

# Probability of network disconnection of water distribution system for maintenance prioritization

Hieu Chi Phan, Ashutosh Sutra Dhar, Premkumar Thodi and Rehan Sadiq

## ABSTRACT

Unexpected pipe breaks in municipal water distribution systems may cause isolation of parts of the network or reduction of redundancy, leading to reduced system reliability. While a network with less redundancy implies less tolerance to further breaks, the isolation of nodes explicitly indicates unavailability of the system to service the nodes. This paper presents a method of measuring these topological changes using algebraic connectivity (AC). AC is a parameter that can be used to assess robustness and redundancy of a network. The changes in AC associated with pipe breaks are compared with the AC of intact networks to assess whether the removal of the pipe causes reduction of redundancy or isolation in the network. An AC of  $1.5625 \times 10^{-5}$  is calculated for an intake of a medium-sized water distribution network (WDN). The fluctuation in AC is used to assess the criticality of each pipe segment to the overall structure of the network. This study also evaluates the failure probability of the WDN, assuming that the network failure probability is equivalent to the probability of isolation of parts of the system due to pipe breaks. The breaks leading to the failure are identified using the method of the minimum cut-sets.

**Key words** | algebraic connectivity, complex network analysis, minimum cut-sets, network failure probability, network reliability, water distribution network

### Hieu Chi Phan

**Ashutosh Sutra Dhar** (corresponding author)  
Faculty of Engineering and Applied Science,  
Memorial University of Newfoundland,  
St John's, NL A1B 3X5,  
Canada  
E-mail: [asdhar@mun.ca](mailto:asdhar@mun.ca)

### Premkumar Thodi

C-CORE,  
St John's, NL,  
Canada

### Rehan Sadiq

The University of British Columbia,  
Okanagan Campus, Kelowna, BC,  
Canada

## INTRODUCTION

The municipal water distribution network (WDN) is the life-line of communities. Since potable water is the most essential element for life, a WDN is designed to provide uninterrupted service to community dwellers. For this purpose, the water distribution system (WDS) is designed with an amount of redundancy using looped networks to provide alternative paths for situations if one or more links go out of service. It is, however, difficult to assess all possible paths from the source to the demand points for a municipal WDN containing thousands of demand nodes and pipes. The performance of the WDS in providing an acceptable level of service to consumers is assessed in terms of system reliability (Tung 1985). Over the last few decades, considerable research has focused on system reliability assessments of the WDN. However, no universally accepted definition and measure of system reliability has been used.

Wagner *et al.* (1988) developed analytical methods for system reliability assessment using network connectivity and approachability. Wagner *et al.* (1988) and Quimpo & Shamsi (1991) incorporated multiple connections from several source nodes to demand nodes for estimating the system reliability. The connectivity and approachability are defined as the connection of all demand nodes to the source nodes, and the connection of a demand node to its source, respectively.

Su *et al.* (1987) and other researchers employed hydraulic failure as the reliability measure for WDS design optimization, which is based on the measure of hydraulic availability of the system needed to provide an acceptable level of service. This approach is rigorous and provides a precise description of a WDS. However, analysis of hydraulic availability involves extensive system simulations and is therefore computer time-intensive. To overcome the

computational problem, several heuristic measures of system reliability were developed (Shamir & Howard 1981; Goulter & Coals 1986; Awumah *et al.* 1991), which are criticized for not evaluating the hydraulic availability (Cullinane *et al.* 1992). Cullinane *et al.* (1992) attempted to balance the two types of measures; hydraulic failure is evaluated for selected conditions. This approach still relies on heuristics and judgement. Shinstine *et al.* (2002) demonstrated that availability measurement provides a means of examining the robustness of a WDS during design, rather than relying on implicit judgement. They employed a hydraulic simulation model (i.e., KYPIPE) to obtain the values of a pressure head at each demand node by closing a pipe or a combination of pipes, which was repeated until all the combinations of the pipes were examined. Pipes were considered as repairable components and an average time to repair of 1 day was assumed. In this study, a WDS containing up to 109 pipes and 89 nodes was analysed. However, repeated hydraulic simulation of a larger WDN containing thousands of nodes would be tedious. Aydin *et al.* (2014) reported 7 days of simulation of a WDN using the EPANET model. Moreover, this approach only simulates the network conditions when valves are closed immediately after pipe breaks. Water disruption due to leaks in water mains is not simulated.

