



## VULNERABILITY ASSESSMENT OF WATER DISTRIBUTION NETWORKS: GRAPH THEORY METHOD

R. Sheikholeslami<sup>1</sup> and A. Kaveh<sup>2,3\*,†</sup>

<sup>1</sup>*Department of Civil and Environmental Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran*

<sup>2</sup>*Centre of Excellence for Fundamental Studies in Structural Engineering, Iran University of Science and Technology, Narmak, Tehran-16, Iran*

<sup>3</sup>*Guest Professor, Institute for Mechanics of Materials and Structures, Vienna University of Technology, Karlsplatz 13, A-1040 Wien, Austria*

### ABSTRACT

The main functional purpose of a water distribution network is to transport water from a source to several domestic and industrial units while at the same time satisfying various requirements on hydraulic response. All the water distribution networks perform two basic operations: firstly the water network needs to deliver adequate amounts of water to meet specific requirements, and secondly the water network needs to be reliable; therefore, the required amount of water needs to be continuously available 24 hours a day and 365 days per year. Due to the inevitable failures of some components such as pump stations, reservoirs and/or pipelines in a large-scale water distribution network, in designing a reliable network, the topological structure with low vulnerability must be achieved. Consequently, the study of connectivity, which is the key graph-theoretical notion, becomes crucial. This paper highlights some fundamental concepts from graph theory for vulnerability assessment of water distribution networks, addresses the mathematical properties of the link and node-deletion problems, and outlines some well-established results on the deterministic measures to assess the fault tolerance of networks.

**Keywords:** water distribution networks; connectivity; system reliability; fault-tolerant networks.

Received: 10 March 2015; Accepted: 25 May 2015

---

\*Corresponding author: Iran University of Science and Technology, Narmak, Tehran-16, Iran

†E-mail address: alikaveh@iust.ac.ir (A. Kaveh)

## 1. INTRODUCTION

Water distribution networks are one of the main components of urban infrastructural systems. The pipeline networks are complex systems that require a high level of investment for their construction and maintenance. Failures of some components and communication lines in the municipal water supply system, which often result the catastrophic consequences, are inevitable. The ability of a faulty network to continue to operate at an acceptable but lower level of performance is referred to as fault tolerance. Generally, a fault-tolerant network with some damaged components or links can usually continue to operate with a smaller throughput.

Risk-based management of these networks encourages water utilities to consider reliability analysis and vulnerability assessment of the system which is useful for achieving robust network management [1]. The reliability analysis of complex water networks is basically investigated by considering two categories of failure events: physical failures and design errors. Most physical failures have mechanical causes such as component wear, pipe breaks, etc, while the latter is mainly due to design errors. Thus, in the design of a water network one of the most fundamental considerations is the physical reliability which can be usually characterized by connectivity of the topological structure of a network.

Physical reliability problems have been broadly investigated in the literature [2,3]; however, assessing vulnerability and robustness of water distribution networks have not been systematically exposed to the analyses by graph theory [4]. On the other hand, the reliability-based analysis and design of water distribution networks have been extensively dealt with using simulations, optimization algorithms and non-deterministic techniques [5]. By way of example, Giustolisi et al. [1] proposed an algorithm for automatic identification of topological changes in water distribution networks coupled with a pressure-driven simulation model. Shuang et al. [6] described a methodology for simulation of cascading dynamics of water distribution network in failure condition and the detection of crucial pipes.

There are few works in graph theory-based reliability analysis of water distribution systems, among them are the studies by Jacobs and Goulter [7,8] who showed that the “regular graphs” with equal number of links incident to each node are most invulnerable to failures. Also Goulter [9] suggested that future work into the definition of suitable topology-based reliability analysis of water distribution networks will use fundamental graph theory concepts. The connectivity and reachability criteria, taken from graph theory, were suggested by Wagner et al. [10] to assess the reliability of a water network. Kessler et al. [11] developed a graph-theoretical methodology for optimal cost design of invulnerable water networks by incorporating reliability in the design procedure. However, the application of purely topological graph theory in vulnerability assessment of water distribution networks was shown to have limited extent. Yazdani and Jeffrey [4] utilized several spectral and statistical measurements of complex networks to quantify vulnerability-related topological properties of water distribution networks. They concluded that networks with smaller graph diameter, higher level of redundancy, and more lattice-like structures are connected in a more optimal configuration; hence, perfect grids and regular graphs are the most invulnerable networks in this respect. Perelman and Ostfeld [12] introduced a new clustering algorithm for topological analysis of large-scale water supply systems. The main assumption in [12] is that a network can be decomposed into some sub-networks with

smaller scales, and sub-networks may split further or merge into another sub-network with a larger scale. Ying et al. [13] proposed a framework for topological analysis of the pipeline networks with variable connectivity. They developed several procedures to derive the incidence matrix and the fundamental circuit matrix of a graph or subgraph directly from those of its parent subgraphs or graph, without the time-consuming inversion of the incidence matrix for determining the fundamental circuits. Gutiérrez-Pérez et al. [14] developed a methodology based on spectral measurements of graph theory to achieve an efficient vulnerability analysis of water distribution systems.

There are a number of measures for fault tolerance assessment of networks which have been developed by several researchers. These quantitative descriptions can be classified into: deterministic measures and probabilistic measures. The first can be obtained using graph-theoretical concepts such as connectivity, while the second utilizes the failure probabilities of the links or nodes in a network. Although the probabilistic measures of fault tolerance seem to model the real-world networks accurately, it has been revealed that all probabilistic measures are belong to the class of non-deterministic polynomial-time hard problems known as NP-hard [15]. The NP-hard means that for an  $n$ -pipe network, the computational time required for a rigorous algorithm is at best an exponential function of  $n$  and is thus enormous even for relatively small networks. For general review of different variants of the probabilistic measures and discussion on their performance in reliability analysis of water distribution networks the reader is referred to [16].

The deterministic measures are generally grouped into two categories of statistical and spectral measurements. Statistical metrics are those which quantify organizational properties of the network based on the most frequent ideas and structural patterns ranging from more basic metrics such as network size, network order, link density, number of independent loops or cycles, and graph diameter, to other metrics including degree distribution, average node-degree, clustering coefficient, connectivity, and edge-connectivity. Spectral metrics are obtained from the spectrum of network adjacency matrix and Laplacian matrix. The second smallest eigenvalue of Laplacian matrix is one of the most important spectral metrics which has many properties and can be used in order to quantify the strength of network connections [4].

