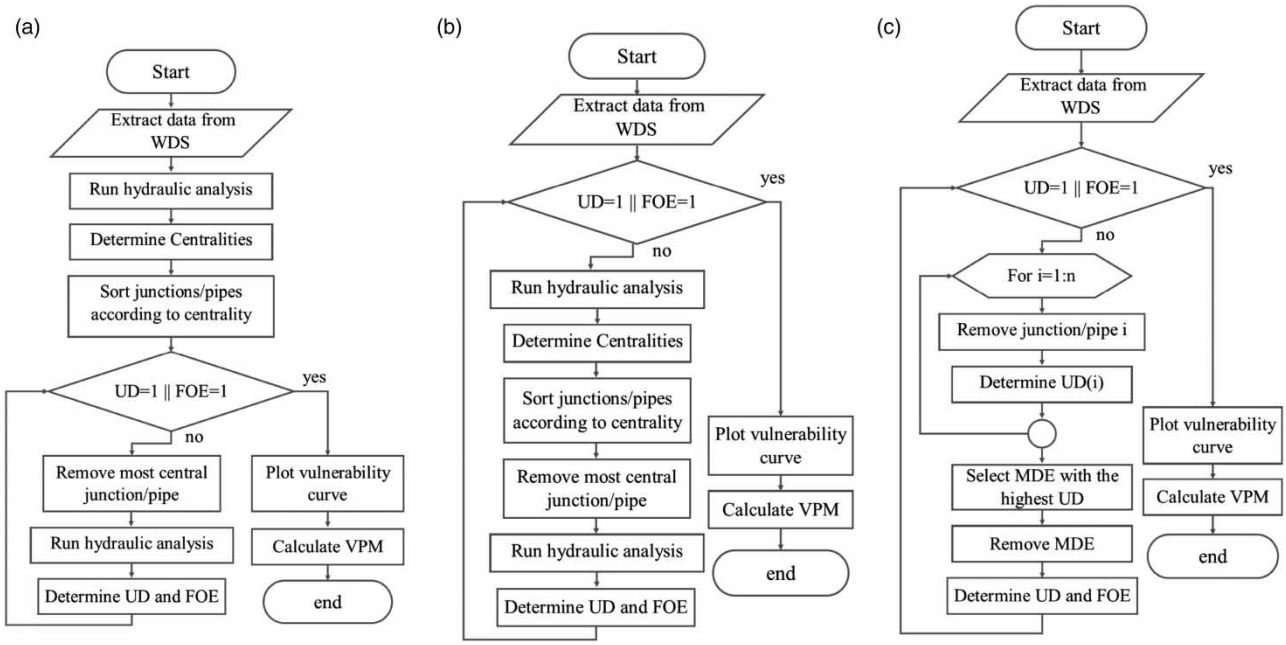

Figure 3 | Summary of algorithms for determination of the vulnerability curve and VPM with different fault strategies: (a) RMCEF, (b) IMCEF, and (c) IMDEF.

5. BENCHMARKS

To compare the different centrality measures and attack strategies, four benchmarks were selected from previous research and reports. The benchmark models are called: the New York Tunnel (NYT), the Hanoi (HAN), the Modena (MOD), and the Balerma (BAL) networks, the former two are considered medium-size networks and the latter two are large-size networks.

Table 3 shows the characteristics of each benchmark selected for this research, from the smallest to the largest based on the data from Centre for Water Systems University of Exeter (2014). The NYT network was originally presented as an extended design problem (Schaake & Lai 1969); the HAN WDS consists of a structure of three loops and one reservoir with a fixed head of 100 m (Kim *et al.* 1994); the MOD WDS has both minimum and maximum pressure requirements for demand nodes with pipes made of cast iron (Bragalli *et al.* 2008); the BAL system is an adaptation of an existing irrigation network in the Sol-Poniente irrigation district, located in Balerma in the province of Almeria (Spain), and it has a fixed water consumption across all the demand nodes (Reca & Martínez 2006).

The selection of benchmarks was based on having diverse characteristics, i.e., medium and large systems with different sizes of pipes and materials. The topologies of the WDS benchmarks are shown in Figure 4, where they are meshed networks, which allow the redirection of the power flow when the links are broken.

6. EVALUATION OF ATTACKS ON BENCHMARKS

This section presents the results of the experiments performed using the vulnerability evaluation framework for the benchmarks selected. The experiments are based on fault profiles for junctions and pipes, using the centrality measures

Table 3 | Characteristics of the benchmarks

Benchmarks	Acronyms	Junctions	Pipes	Reservoirs
New York Tunnel	NYT	20	21	1
Hanoi	HAN	32	34	1
Modena	MOD	272	317	4
Balerma	BAL	447	454	4