Despite the limitations, the hydraulic-based approach is desired for the design optimization of WDS, since it provides water availability at the demand nodes. On the other hand, for repairing deteriorating water mains in an existing WDN, the pipe breaks causing isolation of part of the system should be given priority. If isolated from the network, the demand nodes will not have any available supply. The focus of the present study is the prioritization of water mains for repair/maintenance and to identify the pipes and combinations of pipes causing isolation (disconnection) of parts of the WDN. The failure probability of the system is defined as the probability of disconnection. Although breaks of other pipes may also cause water to be unavailable at certain demand nodes due to low pressure, this is not considered here, to avoid computationally expensive hydraulic simulations.

To identify a pipe and combinations of pipes causing isolation, the minimum cut-set method is used. In the conventional minimum cut-set method, pipes causing system failure are rigorously examined, requiring complex and extensive computation. Yannopoulos & Spiliotis (2013)

employed graph theory using a connectivity matrix to identify minimum cut-sets. The computational time required in this method is also significantly high, as discussed in more detail in the paper. To overcome this computational limitation, this paper employs complex network analysis (CNA) to find the minimum cut-sets of the network. CNA has been used for the quantification of structural properties of networks and to improve understanding of network connectivity and robustness. Complex network models have been applied in various areas including WDNs (Yazdani & Jeffrey 2010, 2011a, 2011b; Nazempour *et al.* 2016). Yazdani & Jeffrey (2010) discussed the metrics for CNA to evaluate the robustness and vulnerability characteristics of a WDN. Among several metrics for CNA, algebraic connectivity (AC) has been proposed and analysed by Fiedler (1973). AC is a well-known measure for evaluating the well-connectedness of a graph (Fiedler 1973; Capocci *et al.* 2005; Ghosh & Boyd 2006; Newman 2006; Yazdani & Jeffrey 2011a, 2011b). The network with higher AC is more robust and more tolerant to the breakage of links. AC provides information on graph partitioning that is closely related to the disconnection of nodes due to pipe breaks (Phan *et al.* 2017). Fiedler (1973) pointed out that an event of disconnection in a system would lead to an increase of AC. On the other hand, decrease of AC corresponds to a reduction of redundancy in the network (Phan *et al.* 2017). AC has been used to develop a framework to evaluate the redundancy or robustness of WDNs (Fiedler 1973; Yazdani & Jeffrey 2010; Phan *et al.* 2017). However, no such comprehensive modelling has ever been performed to take advantage of the characteristics of the AC for network reliability assessment.

The novelty of the current paper is the development of a method for finding minimum cut-sets of complex WDNs using AC. The proposed approach is computationally more efficient than the existing methods of finding minimum cut-sets. A reliability framework is then developed based on AC and the failure probability of pipe components. The proposed method is applied to a real WDN.

## RELIABILITY WITH MINIMUM CUT-SETS

This study assumes that network disconnection is equivalent to system failure and is denoted as an event,  $E$ . The

probability of system failure can be determined using the minimum cut-set approach (Tung 1985; Yannopoulos & Spiliotis 2013). The minimum cut-set is a combination of a minimum number of component failures that lead to system failure (i.e., disconnection, in the current study). If  $E_k$  is an event of system failure due to a set of  $k$  pipes,  $k$  ranges from 1 to  $m-1$  where  $m$  is the number of pipes in the network, the relationship between event  $E$  and events  $E_k$  can be written as:

$$E = E_1 \cup E_2 \cup \dots \cup E_{m-1} \tag{1}$$

Because the events  $E_k$  are mutually exclusive, then:

$$P(E) = \sum_{k=1}^m P(E_k) \tag{2}$$

where  $P(E)$  is the probability of system failure and  $P(E_k)$  is the probability of event  $E_k$  to occur.

$pk$  is defined as the number of minimum cut-sets consisting of  $k$  pipes in the network, and event  $E_k$  is the union of  $pk$  events corresponding to  $pk$  minimum cut-sets:

$$E_k = E_{k,1} \cup E_{k,2} \cup \dots \cup E_{k,pd} \cup \dots \cup E_{k,pk} \tag{3}$$

$$P(E_k) = 1 - \prod_{pd=1}^{pk} (1 - P(E_{k,pd})) \tag{4}$$

$E_{k,pd}$  is the event of the  $pd^{th}$  minimum cut-set to occur, where  $pd$  ranges from 1 to  $pk$ . The relationship of system failure and the event of minimum cut-sets occurring is provided in the fault tree analysis (FTA) diagram shown in Figure 1.