By deeper understanding of the topological structure of a pipeline network and identifying the most crucial components in a network, one can effectively protect the network to prevent failures and build a fault-tolerant network. Given this, the main objectives of current study are: (i) to facilitate a deeper understanding of vulnerability analysis of water distribution networks from a graph-theoretical point of view; (ii) to formulize link and node-deletion problems as standard concepts of graph theory; and (iii) to present some basic results on the fault tolerance assessment of water networks.

## 2. TOPOLOGICAL PROPERTIES OF WTER DISTRIBUTION NETWORKS

Several aspects of our daily life depend excessively on large-scale infrastructural systems such as water networks, road networks, telecommunication networks, world-wide web, etc. Nowadays, as society depends critically on these complex systems, their survivability and vulnerability assessment deserves more attention [17]. Accordingly, engineers encounter

with the most significant challenge, that is, they are required not only to design reliable, convenient, and economical water networks, but to provide various supporting softwares and algorithms matching with them. However, the topological structure in a water network plays a crucial role in the optimal design of these systems. In this section some basic principles and guidelines, for designing reliable water networks, are highlighted from a graph-theoretical perspective.

### 2.1 Mathematical model of water distribution networks

A water network with a specified configuration can be considered as a directed graph with pipes being the members and junctions of pipes being the nodes. Nodes for normal pipe joints are called fixed demand nodes, whereas tanks, sources and destinations are represented by fixed head nodes with constant pressures or pressures specified outside the network. As a result of this interconnection, the variables specifying the behavior of the individual members are inter-related. In order to get familiar with the topological properties of water networks, an introduction to the graph theory seems to be necessary.

A simple graph  $S$  is defined as a set  $N(S)$  of nodes and a set  $M(S)$  of members together with a relation of incidence that associates two distinct nodes with each member, known as its ends. A path is a finite sequence  $P_k = \{n_0, m_1, n_1, \dots, m_p, n_p\}$  whose terms are alternately nodes  $n_i$  and members  $m_i$  of  $S$  for  $1 \leq i \leq p$ , and  $n_{i-1}$  and  $n_i$  are two ends of  $m_i$  provided that no node appears more than once. The distance between nodes of a graph is defined as the number of members of a shortest path between these nodes. The diameter of  $S$ , denoted by  $d(S)$ , is defined as the maximum distance from the shortest path between all pairs of nodes.

A subgraph  $S_i$  of a graph  $S$  is a graph for which  $N(S_i) \subseteq N(S)$  and  $M(S_i) \subseteq M(S)$  and each member of  $S_i$  has the same ends as in  $S$ . A tree  $T$  of  $S$  is a subgraph of  $S$  which has no cycle, a cycle being a closed path. The complements of members of  $T$  in  $S$  are the chords of this tree. If a tree contains all the nodes of  $S$ , it is called a spanning tree of  $S$ . A cycle is a path  $\{n_0, m_1, n_1, \dots, m_p, n_p\}$  for which  $n_0 = n_p$  and  $p \geq 1$ . A cutset in a graph  $S$  is a set of members whose removal from the graph increases the number of connected components of  $S$  [18].

Consider a spanning tree of a graph. Such a tree comprising of  $N(S) - 1$  members is a maximal subgraph which is free from any cycle. Inclusion of any of  $M(S) - N(S) + 1$  members of the original graph that are not contained in the tree will produce a new cycle. The cycles obtained in this way are independent because each non-tree member is exactly contained in one cycle. These cycles form a basis known as a fundamental cycle basis. The cycle-member incidence matrix  $\bar{\mathbf{C}}$  of a graph  $S$ , has a row for each cycle and a column for each member. An entry  $\bar{c}_{ij}$  of  $\bar{\mathbf{C}}$  is 1 if cycle  $C_i$  contains member  $m_j$  and it is 0 otherwise. For a graph  $S$  there exists  $2^{M(S) - N(S) + 1} - 1$  cycles. However, for a cycle basis, a cycle-member incidence matrix becomes a  $(M(S) - N(S) + 1) \times M(S)$  matrix, denoted by  $\mathbf{C}$ , known as the cycle basis incidence matrix of  $S$  [18]. In order to automatic formation of  $\mathbf{C}$  or the identification of a set of independent loops in a network, the reader is referred to [19].

Also a graph can be represented as a node-member incidence matrix. The node-member incidence matrix  $\bar{\mathbf{B}}$  is an  $N(S) \times M(S)$  matrix in which the entry in row  $i$  and column  $j$  is 1 if node  $n_i$  is incident with member  $m_j$ , and is 0 otherwise. The rows of  $\bar{\mathbf{B}}$  are dependent and one row can arbitrarily be deleted to ensure the independence of the rest of the rows. The

node corresponding to the deleted row is called a *datum (reference) node*. The matrix obtained after deleting a dependent row is called an *incidence matrix* of  $S$  and it is denoted by  $\mathbf{B}$  [20].

In a directed graph each member is assigned an orientation. A member is oriented from its initial node to its final node where the initial node is said to be positively incident on the member and the final node negatively incident. All the matrices  $\bar{\mathbf{B}}$ ,  $\mathbf{B}$ ,  $\bar{\mathbf{C}}$  and  $\mathbf{C}$  can be defined as before, with the difference of having +1, -1 and 0 as entries, according to whether the member is positively, negatively and zero incidents with a node or a cycle [20].

In algebraic topology, one can assign mathematical objects such as numbers to nodes (0-simplexes), members (1-simplexes), cutsets and cycles of a graph (1-complex). This leads to an algebraic structure associated with a graph. Consider a graph model of a water network, for this graph the head vector  $\mathbf{h}$  can be defined as following:

$$\mathbf{h} = \{h_j\} = \{h_1, h_2, \dots, h_{N(S)}\} \tag{1}$$

in which  $h_j$  is the head at the  $j$ th node vertically measured from a reference plane; and  $N(S)$  is number of nodes.

The head difference vector  $\mathbf{H}$ , can be expressed by:

$$\mathbf{H} = \{H_i\} = \{h_j - h_k\}, \quad i = 1, 2, \dots, M(S) \tag{2}$$

in which  $h_j$  and  $h_k$  is the heads at the upstream ( $j$ th node) and downstream ( $k$ th node) ends of the  $i$ th pipe, respectively.