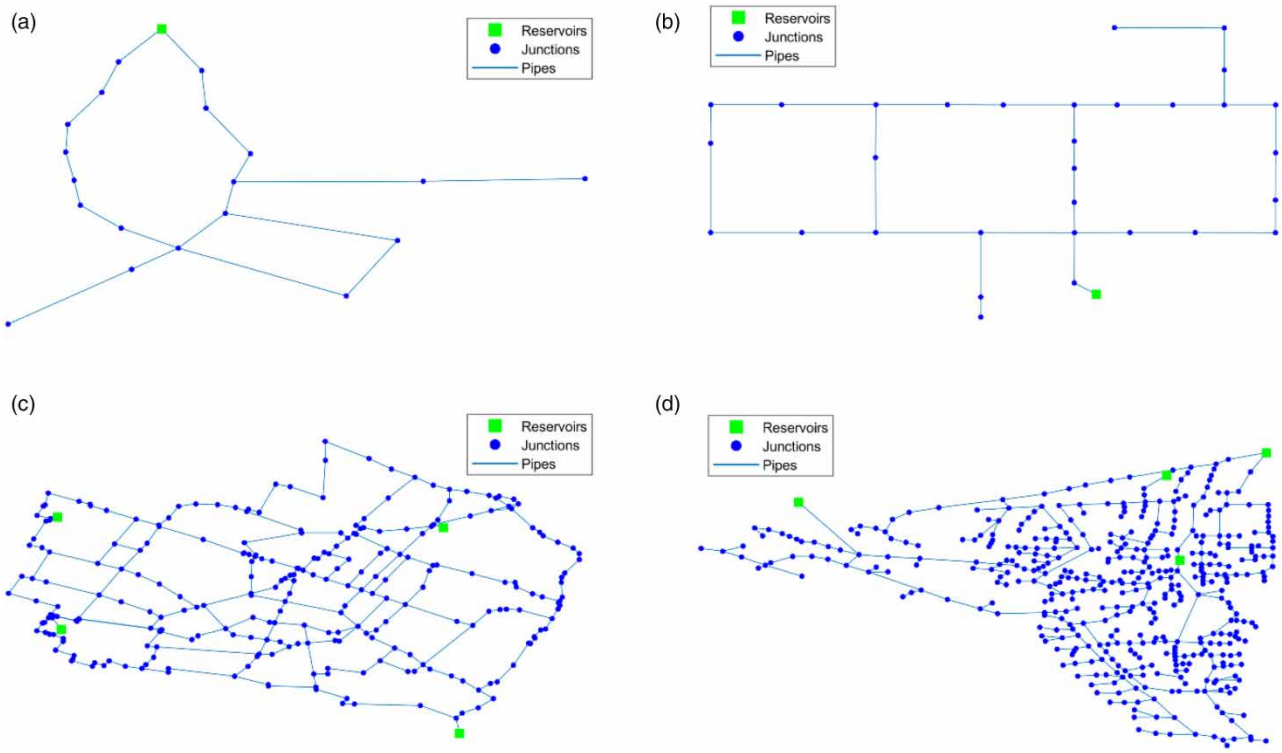

Figure 4 | Network topology: (a) NYT, (b) HAN, (c) MOD, and (d) BAL.

described in Section 3. The results are analyzed in terms of the vulnerability curve, the VPM, and the execution time for each centrality measure to detect the centrality that provides the best prediction for the most damaging attack profile. The calculation of the water flows was made assuming an initial condition, and the variability of the demand is not taken into account. This operating condition can be easily changed to predict the VPM under future circumstances.

6.1. Attacks on junctions

For the attack profiles performed in the junctions of the WDS, the histogram of the centrality measures calculated in the initial operative point is shown in Figure 5. It is noted that degree, eigenvector, and Katz centralities present a similar behavior, where most of the junctions have a small centrality but there are a few that have larger centralities, while for the rest of the centralities there is not a large deviation among all the junctions, which have small and similar centrality measures.

Figure 6 presents three examples of the vulnerability curve for the MOD benchmark, where is it noted that the IMDEF and the IMCEF attack strategies have similar behavior with a steep curve that rises to 100% of unsatisfied water demand after removing less than 20% of the junctions. Meanwhile, the RMCEF strategy for this network had a less steep curve, only reaching 100% of unsatisfied demand after removing more than 80% of the junctions.

Results from VPM and execution times are shown in Table 4 for each experiment using the different fault strategies. It is noted that the IMDEF strategy presented the highest values of VPM in most of the benchmarks, which is expected since this algorithm finds the nodes of the network with the highest impact in the unsatisfied demand. In terms of the strategies based on centrality measures, degree, eigenvector, and Katz centralities also presented the highest VPM values. Furthermore, it is evident in Figure 7 that the VPM for the experiments with degree and Katz centralities were always close to the IMDEF VPM.

Regarding the execution times of each strategy, it is noted that the IMDEF had the highest times, which is expected given the quadratic nature of the algorithm. The RMCEF and the IMCEF presented similar execution times for the medium-size problems (NYT and HAN), but for the larger systems, the IMCEF was more effective. Although the IMCEF algorithm recalculated the centrality measures in each iteration, it reaches the maximum unsatisfied water demand much faster. The shortest execution times were obtained with the degree centrality.