Here,  $E_{k,pd}$  is the intersection of  $k$  basic events  $E_{k,pd,i}$  with  $i = [1:k]$ . In other words,  $E_{k,pd,i}$  is the event of the  $i^{th}$  pipe break in the  $pd^{th}$  minimum cut-set, and thus:

$$E_{k,pd} = E_{k,pd,1} \cap E_{k,pd,2} \cap \dots \cap E_{k,pd,i} \dots \cap E_{k,pd,k} \tag{5}$$

$$P(E_{k,pd}) = \prod_{i=1}^k P(E_{k,pd,i}) \tag{6}$$

Equation (4) can be rewritten as:

$$P(E_k) = 1 - \prod_{pd=1}^{pk} \left( 1 - \prod_{i=1}^k P(E_{k,pd,i}) \right) \tag{7}$$

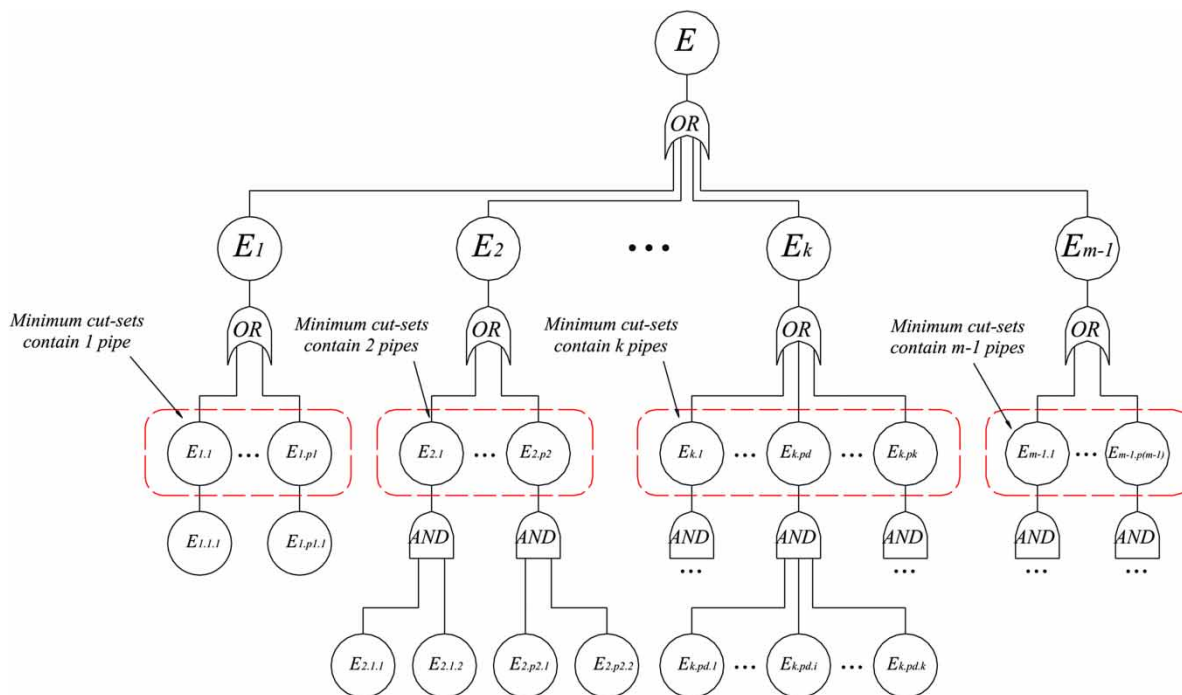


Figure 1 | FTA of the event of disconnection occurrence in the network.

Since  $E_{k,pd,i}$  is the event of an individual pipe breakage, the probability of these basic events can be found directly using the component reliability model.

From Equations (2) and (7), the probability of disconnection (i.e., system failure) in the network can be estimated as:

$$P(E) = \sum_{k=1}^m \left[ 1 - \prod_{pd=1}^{pk} \left( 1 - \prod_{i=1}^k P(E_{k,pd,i}) \right) \right] \quad (8)$$

The failure probability of a simple hypothetical network with its connection of nodes and lengths of pipes as presented in Figure 2 is examined using this approach. For the system failure probability assessment, reliability of the components (the pipes in the current study) in the network is required. Researchers have investigated several risk factors, including age, diameter, length, pipe material, corrosiveness of soil and operating pressure for the prediction of water main breaks and developed relationships with the time-to-failure (Goulter & Coals 1986; Kleiner & Rajani 2010). The major challenge with this approach is the lack of availability of data to develop the model. Municipalities often lack resources to collect data, except for break records. Historic break records are therefore used to predict subsequent breaks using a homogeneous or non-homogeneous Poisson model (Goulter & Kazemi 1988; Kleiner & Rajani 2010). The historic break record-based model is considered here for the hypothetical network as an example. Application of this framework for other failure models (e.g., use of a failure function and Monte Carlo simulation) is demonstrated later in the paper. The reliability model is developed using the average break per year ( $Br$ ) of the network, which is readily available to a city. The average break rate for each pipe segment is assumed to be proportional to the length of the segment. The event of a water main break is assumed to be a Poisson process, which is commonly used to forecast the break patterns in individual water mains (e.g., Goulter &

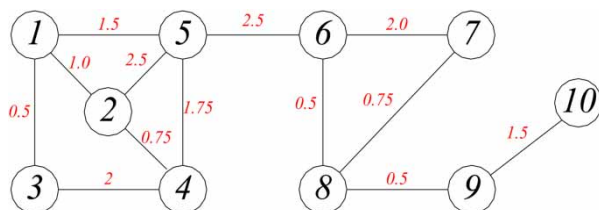


Figure 2 | Example of a hypothetical network.