The head difference vector  $\mathbf{H}$  is associated with the head vector  $\mathbf{h}$  and the node-member incidence matrix  $\mathbf{B}$  as:

$$\mathbf{H} = \mathbf{B}^t \mathbf{h} \tag{3}$$

The link flow vector  $\mathbf{q}$  can be defined as:

$$\mathbf{q} = \{q_i\} = \{q_1, q_2, \dots, q_{M(S)}\} \tag{4}$$

where  $q_i$  is the outflow from  $i$ th link/pipe.

By defining the head vector  $\mathbf{h}$ , the head difference vector  $\mathbf{H}$ , and the link flow vector  $\mathbf{q}$  for a pipeline network, the postulates of Kirchhoff can be written in the following form [21]:

**Postulate 1:** the circuit (cycle basis incidence) matrix  $\mathbf{C}$  and the head difference vector  $\mathbf{H}$  are orthogonal:

$$\mathbf{CH} = \mathbf{0} \tag{5}$$

**Postulate 2:** the node-member incidence matrix  $\mathbf{B}$  and the flow vector  $\mathbf{q}$  are orthogonal:

$$\mathbf{B}\mathbf{q} = \mathbf{0} \quad (6)$$

Eq. (5) is the energy conservation law in each closed loop or between every two nodes of known head. The conservation of energy requires that the sum of the line losses (friction losses) and the minor losses (other losses due valves, etc.) over any path, minus any energy added to the liquid by a pump, minus the difference in grade between the two points of known energy is equal to zero. Application of conservation of energy results in a set of nonlinear equality constraints in terms of flow rates. Eq. (6) expresses the equilibrium equation (mass conservation law) which states that the sum of flows into or out of any junction node minus any external demand must be equal to zero. Application of nodal conservation of mass results in a set of linear equality constraints in terms of pipe flow rates.

At this point, the purely topological aspects of the algebraic structure that characterizes a water network have been described. Two dual sets of variables  $\mathbf{q}$  on one hand, and  $\mathbf{H}$  on the other hand, have been independently studied. However, a distinct characteristic of a network problem is the existence of a relationship between the two sets of variables, expressed as:

$$\mathbf{H} = \mathbf{R}\{\mathbf{q}^\alpha\} \quad (7)$$

with  $\mathbf{H}$ ,  $\mathbf{q}$ ,  $\mathbf{R}$ , and  $\alpha$  are respectively the pipe head loss vector, link flow vector, resistance coefficient (diagonal) matrix, and flow exponent. There are different values for the resistance coefficient and flow exponent depend on the particular head loss formula. The hydraulic head loss by water flowing in a pipe due to friction with the pipe walls can be computed using one of three different formulas: Hazen–Williams, Darcy–Weisbach and Chezy–Manning.

Eq. (5-7) represent sufficient and necessary conditions for the solvability of the network flow problem. The topological and physical properties of a network model are all contained in these equations, thus they provide enough tools for the hydraulic simulation of the water distribution networks. The resulting system of equations is therefore:

$$\begin{bmatrix} \mathbf{B} \\ \mathbf{CR}\mathbf{q}^{\alpha-1} \end{bmatrix} \{\mathbf{q}\} - \begin{bmatrix} \mathbf{Q} \\ \mathbf{null}(p,1) \end{bmatrix} = \mathbf{0} \quad (8)$$

where  $\mathbf{Q}=\{Q_j\}$  ( $j = 1, 2, \dots, N(S)$ ) is the demand in each node (external inflow or outflow), and  $\mathbf{null}(p,1)$  is a  $p \times 1$  zero vector with  $p$  being the  $M(S) - N(S) + 1$  (number of loops). This zero vector indicates that in each loop the summation of pipe's head losses should be zero.

## 2.2 Topological structure of water distribution networks: design principles

When the mathematical model of a water network can be expressed in terms of graph-theoretic notions, the problem of the optimal network design becomes: "Find a graph  $S$  subject to some specified constraints." This problem can be formulated more practically as a multi-criteria optimization problem.

In choosing a proper topological structure for water distribution networks two

optimization criteria must be taken into account: maximum reliability and minimum cost. If one can fulfill these conditions the resulting network is a least-cost invulnerable water distribution network [11]. Historically, many optimization techniques have been applied to water network least-cost design, such as linear programming [22], nonlinear programming [23], and meta-heuristic algorithms [24,25]. The aim of this paper is to consider the topological invulnerability of the water distribution networks, though there are many other parameters influencing reliability and cost of the networks.

As mentioned in section 2.1, topological structure of a network is a graph. In this subsection, using the terminology of graph theory, some primary principles are presented in order to design a reliable water network. Following Ref [26], these conditions can be stated as follows:

**Maximum fault tolerance:** The water distribution network must continue to be functional in case of node (i.e. vertex or junction) or link (i.e. edge, member, or pipe) failures. Different definitions of fault tolerance exist, the simplest one corresponding to connectivity (or edge-connectivity) of the graph, is the minimum number of components (nodes or links) which must be removed in order to eliminate all paths between a pair of nodes. The maximum connectivity is valuable; since, it corresponds to not only the maximum fault tolerance of the water network but also the maximum number of internally (or edge-) disjoint paths between any two distinct nodes.

**Embedding of other topologies or rehabilitation capability:** This important characteristic deals with the ability of a given topology to match the replacement (rehabilitating, duplicating, or repairing) of various components in a network which improves the performance of the water supply system. In other words, when a graph is used as an interconnection network, it should contain certain subgraph structures. The existence of these structures has special importance for executing certain algorithms.

**Extendibility:** It should be possible to construct a larger water networks from existing one. When a small pipeline network is extended, some desirable properties should be remained and some useful parameters should be calculated easily.

**Symmetry and regularity:** In a complex network if almost all the components behave in the same manner and communicate in similar ways it implies at least some regularity and some symmetric properties on the graph. A highly symmetry water network is desirable; since, they can be modeled by some efficient and useful algorithms.