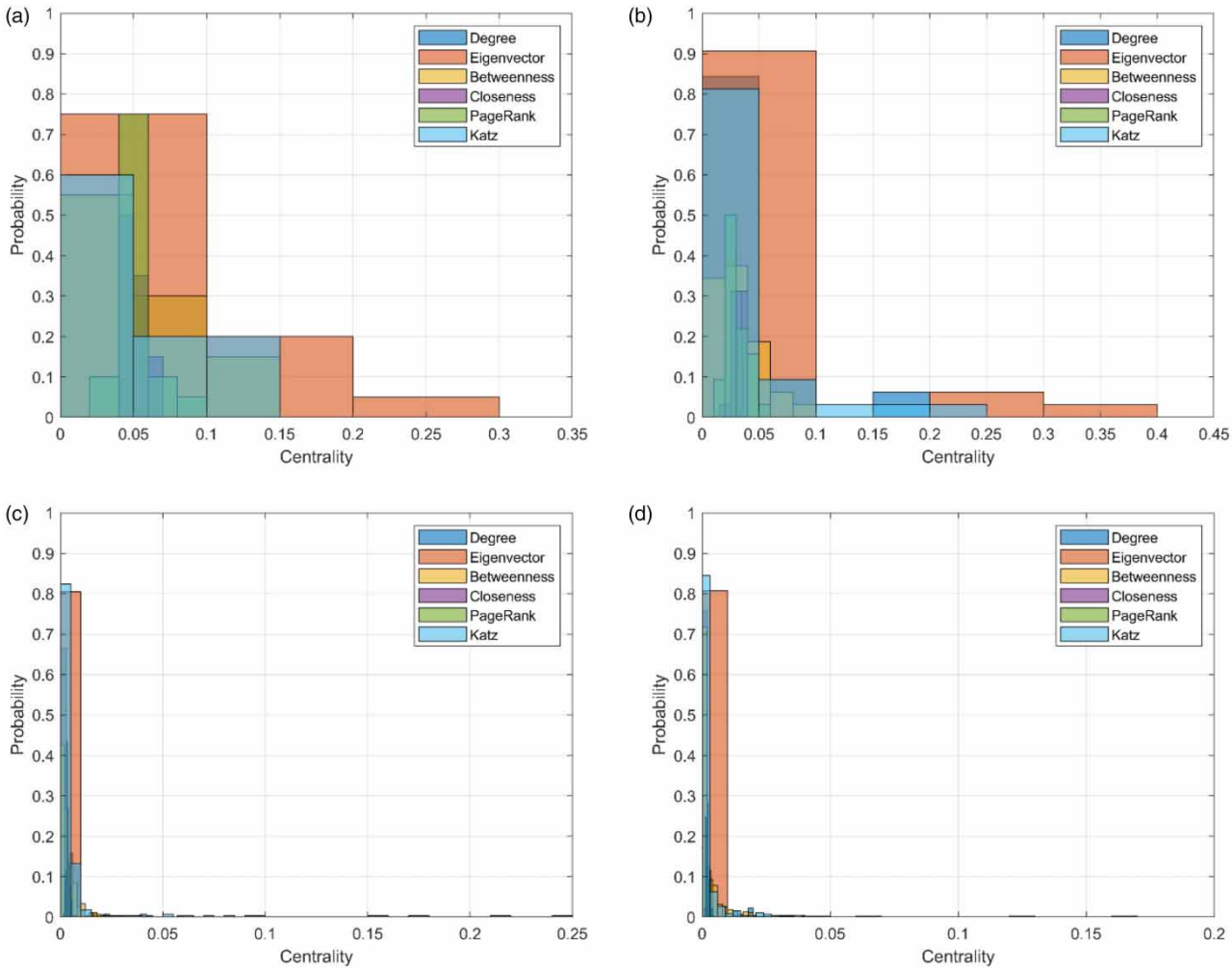

Figure 5 | Centralities for Junctions: (a) NYT, (b) HAN, (c) MOD, and (d) BAL.

6.2. Attacks on pipes

The three attack strategies (IMDEF, RMCEF, and IMCEF) were also applied to pipes, by employing the transformation to the line graph detailed in Section 4. The histograms of the pipe centrality measures for the benchmarks is shown in Figure 8, where it can be noted that the behavior is similar to the junction centralities, having a larger deviation between values of degree, eigenvector, and Katz centralities than for the rest.

Examples of vulnerability curves are shown in Figure 9 for the MOD benchmark, where it is noted that the steepest curve was obtained for the IMCEF attack strategy, which reaches 100% of unsatisfied water demand before removing 20% of the pipes. This behavior differs from the attacks on junctions, since this time the IMDEF attack strategy does not guarantee that the network reaches the total unsatisfied demand faster, which is because different pipes cause the same damage, and the algorithm chooses one of them without a particular criterion that allows predicting the overall most damaging path.

Examining the VPM values for each experiment, it is noted that the IMCEF attack strategy overall presented equal or higher VPM values than the rest of the strategies for the same WDS. Furthermore, the RMCEF strategy in all the studied cases presented the second highest VPM. The degree, eigenvector, and Katz centralities presented the highest VPM in most of the cases in the same way as for the junction attacks (Table 5). The IMCEF strategy allows high VPM values to be obtained, so it predicts the most damaging attacks on pipes and has the lowest execution times for the largest benchmarks (MOD and BAL).