Coals 1986; Kleiner & Rajani 2010; Yannopoulos & Spiliotis 2013). A homogeneous Poisson process with a constant annual failure rate is considered. Note that the annual break rate is not necessarily constant for a WDS. In this case, a non-homogeneous Poisson model or other models can be developed for pipe break assessments. The development of a component failure model is not within the scope of the current study.

Based on these assumptions, the average annual break rate for pipe linking node  $i$  and node  $j$  (from now on, written as pipe  $ij$ ) can be found as follows:

$$\lambda_{ij} = \frac{Br}{l_{total}} l_{ij} \quad (9)$$

where  $Br$  is the average annual break rate in the WDN,  $l_{total}$  is total length of pipes in the network,  $l_{ij}$  is length of pipe  $ij$  and  $\lambda_{ij}$  is break rate of pipe  $ij$ .

Since the pipe breaking is assumed to be a Poisson process, the annual pipe failure probability can be estimated using an exponential model with a constant failure rate. This comes from the assumption that the average annual break rate is a constant. Therefore, the component reliability of pipe  $ij$  at time  $t$ ,  $R_{ij}(t)$ , is the probability of a pipe to function at a desired level within the period of time  $[0:t]$ , where:

$$R_{ij}(t) = e^{-\lambda_{ij}t} \quad (10)$$

Thus,

$$P_{ij}(t) = 1 - e^{-\lambda_{ij}t} \quad (11)$$

For the hypothetical network in Figure 2, the average break per year ( $Br$ ) is randomly chosen as three breaks per year. From Equations (9) and (10), the failure probability of each pipe within 1 month ( $t = 1/12$ ) is calculated and given in Table 1. Consequently, the network disconnection probability is found as 0.063 with the details of calculating process given in Table 2.  $E_5$  (empty) and  $E_{13}$  (empty) in Table 2 indicate that there are no minimum cut-sets with five and 13 pipes, respectively.

In Table 1, the pipe is named for the first and second nodes that it links. For example, pipe 5–6 means the pipe connects nodes 5 and 6. The numbers alongside in Figure 2 indicate the assumed geometric length of the corresponding pipe. For example, pipe 5–6 has 2.5 units length.

**Table 1** | Failure probability of pipes in a period of 1 month

Pipe	Length	$\lambda_{ij}$	$P_{ij}(t)^a$
1-2	1.0	0.169014	0.014085
1-3	0.5	0.084507	0.007042
1-5	1.5	0.253521	0.021127
2-4	0.75	0.126761	0.010563
2-5	2.5	0.422535	0.035211
3-4	2.0	0.338028	0.028169
4-5	1.75	0.295775	0.024648
5-6	2.5	0.422535	0.035211
6-7	2.0	0.338028	0.028169
6-8	0.5	0.084507	0.007042
7-8	0.75	0.126761	0.010563
8-9	0.5	0.084507	0.007042
9-10	1.5	0.253521	0.021127

<sup>a</sup> $t = 1$  month.

It can be observed from Table 2 that  $P(E_1) \gg P(E_2) \gg P(E_3) \gg P(E_4)$ . In general,  $P(E_k) \gg P(E_h)$  (where  $h > k$ ). This is due to the fact that  $P(E_h)$  requires an intersectional condition for  $h$  (i.e., larger than  $k$ ) simultaneous pipe breaks. Similarly, the intersectional condition indicates

the fact that  $P(E_{k-1}) \ll P(E_k)$ , since the failure probability of each component is less than 1. The difference of  $P(E_k)$  and  $P(E_h)$  is more significant when components  $P(E_{k,pd,i})$  in Equation (7) are relatively small. The average failure probability,  $P_{Br,av}$ , is useful to quantify this difference. The average failure probability can be obtained by taking the ratio of expected pipe breaks per total number of pipes in the network. This ratio is positively correlated with the breakage probability of the pipe components. Assuming that the pipe break is a Poisson process, the average failure probability within a period  $\Delta t$  can be written as:

$$P_{Br,av}(\Delta t) = \frac{Br \times t}{m} \tag{12}$$

where  $Br$  is the average annual break rate in the WDN and  $m$  is the number of pipes in the network.