In order to design a topological structure for graphs (configuration processing), there are three general techniques, namely *line graphical* method, *Cayley* method and *graph products* method. The line graph method is a very useful technique for constructing a larger graph from a given graph. The graph constructed by this method can easily preserve many valuable characteristics from the primary graph, such as degree, diameter, connectivity, Eulericity, Hamiltonicity, etc. Cayley method introduces a very useful graph-theoretic procedure for designing and analyzing symmetric networks. Cayley method is based on a finite group, and, thus; it is also known as a group-theoretic method or algebraic method in the literature. The graph products method is a very effective method for constructing a larger graph from specified small graphs. The graph constructed by this way can contain the factor graphs as its subgraphs and obtain many significant properties of the factor graphs. A number of important graph-theoretic characteristics, such as degree, diameter and connectivity can be easily extracted from the factor graphs. This method is an important technique for designing large-scale complex networks [26,27].

### 3. ASSESSING VULNERABILITY OF WATER DISTRIBUTION NETWORKS UNDER NODE AND LINK FAILURES

Failures are quite common in water distribution networks. There are various node and link failures due to mechanical, software, and operational problems. As the network size grows, individual failures may become the norm rather than exception. Fault tolerance in water networks requests for both redundancy and robust mechanisms, in physical connectivity and design procedures. Therefore, characterizing damaged or disconnected components in water supply systems is an essential task. Water distribution networks could be potentially weighted bi-directional graphs, given the pipes and nodes physical qualities and the possibility of flow redirection which may take place as a result of the failures of different components. Generally, the direction of network flow and weighting inserted to graph links or nodes is identified by physical and operational parameters and the costs associated with water supply. In networks under study for which no data was available, the networks are treated as undirected and un-weighted graphs as indicated by [4]. In addition, it is beneficial to consider the water networks as planar graphs so that graph links intersect only at a node mutually incident with them. The first necessary condition for an invulnerable water network is the existence of two or more distinct paths between the source and every demand node. This property, known as topological invulnerability, is a well-studied topic in graph theory with respect to graph connectivity [11].

In this section, the link and node failures, as the graph-theoretic problems, are formulized and some basic studies on these problems in graph theory are proposed. Our treatment is based mainly on the Ref [28]. Before going any further, we need to introduce some necessary definitions from graph theory:

**Definition 1.** A graph is *connected* if there exists a walk between every pair of its nodes. A graph that is not connected is called *disconnected*.

**Definition 2.** Let  $S = (N, M)$  be a graph and  $U \subset N$ . The node-deletion subgraph  $S - U$  is the graph obtained from  $S$  by deleting from  $S$  the nodes in  $U$ . That is,  $S - U$  is the subgraph induced on the node subset  $N - U$ . If  $U = \{u\}$ , we simply write  $S - u$ .

**Definition 3.** Let  $S = (N, M)$  be a graph and  $F \subset M$ . The edge-deletion subgraph  $S - F$  is the subgraph obtained from  $S$  by deleting from  $S$  the edges in  $F$ . Thus,  $S - F = (N, M - F)$ . If  $F = \{e\}$ , it is customary to write  $S - e$ .

**Definition 4.** The *connectivity*  $\kappa(S)$  of a graph  $S$  is the minimum number of nodes whose removal from  $S$  generates a disconnected or a trivial graph.

**Definition 5.** The *edge-connectivity*  $\lambda(S)$  of a nontrivial graph  $S$  is the minimum number of edges whose removal from  $S$  results in a disconnected graph.

**Definition 6.** The *diameter*  $d(S)$  of a graph  $S$  is the greatest distance between any pair of nodes.

**Theorem 1.** The minimum nodal degree  $\delta(S)$  of a graph  $S$  is an upper bound on both the node and the edge-connectivity (Whitney's inequality):  $\kappa(S) \leq \lambda(S) \leq \delta(S)$ .

Mathematically, if  $S$  is a graph of the pipeline network, then the connectivity  $\kappa(S)$  (or edge-connectivity  $\lambda(S)$ ) of  $S$  becomes the smallest number of nodes (or communication links) whose collapse would reduce the safety in the network. Therefore, to remain connected the water network can simultaneously tolerate  $\kappa(S) - 1$  component failures or  $\lambda(S) - 1$  link failures which means that the higher the connectivity and edge-connectivity,



the more reliable the network. The water distribution network can be considered as functional as long as there is a non-faulty communication path between each pair of non-faulty components. In other words, the topological structure of the network remains connected in the occurrence of certain failures.

In vulnerability analysis, robustness is the ability of components and system to tolerate a given level of stress or demand without loss of function [4]. Hence, the connectivity and edge-connectivity are significant metrics for evaluation of robustness because they quantify the extent to which a graph can accommodate to node and link failures. The graph diameter  $d(S)$  is another interesting measure of robustness and efficiency. A graph's diameter is the largest number of nodes which must be traversed in order to travel from one node to another when paths which backtrack, detour, or loop are excluded from consideration. Small value of graph diameter  $d(S)$  describes this property that the water network nodes are mutually reachable and the network is ordered in a manner which helps with more efficient and balanced distribution of flow and smaller consequences as a result of failure of the most central components [29]. Here, some basic results on diameter, connectivity, and edge-connectivity of a graph are presented (the proofs can be found in [26]):

**Definition 7.** Suppose that  $S = (N, M)$  is a  $\lambda$ -edge-connected graph where  $\lambda(S) \geq \lambda \geq 1$ , and  $F$  is a set of edges in  $S$  with  $F \subset M$  and  $|F| < \lambda$ . The  $(\lambda - 1)$ -edge fault-tolerant diameter of  $S$ , denoted by  $D_\lambda(S)$ , is defined as:

$$D_\lambda(S) = d(S - F) : F \subset M, |F| = \lambda - 1 \tag{9}$$

where  $D_\lambda(S)$  is the diameter of a edge-deletion subgraph  $S - F$  which is obtained by removing  $\lambda - 1$  links from  $S$ .

**Theorem 2.** For any  $\lambda$ -edge-connected graph  $S$  it is clear that:

$$d(S) = D_1(S) \leq D_2(S) \leq \dots \leq D_{\lambda-1}(S) \leq D_\lambda(S) \tag{10}$$

**Definition 8.** Let  $S = (N, M)$  be a graph is a  $\kappa$ -connected graph where  $\kappa(S) \geq \kappa \geq 1$ , and  $U$  is a set of nodes in  $S$  with  $U \subset N$  and  $|U| < \kappa$ . The  $(\kappa - 1)$ -fault-tolerant diameter of  $S$ , denoted by  $D_\kappa(S)$ , is defined as:

$$D_\kappa(S) = d(S - U) : U \subset N, |U| = \kappa - 1 \tag{11}$$

where  $D_\kappa(S)$  is the diameter of a node-deletion subgraph  $S - U$  which is obtained by removing  $\kappa - 1$  nodes from  $S$ .