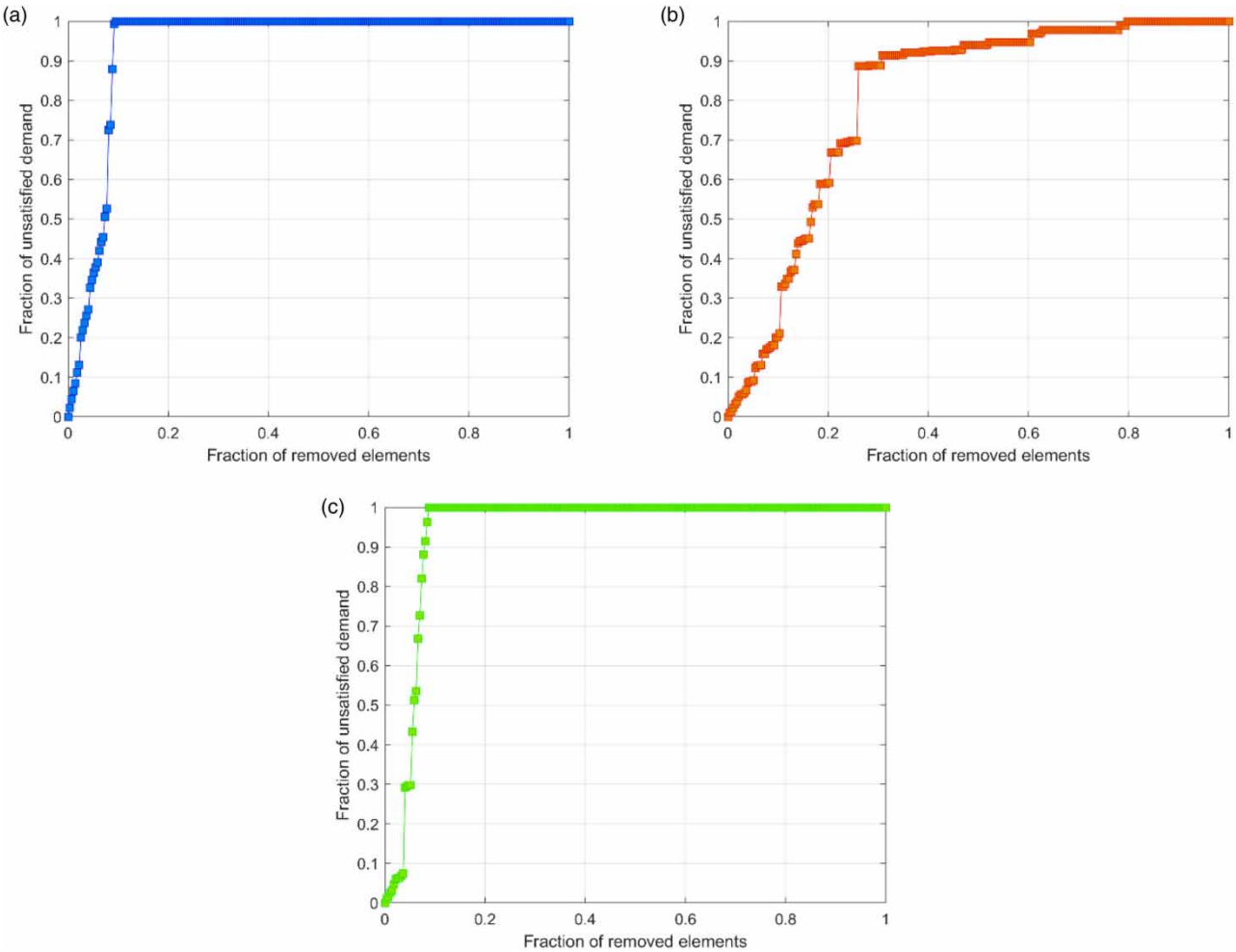

Figure 6 | Vulnerability curves for junction faults in the MOD network: (a) IMDEF, (b) RMCEF, and (c) IMCEF with betweenness centrality.

Table 4 | VPM and execution time from different attack profiles performed on junctions

		VPM				Execution time (s)			
Strategy	**Centrality**	**NYT**	**HAN**	**MOD**	**BAL**	**NYT**	**HAN**	**MOD**	**BAL**
IMDEF	–	0.975	0.984	0.939	0.995	26.504	39.319	10,012.663	5,980.790
RMCEF	Degree	0.882	0.984	0.966	0.972	0.044	0.006	9.101	93.286
	Eigenvector	0.882	0.984	0.628	0.899	0.046	0.005	120.880	475.353
	Betweenness	0.789	0.952	0.806	0.966	0.158	0.433	114.339	338.913
	Closeness	0.729	0.965	0.608	0.894	0.189	0.276	131.259	510.439
	PageRank	0.814	0.972	0.946	0.974	0.117	0.196	13.857	232.413
	Katz	0.882	0.984	0.962	0.971	0.046	0.006	10.829	99.934
IMCEF	Degree	0.927	0.984	0.987	0.996	0.044	0.014	3.321	16.442
	Eigenvector	0.925	0.984	0.987	0.996	0.071	0.020	3.223	23.794
	Betweenness	0.911	0.983	0.943	0.994	0.078	0.048	21.385	42.349
	Closeness	0.911	0.983	0.934	0.993	0.066	0.044	24.382	36.578
	PageRank	0.860	0.983	0.924	0.982	0.146	0.051	22.207	79.563
	Katz	0.925	0.984	0.987	0.996	0.064	0.026	3.337	22.435