If the first component in Equation (2) (i.e.,  $P(E_1)$ ) significantly overwhelms the summation of the others, then network disconnection probability  $P(E)$  can be approximated by  $P(E_1)$ . That is:

$$P(E) \approx P(E_1) \tag{13}$$

**Table 2** | Calculations of disconnection probability of network in a period of 1 month

Min. cut-sets contain $k$ pipes $E_k$	$pk$	Min. cut-set ( $E_{k,pd}$ )	Pipes in the min. cut-set $E_{k,pd}$ ( $E_{k,pd,1pd,1} ; \dots ; E_{k,pd,1pd,k}$ )	$P(E_{k,pd})$	$P(E_k)$
$E_1$	3	$E_{1,1}$	5-6	0.035211	0.0622
		$E_{1,2}$	8-9	0.007042	
		$E_{1,3}$	9-10	0.021127	
$E_2$	4	$E_{2,1}$	1-3; 3-4	0.000198	$7.6849 \times 10^{-4}$
		$E_{2,2}$	6-7; 6-8	0.000198	
		$E_{2,3}$	6-7; 7-8	0.000297	
		$E_{2,4}$	6-8; 7-8	$7.439 \times 10^{-5}$	
$E_3$	6	$E_{3,1}$	1-2; 1-3; 1-5	$2.0955 \times 10^{-6}$	$4.3219 \times 10^{-5}$
		$E_{3,2}$	1-2; 1-5; 3-4	$8.3820 \times 10^{-6}$	
		$E_{3,3}$	1-2; 2-4; 2-5	$5.2387 \times 10^{-6}$	
		$E_{3,4}$	1-3; 2-4; 4-5	$1.8336 \times 10^{-6}$	
		$E_{3,5}$	1-5; 2-5; 4-5	$1.8336 \times 10^{-6}$	
		$E_{3,6}$	2-4; 3-4; 4-5	$7.3342 \times 10^{-6}$	
$E_4$	5	$E_{4,1}$	1-2; 1-3; 2-5; 4-5	$8.6082 \times 10^{-8}$	$7.8458 \times 10^{-7}$
		$E_{4,2}$	1-2; 1-5; 2-4; 4-5	$7.7474 \times 10^{-8}$	
		$E_{4,3}$	1-2; 2-5; 3-4; 4-5	$3.4433 \times 10^{-7}$	
		$E_{4,4}$	1-3; 1-5; 2-4; 2-5	$5.5339 \times 10^{-8}$	
		$E_{4,5}$	1-5; 2-4; 2-5; 3-4	$2.2136 \times 10^{-7}$	
$E_5$ (empty)	0	-	-	-	0
$E_{13}$ (empty)	0	-	-	-	0
			Sum:		0.0630

To investigate the relationship of  $P_{Br.av}$  and the difference between  $P(E_1)$  and  $P(E)$ , the network in Figure 2 has been investigated by changing the average break ( $Br$ ) within a fixed period of time,  $\Delta t$ . Figure 3 compares the results. Figure 3(a) shows that  $P(E)$  is very close to  $P(E_1)$ . This is supported by Figure 3(b), in which the ratio between  $P(E_1)$  and  $P(E)$  is consistently larger than 0.89. This demonstrates that  $P(E_1)$  accounts for more than 89% of  $P(E)$  in all cases. It also implies that the approximation in Equation (13) can be more accurate with small  $P_{Br.av}$ . In other words, the smaller the  $P_{Br.av}$ , the closer the distance between  $P(E_1)$  and  $P(E)$  becomes. It is observed from Figure 3 that  $P(E_1)$  can be used to predict  $P(E)$  within an acceptable small value of  $P_{Br.av}$ . In the case of a WDN, containing thousands of pipes combining with relatively small  $Br$ , the system  $P_{Br.av}$  is expected to be small. This leads to the approximation presented in Equation (13) being reasonable.

## ALGEBRAIC CONNECTIVITY

One of the major limitations of using the minimum cut-set approach for evaluating the disconnection probability of a large network is the high computational cost needed to determine all minimum cut-sets. Applying the approximation presented in Equation (13), the computational cost can be significantly reduced. The conventional approach to solve the minimum cut-set problem is to count all the paths connecting two different nodes of the network. Yannopoulos & Spiliotis (2013) used the connectivity

matrix from graph theory to record all the paths. If any element of the connectivity matrix is zero, then there is a disconnection between the concerned nodes. Subsequently, the minimum cut-sets are found by removing (idealized for breaking) a set of pipes and recalculating the resulting connectivity matrix. If the connectivity matrix of the break-containing network comprises zero elements, then disconnection occurs. However, computational cost of such an approach may be significantly increased when the size of the network is large. In the following sections, the characteristic of the AC is investigated to determine the minimum cut-sets for a single pipe break and multiple pipe breaks.