**Theorem 3.** For any  $\kappa$ -connected graph  $S$  it is clear that:

$$d(S) = D_1(S) \leq D_2(S) \leq \dots \leq D_{\kappa-1}(S) \leq D_\kappa(S) \tag{12}$$

### 3.1 The node and link-deletion problems

The node and link-deletion problems can be stated as the following graph-theoretic form: "Given a graph  $S$  and a graph property  $\Pi$ , the problem is to find a subset of edges or nodes,

the deletion of which results in a subgraph satisfying the property  $\Pi$ ." In 1981 Yannakakis proved that for a general class of properties this problem is NP-complete [28].

After defining connectivity, edge-connectivity, and diameter of a graph as the statistical metrics for fault tolerance assessment of networks, which indicate the robustness and efficiency, it is very useful to consider these metrics in the property  $\Pi$ . From **Theory 2** and **3** it can be concluded that when failures occur in a water network, its diameter  $d(S)$  may increase. Hence, a pipeline network with a small value of diameter is the most invulnerable network. On the other hand, when the connectivity (or edge-connectivity) of a water network is high, the number of nodes (or links) that have to be eliminated in order to disconnect a water network is a large number which means that the network is reliable under failure conditions. Thus, the above-mentioned node and link-deletion problems can be expressed in the following forms:

**Node-deletion problem:** Suppose that  $S$  is a  $\kappa$ -connected graph. Then the diameter of the graph  $G$  obtained from  $S$  by removing  $\kappa - 1$  of its nodes is well defined. The problem is finding the maximum diameter of  $G$ , denoted by  $D_\kappa$ .

**Link-deletion problem:** Suppose that  $S$  is a  $\lambda$ -edge-connected graph. Then the diameter of the graph  $G$  obtained from  $S$  by removing  $\lambda - 1$  of its links is well defined. The problem is finding the maximum diameter of  $G$ , denoted by  $D_\lambda$ .

The proposed re-formulation of node and link-deletion problems is appealing for its practical implementation because there are several efficient frameworks to evaluate the diameter and connectivity of a graph. In order to cope with the proposed problems and finding the upper bounds for  $D_\kappa(S)$  and  $D_\lambda(S)$ , practically the vulnerability assessment of water distribution networks based on graph theory can be performed using some open-source graph softwares such as igraph [30], MATLAB BGL [31], MATGRAPH [32], etc which are available free online. In this case, determining the network measurements take place in few simple algorithmic steps by utilizing these frameworks.

In order to assess the vulnerability of water networks and identifying their connectivity strength, it is essential to calculate  $\lambda(S)$ ,  $\kappa(S)$ , and  $d(S)$  of a graph. Computing these fundamental parameters has been extensively studied in the literature. Here, two straightforward procedures are applied, as described in Figure 1 and 2:

<b>Algorithm 1.</b>
Input: A connected non-trivial graph $S = (N, M)$ .
Output: Value of $d(S)$ .
1. Select an arbitrary pair of nodes as $(v, w) \in N$ .
2. Using Dijkstra's algorithm find the shortest path between $v$ and $w$ , denoted by $d(v, w)$ .
3. For each pair of nodes in $N$ , compute $d(v, w)$ .
4. Assign $d(S) \leftarrow \max \{ d(v, w) \mid (v, w) \in N \}$ .

Figure 1. The pseudo-code for computing  $d(S)$

In Algorithm 1 an efficient graph theory algorithm, called Dijkstra's algorithm, is employed for finding shortest path between nodes in a graph. Generally, the computational complexity for the Dijkstra's algorithm implementation on a graph  $S$  with  $n$  nodes and  $m$  edges, is  $O(n^2 + m)$ . However, the graph representing a water network is sparse, thus this algorithm can be implemented more efficiently by storing the graph in the form of linked

lists. In this case, the computational complexity time is  $O((n + m).log[n])$  [33]. The pseudo-code of Dijkstra’s algorithm is provided in Appendix A.

Many algorithms for the computation of  $\lambda(S)$  and  $\kappa(S)$  of graphs have been developed over the past four decades. Most of these algorithms are based on Menger's theorem, which determines the  $\lambda(S)$  and  $\kappa(S)$  of a graph in terms of the number of independent paths between nodes [34]. We define local edge-connectivity  $\lambda(v,w)$  as the minimum number of edges whose deletion from  $S$  would destroy every path between  $v$  and  $w$ . Similarly, local connectivity  $\kappa(v,w)$  can be defined as the minimum number of nodes, whose deletion from  $S$  would destroy every path between  $v$  and  $w$ . A collection of paths between  $v$  and  $w$  is called *independent* if no two of them share a node (other than  $v$  and  $w$  themselves). Also, the collection is edge-independent if no two paths in it share an edge. The number of mutually independent paths between  $v$  and  $w$  is denoted by  $\kappa'(v,w)$ , and the number of mutually edge-independent paths between them is written as  $\lambda'(v,w)$ . One of the most important facts about connectivity is Menger's theorem. This theory states that: (i)  $\kappa(v,w) = \kappa'(v,w)$  and (ii)  $\lambda(v,w) = \lambda'(v,w)$ , for every pair of nodes  $v$  and  $w$ . This fact is actually a special case of the max-flow theorem. By this theorem, for any two nodes in a connected graph, the values of  $\kappa(v,w)$  and  $\lambda(v,w)$  can be obtained efficiently using the max-flow algorithm. The  $\kappa(S)$  and  $\lambda(S)$  can then be determined as the minimum values of  $\kappa(v,w)$  and  $\lambda(v,w)$ , respectively. In optimization theory, the max-flow theorem asserts that in a flow network, the maximum amount of flow passing from the source to the sink is equal to the minimum capacity that, when removed in a particular way from the network, causes the situation that no flow can transport from the source to the sink. This theory is a special case of the duality theorem for linear programs and can be used to derive Menger's theorem and the König–Egerváry theorem [28]. The Menger's theorem led to the following algorithm (Figure 2) for computing  $\lambda(S)$  (an algorithm for determination of  $\kappa(S)$  is presented in Appendix B):

<b>Algorithm 2.</b>
Input: A connected non-trivial graph $S = (N,M)$ .
Output: Value of $\lambda(S)$ .
1. If $S$ is an undirected graph, replace each edge $xy \in M$ with arcs $(x, y)$ and $(y, x)$ .
2. Assign $v$ as the source vertex and $w$ as the sink vertex.
3. Assign the capacity of each arc to 1, and call the resulting network $S'$ .
4. Find a max-flow function $f$ in $S'$ .
5. Set $\lambda(v,w)$ equal to the total flow of $f$ .
6. Select an arbitrary vertex $u \in N$ , and let $X = N - \{u\}$ .
7. Compute $\lambda(u,t)$ for every $t \in X$ .
8. Assign $\lambda(S) \leftarrow \min\{ \lambda(u,t) \mid t \in X \}$ .