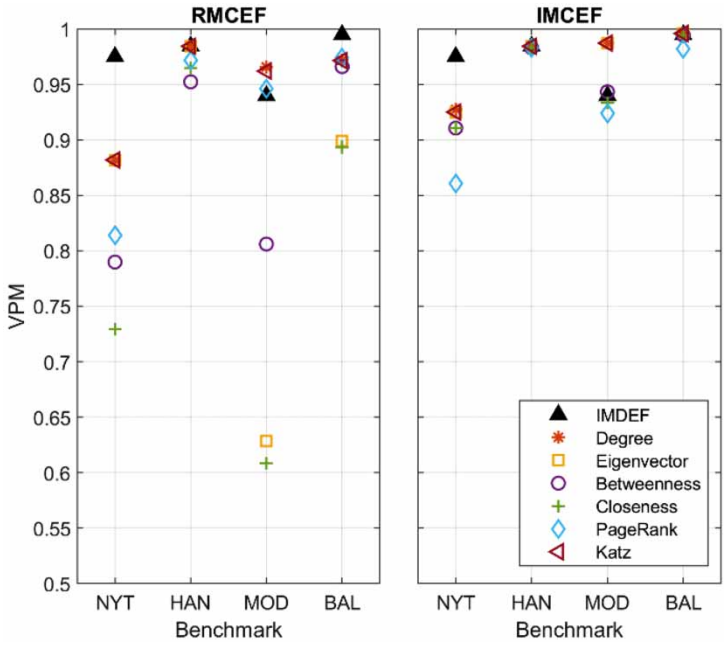

Figure 7 | VPM from attacks based on junction removal under different fault strategies.

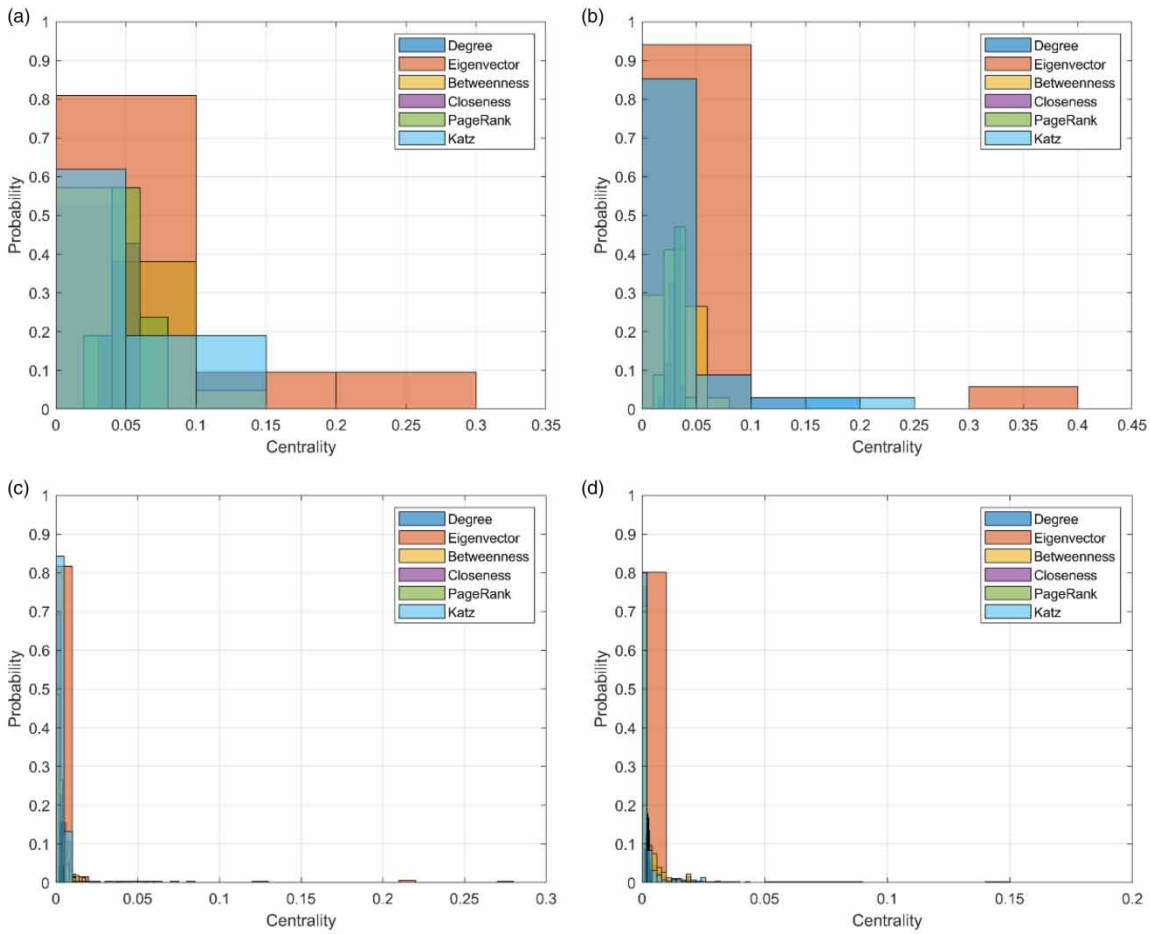

Figure 8 | Centralities for pipes: (a) NYT, (b) HAN, (c) MOD, and (d) BAL.