## AC for single pipe breaks

AC is a parameter used in graph theory to determine the strength of connection between the nodes in a network. Mathematically, it is defined as the second smallest Eigenvalue of the Laplacian matrix of the connected graph (Fiedler 1973). Municipal WDNs can be considered as graphs of connected networks, where a number of pipes connect the nodes at their intersections. Then, the graph of WDN can be described as  $G = G(V, P)$ , where  $V$  is a set of  $n$  nodes (intersections) and  $P$  is a set of  $m$  pipes. An adjacent matrix  $A$  of  $G$  is used to describe the link between the nodes, where:

$$A = a_{ij} \quad (14)$$

$a_{ij} = 1$ , if there is a link (pipe) between node  $i$  and node  $j$ .  
 $a_{ij} = 0$ , if there is no link (pipe) between node  $i$  and node  $j$ .

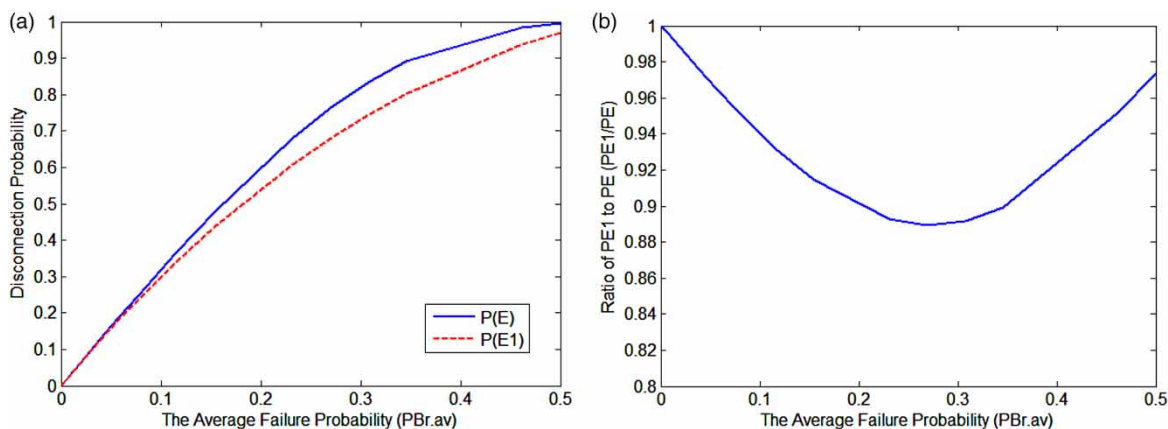


Figure 3 | The difference between  $P(E)$  and  $P(E_1)$  with the change of  $P_{Br.av}$ . (a) Comparison of  $P(E)$  and  $P(E_1)$ . (b) Ratio of  $P(E_1)/P(E)$ .

The node-degree matrix is a diagonal matrix, which contains the information about the number of connections (node-degree) at each node, and is defined as:

$$D = \text{diag} (d_i)$$

$d_i$  = number of connection (node-degree) of node  $i$ , where:

$$d_i = \sum_{j=1}^n a_{ij}$$

Then, the Laplacian matrix is given by Equation (15):

$$L = D - A \quad (15)$$

indicates that if a break does not isolate a node or cluster of nodes, the  $AC$  will decrease from the value prior to the break.

However, when a node or a cluster of nodes is isolated due to pipe breakage, the Laplacian matrix of the network becomes a combination of two (or more) non-connected sub-networks. Then, the  $AC$  will be higher. For example, assume a network has two clusters: Cluster 1 includes nodes from 1 to  $i-1$  and Cluster 2 includes nodes from  $i+1$  to  $n$ . The clusters of nodes are only linked by  $i^{\text{th}}$  node. If links to this node are all broken, then the network Laplacian matrix will have the following form:

$$L = \begin{bmatrix} [\text{Laplacian Matrix of Cluster 1}] & [0] & [0] & \\ & [0] & 0 & [0] \\ & [0] & [0] & [\text{Laplacian Matrix of Cluster 2}] \end{bmatrix} \begin{matrix} \text{Node 1 to } i-1 \\ \text{Node } i \\ \text{Node } i+1 \text{ to } n \end{matrix} \quad (16)$$

The Laplacian matrix  $L$  for the undirected network is usually symmetrical and the sum of rows (and columns) is zero. This characteristic leads to the fact that the first (i.e., smallest) Eigenvalue ( $\lambda_1$ ) of the matrix is zero, corresponding to the Eigen vector of  $(1,1, \dots, 1)^T$ . The second smallest Eigenvalue ( $\lambda_2$ ) of the Laplacian matrix is the  $AC$ , which is greater than 0 if  $G$  is a connected graph. The Eigenvalues of a network with  $n$  nodes of connected graph are:  $\lambda_1 = 0 \leq \lambda_2 = AC \leq \lambda_3 \leq \dots \leq \lambda_n$ .