Figure 2. The pseudo-code for computing  $\lambda(S)$

The Algorithm 2 works by making a number of calls to a max-flow subroutine which identify the computational overheads of the algorithm. In a graph  $S$  with size  $n$ , the above

algorithm reduces the number of calls from  $\frac{n(n-1)}{2}$  to  $n-1$  by performing step (1) and (6),

where an undirected graph is changed to a directed graph, which makes the algorithm computationally efficient. For this algorithm the computational complexity time is  $O([nm].\min\{n^{2/3}, m^{1/2}\})$ .

#### 4. SIMULATION RESULTS

In this section, three benchmark water distribution networks are studied (Figure 3), each representing different organizational patterns. Here, the algorithms were implemented in MATLAB using the MATLAB BGL v1.0 [31] open-source package and simulation runs were carried out on a computer with an Intel Core i5 CPU, 2.53 GHz processor, and 3.00 GB RAM.

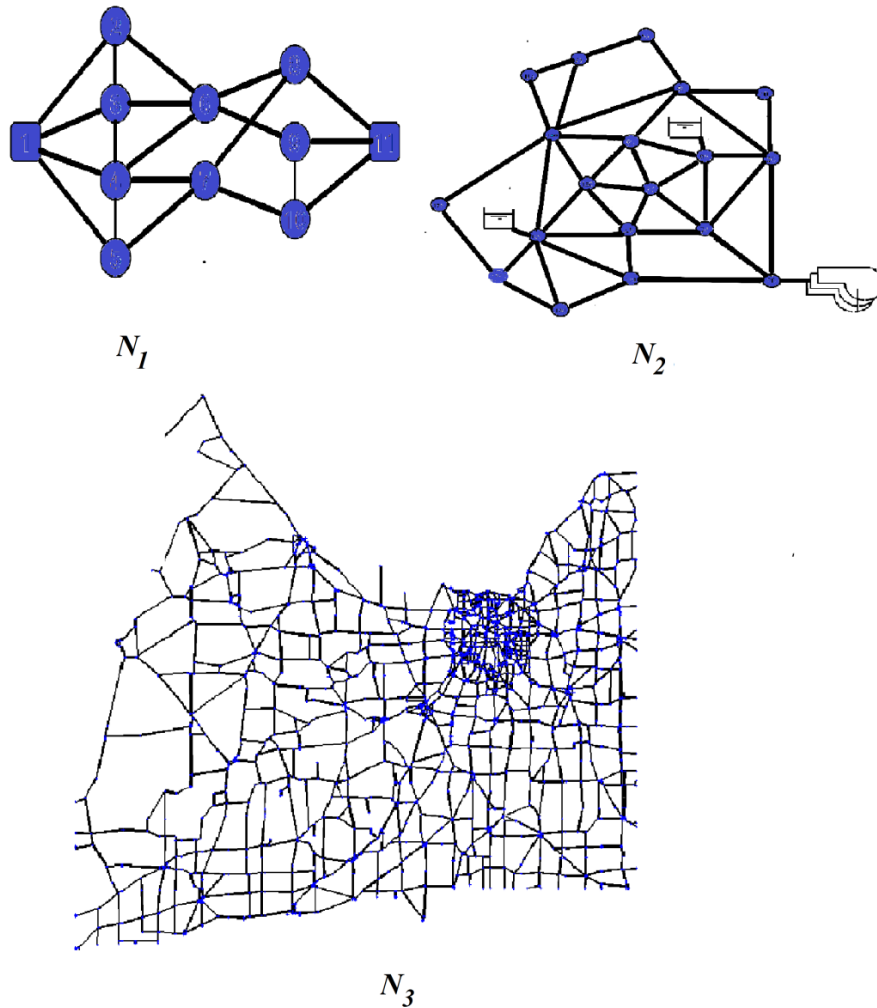


Figure 3. Layout of the studied water distribution networks

The MATLAB BGL is a MATLAB toolbox for working with graphs. It utilizes the Boost Graph Library which is a powerful set of algorithms on graph data structures to efficiently implement the graph algorithms. The numerical results and connectivity properties of each network are presented in Table 1.

Table 1: Network metrics for the studied water distribution networks

Network metrics	$N_1$ (small)	$N_2$ (medium)	$N_3$ (large)
Number of nodes ( $n$ )	11	25	2642
Number of links ( $m$ )	19	44	6606
Edge-connectivity ( $\lambda$ )	3	2	2
Connectivity ( $\kappa$ )	2	2	2
Graph diameter ( $d$ )	4	5	89

The  $N_1$  case study is a small hypothetical water network with 11 nodes and 19 links. Nodes with degrees 3 and 4 are the most frequent types of connection. As shown in Table 1, the small diameter ( $d = 4$ ) and the triangular loops are the main characteristics of the network. In addition, the minimum number of edges whose removal from  $N_1$  results in a disconnected graph is  $\lambda = 3$  (3-edge-connected graph). For this example the minimum number of nodes whose deletion generates a disconnected graph is  $\kappa = 2$ .

The second benchmark example is the Anytown water distribution network ( $N_2$ ). The  $N_2$  network is a medium size network with 38 pipes, 2 tanks, and 3 identical pumps delivering water from the treatment plant into the system. To meet the city expansion and increasing demands, up to 2 new tanks and 6 new pipes can be added to the system in the design problem. The particular all-channel structure of  $N_2$  causes high redundancy which improves network connectivity and makes the network failure-tolerant. Nodes with degrees 2 and 5 are the most frequent types of connection. The  $N_2$  network has also a small diameter of  $d = 5$ . For this case study the edge-connectivity of  $\lambda = 2$  (2-edge-connected graph) is obtained. This indicates that the  $N_1$  network is more invulnerable than  $N_2$ . Furthermore, both  $N_2$  and  $N_1$  are 2-connected graphs.