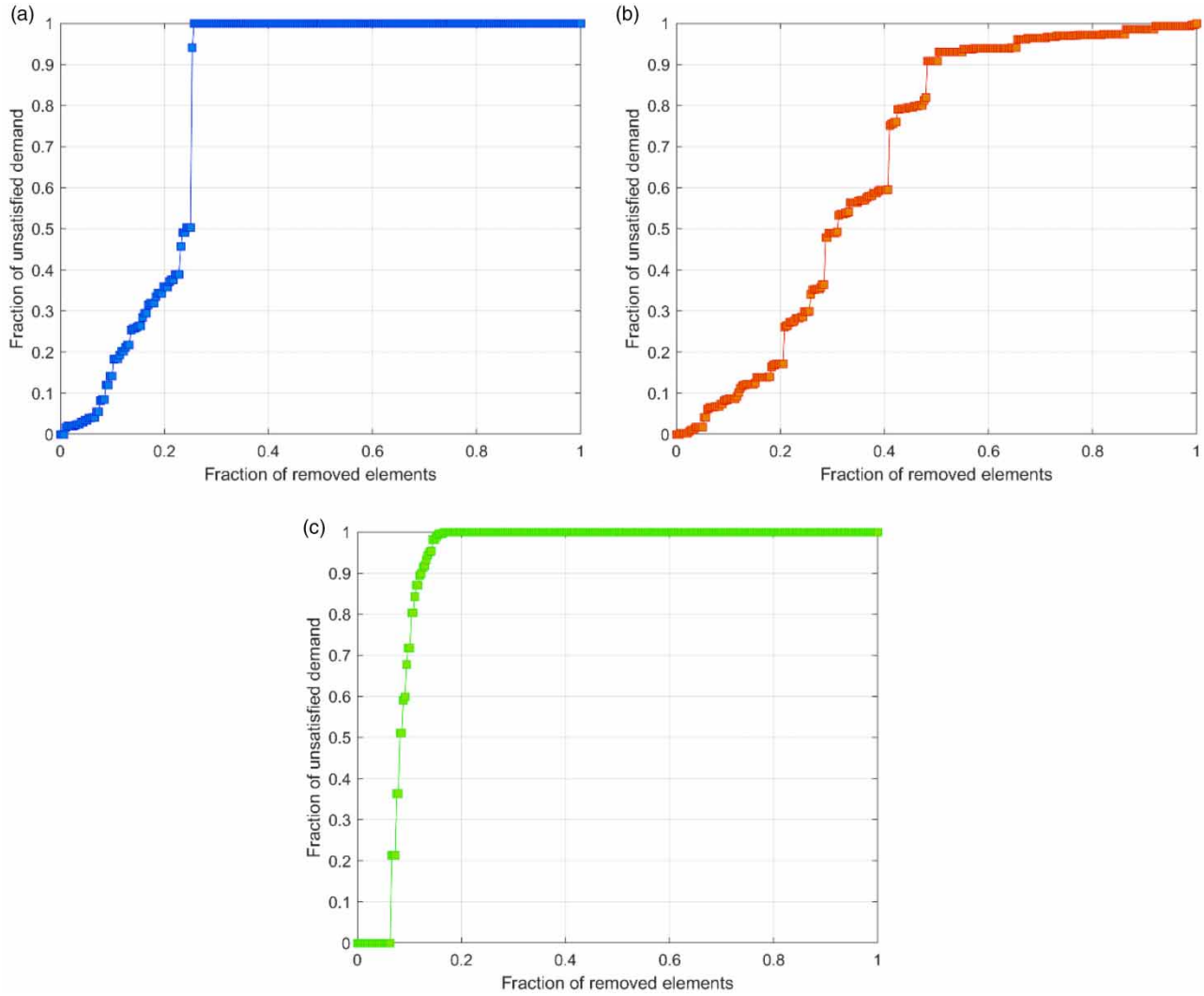

Figure 9 | Vulnerability curves for pipe faults in the MOD network: (a) IMDEF, (b) RMCEF, and (c) IMCEF with betweenness centrality.