A network with higher  $AC$  implies that the network is better connected (Newman 2010; Yazdani & Jeffrey 2010, 2011a). The addition of links to a network has been observed to increase the  $AC$ , as reported in Ghosh & Boyd (2006) and Phan et al. (2017). Ghosh & Boyd (2006) have reported a method to maximize the  $AC$  with a set of pipes to increase the redundancy of the network when the number of nodes remains the same. On the contrary, the process of pipe breakage may be considered as a reverse process of decreasing the  $AC$  with the break or removal of pipes from the network. As long as the disconnection has not occurred, the  $AC$  tends to decrease, implying that the redundancy of the network is reduced after a break event. This

The separation above is called graph partitioning. In the case of graph partitioning, the  $AC$  of the network is a combination of the isolated clusters. Fiedler (1973) mathematically proved that:

$$\lambda_2(G) \leq \lambda_2(G_m) \leq \lambda_2(G_{rp}) \quad (17)$$

where  $\lambda_2(G)$  is the  $AC$  of the connected network,  $\lambda_2(G_m)$  is the  $AC$  of the network after the disconnection with removal of nodes, and  $\lambda_2(G_{rp})$  is the  $AC$  of the network after the disconnection with removal of pipes.

The disconnection may be viewed as the direct failure of the network, considering the system failure as the state of not properly serving water to every demand node (i.e., the communities). The reduction of redundancy might lead to the disabling of service to demand nodes, but might not directly result in a network failure event. Thus, the  $AC$  will change with the removal of a pipe in the network, where:

- $AC$  will increase if an isolation of nodes occurs, and
- $AC$  will decrease if the redundancy is reduced.

Now, if the removal of a pipe leads to disconnection of the network (e.g.,  $AC$  increases), such a pipe is considered as

a minimum cut-set. Thus, by observing the change in  $AC$  by removing pipes one by one within the network, all minimum cut-sets can be determined.

For illustration, the  $AC$  for a hypothetical network shown in Figure 2 is estimated to demonstrate the numerical values of the parameter in describing robustness of the network. The network consists of 10 nodes and 13 links (i.e., pipes) with the  $AC$  of 0.19781 (i.e.,  $AC_{intact} = 0.19781$ ). It is representative of a WDN connecting two distinct communities. Community 1 contains nodes 1 to 5 and Community 2 contains nodes 6 to 10. The communities are connected by pipe 5–6. In other words, pipe 5–6 has a high betweenness.

Table 3 ranks 13 events corresponding to the removals of 13 pipes from the network. The difference between  $AC$  before and after removal of a pipe is presented in a descending order. There are three events with  $AC$  increases, resulting in the network disconnection. These are events 1.1, 1.2 and 1.3, with the removals of pipes 5–6, 8–9 and 9–10, respectively. These pipes are the three minimum cut-sets that contain only one break corresponding to events  $E_{1.1}$ ,  $E_{1.2}$  and  $E_{1.3}$ , shown in Table 2.

As seen in Figure 2, the removal of pipes 5–6, 8–9 and 9–10 will result in the disconnection of a node or cluster of nodes. Removal of pipe 5–6 caused the highest increase (i.e., ranked first) since it has high betweenness. Pipe 9–10 has the least increase because the removal of 9–10 leads to

the isolation of a single node (node 10). Phan *et al.* (2017) have concluded from a previous study that the size and number of disconnected clusters greatly affect the increase of the  $AC$ .

The other ten events (i.e., event 1.4 to 1.13) in Table 3 show a decrease of the  $AC$  with removal of a single pipe, indicating no disconnection. However, the redundancy of the network is decreased, and thus the change in the  $AC$  shows a positive correlation to redundancy. Similar observations were reported in previous studies (Yazdani & Jeffrey 2011a, 2011b). These ten events can further be divided into two groups: the first group consists of events 1.4 to 1.10 with the removal of pipes in Community 1; the second group is the events 1.11 to 1.13 with the removal of pipes in Community 2, which has a higher absolute change in  $AC$  ( $\Delta AC$ ) compared to the first group. This difference is due to the fact that Community 2 (with  $AC = 0.5188$ ) has less redundancy compared to Community 1 (with  $AC = 2.000$ ).