The  $N_3$  case study is a large-scale complex water distribution network with 2642 nodes and 6606 links. This case study has a mesh-like structure with redundant rectangular loops more than triangular loops. There are some structural holes on the left side of the  $N_3$  configuration, when viewed more globally, which makes the network vulnerable because in case of failure events, large compartments of the network can be separated. For the  $N_3$  case study the node and edge-connectivity of  $\kappa = 2$  (2-connected) and  $\lambda = 2$  (2-edge-connected) are obtained by Algorithm 2, respectively. These results reveal that deletion of relatively small number of nodes and links would reduce the safety in the network. By implementing Algorithm 1, after 1374.081 seconds (about 23 minutes) the proposed algorithm finds the diameter of  $d = 89$  for this case study.

## 5. CONCLUDING REMARKS

This paper provides an overview and a conceptual study framework based on graph theory for vulnerability analysis of water distribution networks and establishes some necessary but insufficient conditions to assess fault tolerance of water networks. Moreover, some basic principles are presented in order to design an invulnerable water network. Here, water networks are considered as interconnected, undirected, and planar graphs, and simple but efficient algorithms are applied to compute the statistical measurements of network topology including: connectivity ( $\kappa$ ), edge-connectivity ( $\lambda$ ), and diameter ( $d$ ).

In this work, three water distribution networks with increased scales and complexity are studied as benchmark networks and their reliability, in terms of efficiency and robustness, are evaluated. The numerical results revealed that node and edge-connectivity are important for the robustness because they quantify the extent to which a graph can accommodate to node and link failures. The larger the connectivity is, the more difficult it is to cut a graph into independent components. In the complex network models, the minimum nodal degree is a strict upper bound on both the node and the edge-connectivity. Therefore, the minimum nodal degree is a valuable estimate of the minimum number of nodes or links whose collapse results into a disconnected graph. Furthermore, the graph diameter is a measure of maximum graph eccentricity as well as a quantitative description for efficiency and robustness. A smaller graph diameter indicates a more efficient and robust network.

Nowadays, in urban communities due to the increasing dependency on water supply systems, vulnerability assessment of these infrastructural systems, in an efficient model-based method, becomes more crucial. Such models help to design a fault-tolerant water network with more safety. As illustrated in this study, graph theory methods are effective and computationally efficient approaches for modeling water networks appropriately. Further research is needed in areas like studying the algebraic connectivity in complex water distribution networks and realizing how various failures of the water network will affect the algebraic connectivity.

## APPENDIX A: DIJKSTRA'S ALGORITHM

For each node  $v \in N$ , Dijkstra's algorithm maintains an index  $\delta[v]$ , also known as a shortest-path estimate which is an upper bound on the weight of a shortest path from the source to  $v$ . Initially, the shortest-path estimates of all nodes other than the source are set to be  $\infty$ . Dijkstra's algorithm also maintains a set  $U$  of nodes whose final shortest-path weights from the source have not yet been determined. In a relaxation step, the algorithm repeatedly selects the node  $u \in U$  with the minimum shortest-path estimate and re-computes the  $\delta[v]$  of the nodes adjacent to  $u$ . Once a node is removed from  $U$ , its shortest-path weight from the source is determined and finalized, Figure 4. If  $S$  is an undirected graph, one can replace each undirected edge by 2 directed edges [33].

**Dijkstra's Algorithm**

Input: A weighted directed graph  $G = (N, M, w)$ ; a source node  $s$

Output: A shortest-paths spanning tree  $T$  rooted at  $s$

```

for each node  $v \in N$  do
   $\delta[v] \leftarrow \infty$  (infinity)
   $\pi[v] \leftarrow$  undefined
 $\delta[s] \leftarrow 0$ 
 $T \leftarrow \emptyset$ 
 $U \leftarrow N$ 
while  $U \neq \emptyset$  do
  Choose  $u \in U$  with minimum  $\delta[u]$ 
   $U \leftarrow U - \{u\}$ 
  if  $u \neq s$  then  $T \leftarrow T \cup \{(\pi[u], u)\}$ 
  for each node  $v$  adjacent to  $u$  do
    if  $\delta[v] > \delta[u] + w(u, v)$  then
       $\delta[v] \leftarrow \delta[u] + w(u, v)$ 
       $\pi[v] \leftarrow u$ 

```

Figure 4. The Dijkstra's algorithm

**APPENDIX B: COMPUTING NODE CONNECTIVITY**

As previously described in section 3.1,  $\kappa(v, w)$  for a pair of non-adjacent nodes  $v$  and  $w$  can be computed by solving a max-flow problem in a particular network. Therefore, the value of  $\kappa(S)$  can be determined as described in Algorithm 3 of Figure 5 [34]:

**Algorithm 3.**

Input: A connected non-trivial graph  $S = (N, M)$ .

Output: Value of  $\kappa(S)$ .

1. Replace each edge  $xy \in M$  with arcs  $(x, y)$  and  $(y, x)$ , and call the resulting digraph  $G$
2. Select a pair of non-adjacent vertices  $v$  and  $w$ .
3. For each node  $u$  other than  $v$  and  $w$  in  $S$ , replace  $u$  with two new nodes  $u_1$  and  $u_2$ , and then add the new arc  $(u_1, u_2)$ . Connect all the arcs that were coming to  $u$  in  $S$  to  $u_1$ , and similarly, connect all the arcs that were going out of  $u$  in  $S$  to  $u_2$  in  $G$ .
4. Assign  $v$  as the source node and  $w$  as the sink node.
5. Assign the capacity of each arc to 1, and call the resulting network  $H$ .
6. Find a max-flow function  $f$  in  $H$ .
7. Set  $\kappa(v, w)$  equal to the total flow of  $f$ .
8. Select an arbitrary node  $t$  of minimum degree from  $S$ .
9. Assign  $V_t \leftarrow$  the set of all vertices adjacent to  $t$
10. Compute  $k_1 = \min \{ \kappa(t, s) \mid s \in N - \{t\}, \text{ and } s \text{ is not adjacent to } t \}$ .
11. Compute  $k_2 = \min \{ \kappa(x, y) \mid x, y \in V_t; x \text{ and } y \text{ are non-adjacent} \}$ .
12. Assign  $\kappa(S) \leftarrow \min \{ k_1, k_2 \}$ .