Table 5 | VPM and execution time from different attack profiles performed on pipes

		VPM				Execution time (s)			
Strategy	**Centrality**	**NYT**	**HAN**	**MOD**	**BAL**	**NYT**	**HAN**	**MOD**	**BAL**
IMDEF	–	0.845	0.984	0.801	0.967	95.833	42.097	32,703.512	31,233.684
RMCEF	Degree	0.839	0.984	0.942	0.962	0.059	0.036	17.560	134.336
	Eigenvector	0.839	0.984	0.601	0.896	0.053	0.033	132.849	421.308
	Betweenness	0.738	0.779	0.666	0.949	0.188	0.651	162.252	498.101
	Closeness	0.621	0.895	0.546	0.874	0.246	0.463	157.488	525.757
	PageRank	0.673	0.914	0.820	0.964	0.176	0.636	158.097	492.595
	Katz	0.839	0.984	0.945	0.962	0.054	0.035	15.776	134.281
IMCEF	Degree	0.922	0.984	0.976	0.995	0.087	0.054	8.009	14.194
	Eigenvector	0.922	0.984	0.967	0.994	0.092	0.056	11.954	17.798
	Betweenness	0.896	0.907	0.911	0.992	0.109	0.217	36.029	29.919
	Closeness	0.848	0.907	0.886	0.985	0.133	0.203	44.062	44.049
	PageRank	0.702	0.955	0.815	0.982	0.211	0.082	62.484	55.829
	Katz	0.922	0.984	0.973	0.994	0.259	0.068	9.922	17.768

7. DISCUSSION

Analyzing the results obtained for attack profiles on junctions and pipes of the benchmarks, it is noted that the IMDEF strategy worked better for attacks on junctions by predicting the most damaging attacks; however, it presented high execution times and it did not predict the highest VPM for the attacks on junctions. Therefore, the IMCEF strategy was more efficient for attacks in pipes and junctions having high VPM and lower execution times for the larger WDS.

In terms of the centrality measures, the behavior was similar to that obtained for electric power systems in previous works, where the degree, eigenvector, and Katz centralities are representative of the elements that generate the highest impact in the unsatisfied demand. These centralities allow the identification of important elements in the grid considering the information on the calculated water flow.

These experiments assume that the junctions and pipelines can be disconnected with isolation valves, but in practical scenarios, the disconnection will occur through segment networks. That condition can be taken into account for extending the work to more realistic conditions and evaluating the vulnerability in the segment networks using models with data from isolation valves.

8. CONCLUSIONS

The main contribution of this work is to apply the water flows to the calculation of centrality measures for evaluating the vulnerability in different WDS benchmarks. The vulnerability framework is based on a graph model integrating the topological structure of the network and the operational water flow calculated with EPANET and MATLAB.

The vulnerability framework determines the vulnerability curve and the VPM according to an attack profile based on three different strategies. The RMCEF consists of removing the elements according to the centrality measured calculated in the initial operating point, the IMCEF removes the most central element after recalculating the centrality for each stage, and the IMDEF strategy consists of removing the MDE first. These strategies were applied to junctions and pipes of four different WDS benchmarks: the NYT, the HAN, the MOD, and the BAL networks.

After analyzing the results of all the experiments performed, it is concluded that the IMCEF had the best behavior for evaluating the most damaging attacks on the WDS, since it presented high VPM values both for removing junctions and pipes. Moreover, it presented low execution times for the largest benchmarks. This indicates that the recalculation of the centrality measures after removing the most central elements helps to identify new important elements and the initial centrality measures lose relevance after such removal.

Another important contribution is identifying the centralities that are more predictive in terms of the MDEs, which were the degree, eigenvector, and Katz centralities under the conditions of this research. The degree centrality additionally presents the advantage that its calculation is simpler; thus, it is recommended for the larger WDS.

It was demonstrated that the vulnerability framework can be applied to the WDS in the same way in which it was previously applied to electric power systems. This work can be extended to consider the variability of the demand, allowing us to obtain the vulnerability under different operating conditions and the segment graph. Future work will be oriented to the design of the WDS by optimization by minimizing the vulnerability of the system under attacks that can be generated by droughts and other extreme weather conditions.

ACKNOWLEDGEMENT

The authors extend their appreciation to the deputyship for the Research & Innovation Ministry of Education in Saudi Arabia for funding this research work through the project number (IFP-202O.13).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Agathokleous, A., Christodoulou, C. & Christodoulou, S. E. 2017 Topological robustness and vulnerability assessment of water distribution networks. *Water Resour. Manage.* **31** (12), 4007–4021.

Albarakati, A. 2021a Evaluation of the most harmful malicious attacks in power systems based on a new set of centralities. *J. Electr. Eng. Technol.* **16**, 1929–1939.

Albarakati, A. 2021b Influence of generation and load variations in the vulnerability of power systems. *J. Electr. Eng. Technol.* **16**, 2397–2406.
Albarakati, A. & Bikdash, M. 2021 Empirical electrical-based framework to judge the ability of centrality measures in predicting grid vulnerability. *J. Electr. Eng. Technol.* **16**, 1917–1927.
Albarakati, A., Bikdash, M. & Dai, X. 2017 Line-graph based modeling for assessing the vulnerability of transmission lines. In: *SoutheastCon 2017*. IEEE, pp. 1–8.
Alexiou, D. & Tsouros, C. 2017 Design of an irrigation network system in terms of canal capacity using graph theory. *J. Irrig. Drain. Eng.* **143** (6), 06017002.
Barabási, A.-L. 2013 Network science. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **371** (1987), 20120375.
Barabási, A.-L. & Albert, R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512.
Bloch, F. & Jackson, M. O. 2021 Centrality measures in networks. *SSRN Electron. J.* Available at SSRN: https://ssrn.com/abstract=2749124 or http://dx.doi.org/10.2139/ssrn.2749124.
Bonacich, P. 2007 Some unique properties of eigenvector centrality. *Social Networks* **29** (4), 555–564.
Bragalli, C., Ambrosio, C. D., Lee, J., Lodi, A. & Toth, P. 2008 IBM research report water network design by MINLP water network design by MINLP. Rep. No. RC24495, IBM Research, Yorktown Heights, NY.
Cabrera, E. & García-Serra, J. 1999 *Drought Management Planning in Water Supply Systems*, Vol. 32, no. 9. Springer, The Netherlands.
Centre for Water Systems University of Exeter 2014 *Towards the Best-Known Approximation to the True Pareto Front*.
Christie, M., Cliffe, A., Dawid, P. & Senn, S. (eds) 2011 *Simplicity, Complexity and Modelling*. John Wiley & Sons, Ltd, Chichester, UK.
Eliades, D. G., Kyriakou, M., Vrachimis, S. G. & Polycarpou, M. M. 2016 EPANET-MATLAB Toolkit: an open-source software for interfacing EPANET with MATLAB. *Comput. Control Water Ind.* **8**, 1–8.
Giudicianni, C., Di Nardo, A., Di Natale, M., Greco, R., Santonastaso, G. F. & Scala, A. 2018 Topological taxonomy of water distribution networks. *Water* **10** (4), 1–19.
Giudicianni, C., Herrera, M., Di Nardo, A., Greco, R., Creaco, E. & Scala, A. 2020 Topological placement of quality sensors in water-distribution networks without the recourse to hydraulic modeling. *J. Water Resour. Plan. Manage.* **146** (6), 04020030.
Giudicianni, C., Di Nardo, A., Greco, R. & Scala, A. 2021 A community-structure-based method for estimating the fractal dimension, and its application to water networks for the assessment of vulnerability to disasters. *Water Resour. Manage.* **35** (4), 1197–1210.
Giustolisi, O., Ridolfi, L. & Simone, A. 2019 Tailoring centrality metrics for water distribution networks. *Water Resour. Res.* **55** (3), 2348–2369.
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. *Math. Comput. Modell.* **57** (7–8), 1853–1859.
Jeong, G., Lim, G. & Kang, D. 2021 Identification of unintended isolation segments in water distribution networks using a link-by-link adjacency matrix. *J. Water Resour. Plan. Manage.* **147** (2), 06020013.
Katz, L. 1953 A new status index derived from sociometric analysis. *Psychometrika* **18** (1), 39–43.
Kim, J. H., Kim, T. G., Kim, J. H. & Yoon, Y. N. 1994 A study on the pipe network system design using non-linear programming. *J. Korean Water Resour. Assoc.* **27** (4), 59–67.
Mortula, M. M., Ahmed, M. A., Sadri, A. M., Ali, T., Ahmad, I. & Idris, A. 2020 Improving resiliency of water supply system in arid regions: integrating centrality and hydraulic vulnerability. *J. Manage. Eng.* **36** (5), 05020011.
Murthy, A. L. N. & Murthy, G. S. R. 2012 A network flow model for irrigation water management. *Algorithmic Oper. Res.* **7** (2), 83–93.
Nicolini, M. 2020 Complex networks theory for evaluating scaling laws and WDS vulnerability for potential contamination events. *Water* **12** (5), 1296.
Page, L., Brin, S., Motwani, R. & Winograd, T. 1999 *The PageRank Citation Ranking: Bringing Order to the Web*. Stanford InfoLab.
Reca, J. & Martínez, J. 2006 Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resour. Res.* **42** (5), 1–9.
Riquelme, F., Gonzalez-Cantergiani, P., Molinero, X. & Serna, M. 2018 Centrality measure in social networks based on linear threshold model. *Knowledge-Based Syst.* **140**, 92–102.
Schaake, J. C. & Lai, F. H. 1969 *Linear Programming and Dynamic Programming Application to Water Distribution Network Design*. M.I.T. Hydrodynamics Laboratory, MIT, Cambridge, MA, USA.
Shuang, Q., Zhang, M. & Yuan, Y. 2014 Node vulnerability of water distribution networks under cascading failures. *Reliab. Eng. Syst. Saf.* **124**, 132–141.
The Mathworks Inc. 2020 *Graph and Network Algorithms*.
United Nations 2020 *Goal 6: Ensure Access to Water and Sanitation for All*.
Wéber, R., Huzsvár, T. & Hős, C. 2020 Vulnerability analysis of water distribution networks to accidental pipe burst. *Water Res.* **184**, 116178.
Yazdani, A. & Jeffrey, P. 2012 Water distribution system vulnerability analysis using weighted and directed network models. *Water Resour. Res.* **48** (6), 1–10.
Yoshida, T. 2013 Weighted line graphs for overlapping community discovery. *Soc. Netw. Anal. Min.* **3** (4), 1001–1013.

First received 5 August 2021; accepted in revised form 1 October 2021. Available online 14 October 2021