Figure 5. The pseudo-code for computing  $\kappa(S)$

## REFERENCES

1. Giustolisi O, Kapelan Z, Savic D. Algorithm for automatic detection of topological changes in water distribution networks, *J Hydraul Eng* 2008; **134**(4): 435-46.
2. Xu C, Goulter, IC. Probabilistic model for water distribution reliability, *J Water Resour Plann Manage* 1998; **124**(4): 218-28.
3. Gargano R, Pianese D. Reliability as tool for hydraulic network planning, *J Hydraul Eng* 2000; **126**(5): 354-64.
4. Yazdani A, Jeffrey P. Robustness and vulnerability analysis of water distribution networks using graph theoretic and complex network principles, *Proceedings of the Water Distribution Systems Analysis, WDSA2010*, Tucson, AZ, USA, 2010, pp. 933-945.
5. Ostfeld A. Reliability analysis of water distribution systems, *Journal of Hydro Informatics* 2004; **6**(4): 281-94.
6. Shuang Q, Zhang M, Yuan Y. Performance and reliability analysis of water distribution systems under cascading failures and the identification of crucial pipes, *PLOS ONE* 2014; **9**(2): e88445.
7. Jacobs P, Goulter IC. Evaluation of methods for decomposition of water distribution networks for reliability analysis, *Civ Eng Syst* 1988; **5**(2): 58-66.
8. Jacobs P, Goulter IC. Optimization of redundancy in water distribution networks using graph theoretic principles, *Eng Optimiz* 1989; **15**(1): 71-82.
9. Goulter IC. Current and future use of systems analysis in water distribution network design. *Civ Eng Syst* 1987; **4**(4): 175-84.
10. Wagner JM, Shamir U, Marks, DH. Water distribution reliability: simulation methods, *J Water Resour Plann Manage* 1988; **114**(3): 276-94.
11. Kessler A, Ormsbee L, Shamir U. A methodology for least-cost design of invulnerable water distribution networks. *Civ Eng Syst* 1990; **7**(1): 20-28.
12. Perelman L, Ostfeld A. Topological clustering for water distribution systems analysis, *Environ Modell Softw* 2011; **26**: 969-72.
13. Ying K, Chu J, Qu J, Luo Y. A model and topological analysis procedures for a pipeline network of variable connectivity, *Adv Eng Softw* 2012; **48**: 40-51.
14. Gutiérrez-Pérez, JA, Herrera M, Pérez-García R, Ramos-Martínez E. Application of graph-spectral methods in the vulnerability assessment of water supply networks, *Math Comput Model* 2013; **57**(7): 1853-9.
15. Ball MO. Complexity of network reliability computation, *Networks* 1980; **10**: 153-65.
16. Bao Y, Mays LW. Model for water distribution system reliability, *J Hydraul Eng* 1990; **116**(9): 119-37.
17. Avritzer A, Carnevali L, Ghasemieh H, Happe L, Haverkort BR, Koziolk A, Menasche D, Remke A, Sedigh-Sarvestani S, Vicario E. Survivability evaluation of gas, water and electricity infrastructures, *Electr Notes Theor Comput Sci* 2105; **310**: 5-25.



18. Kaveh A. *Structural Mechanics: Graph and Matrix Methods*, 2nd edition, Research Studies Press Ltd, John Wiley & Sons Inc, 1995.
19. Kaveh A. An efficient program for generating subminimal cycle bases for the flexibility analysis of structures, *Commun Appl Numer Meth Eng* 1986; **2**(4): 339-44.
20. Kaveh A. *Computational Structural Analysis and Finite Element Methods*, Springer, Switzerland, 2014.
21. Rahal H. A co-tree flows formulation for steady state in water distribution networks, *Adv Eng Softw* 1995; **22**(3): 169-78.
22. Fujiwara O, Jenchaimahakoon B, Edirisinghe NCP. A modified linear programming gradient method for optimal design of looped water distribution networks, *Water Resour Res* 1987; **23**(6): 977-82.
23. Fujiwara O, Khang DB. A two-phase decomposition method for optimal design of looped water distribution networks, *Water Resour Res* 1990; **26**(4): 539-49.
24. Tahershamsi A, Kaveh A, Sheikholeslami R, Talatahari S. Big bang-big crunch algorithm for least-cost design of water distribution systems, *Int J Optim Civil Eng* 2012; **2**(1): 71-80.
25. Sheikholeslami R, Kaveh A, Tahershamsi A, Talatahari S. Application of charged system search algorithm to water distribution networks optimization, *Int J Optim Civil Eng* 2014; **4**(1): 41-58.
26. Xu JM. *Topological Structure and Analysis of Interconnection Networks*, Kluwer Academic Publishers, Chicago, 2001.
27. Kaveh A. *Optimal Analysis of Structures by Concepts of Symmetry and Regularity*, Springer, Wien-New York, 2013.
28. Gross JL, Yellen J, Zhang P. (eds.). *Handbook of Graph Theory*, 2nd edition. CRC Press, 2014.
29. Pandit A, Crittenden JC. Index of network resilience (INR) for urban water distribution systems, *Proceedings of Critical Infrastructure Symposium*, Arlington, VA, 2012.
30. Csardi G, Nepusz T. The igraph software package for complex network research, *Int J Complex Syst* 2006; **1695**(5): 1-9.
31. Gleich D. *Matlab BGL: A Matlab Graph Library*. Institute for Computational and Mathematical Engineering, Stanford University, 2008. URL: <http://dgleich.github.io/matlab-bgl/>.
32. Scheinerman ER. *Matgraph: A Matlab Toolbox for Graph Theory*. 2007. URL: <http://www.ams.jhu.edu/~ers/matgraph/>.
33. Federickson GN. Fast algorithms for shortest paths in planar graphs, with applications, *SIAM J Comput* 1987; **16**(6): 1004-22.
34. Esfahanian AH, Hakimi SL. On computing the connectivities of graphs and digraphs, *Networks* 1984; **14**: 355-66.