From the analysis it can be observed that the  $\Delta AC$  can be used for the following:

1. Identifying the disconnection and reduction of redundancy by observing the sign of  $\Delta AC$ , where a positive sign (+) denotes network disconnection and a negative sign (–) indicates reduction of redundancy.
2. Identifying the components causing most significant disconnection based on the highest positive  $\Delta AC$ .

Table 3 |  $AC$  with removal of a link

Sub-event	Removed pipe ID	$AC$	Change in $AC$ $\Delta AC = AC - AC_{intact}$	Type of network struct. change	Grouped as	Note
1.1	5–6	0.518806	0.32100	Disconnection	Group 1	$E_{1.1}$
1.2	8–9	0.351227	0.15342	Disconnection	Group 1	$E_{1.2}$
1.3	9–10	0.266187	0.06838	Disconnection	Group 1	$E_{1.3}$
1.4	1–2	0.197626	–0.00018	Less redundancy	Group 2	
1.5	2–4	0.197626	–0.00018	Less redundancy	Group 2	
1.6	1–3	0.193938	–0.00387	Less redundancy	Group 2	
1.7	3–4	0.193938	–0.00387	Less redundancy	Group 2	
1.8	2–5	0.186393	–0.01141	Less redundancy	Group 2	
1.9	1–5	0.182328	–0.01548	Less redundancy	Group 2	
1.10	4–5	0.182328	–0.01548	Less redundancy	Group 2	
1.11	7–8	0.181990	–0.01582	Less redundancy	Group 2	
1.12	6–7	0.167151	–0.03066	Less redundancy	Group 2	
1.13	6–8	0.134125	–0.06368	Less redundancy	Group 2	



3. Identifying the components causing the most significant decrease in redundancy, based on the highest negative  $\Delta AC$ .

The  $AC$ -based method to determine the minimum cut-set involving a single pipe break, as discussed above, is based on solving the Eigenvalue problem. This is basically different from the method of counting paths as described in Yannopoulos & Spiliotis (2013). To identify whether a single pipe or a set of pipes breaking is a minimum cut-set or not, the method reported in Yannopoulos & Spiliotis (2013) requires finding the associated connectivity matrix  $B$ , as presented below:

$$B = AA^1 + AA^2 + \dots + AA^{n-1} \quad (18)$$

where  $AA$  is the adjacent matrix  $A$  (see Equation (14)) with the removal of the concerned pipe or set of pipes.

Since the computational cost of multiplying two matrices with size  $n \times n$  is  $O(n^3)$  for a naive algorithm, the computational cost of the last component of matrix  $B$  will be  $O(n^4)$ . Even for the improved algorithm proposed by Coppersmith & Winograd (1990), the computational cost of estimating  $A^{n-1}$  will be approximately  $(n-2) \times O(n^{2.375477}) = O(n^{3.375477})$ . However, using the Eigenvalue problem, it can be reduced to  $O(n^2)$  for each step of the interaction, for calculating only Eigenvalues (Flannery et al. 1992), compared to the cost of  $O(n^3)$  otherwise. Thus, calculating the last component in Equation (18) itself is more computationally expensive than the Eigenvalue problem. Consequently, this method of counting paths is more computationally expensive than the method based on the Eigenvalue problem. Using the available functions of Matlab<sup>R</sup> software to solve the Eigenvalue problem consisting of 5,000 nodes and 5,000 pipes, which is close to the size of the WDN in the city of Mount Pearl in Newfoundland and Labrador, the time required for the Eigenvalue problem is just 130 seconds, and calculating component  $AA^{n-1}$  only takes 430 seconds.

### AC for multiple pipe breaks

However, for multiple breaks, it is challenging to compare  $AC$ s. First, the number of computations increases significantly since many more sets of pipes are to be considered when the size of minimum cut-sets (i.e., number of pipes

in the minimum cut-set) is large. For example, if finding the minimum cut-sets containing one pipe break requires  $m$  Eigenvalue problems to be solved, approximately  $m(m-1)/2$  Eigenvalue problems need to be solved to determine all the minimum cut-sets containing two simultaneous pipe breaks. Second, a combination of breaks can lead to a 'noisy' situation, in which both the disconnection and reduction of redundancy may occur. Once this 'noise' happens, the  $\Delta AC$  may decrease (e.g.,  $E_{2,3}$  in Table 2) or increase (e.g.,  $E_{2,1}$  in Table 2), depending on the network structure. If  $\Delta AC$  decreases, it may fail to recognize the disconnection. This is further demonstrated below through application of the hypothetical network in Figure 2.

Table 4 presents the  $AC$ s for all possible minimum cut-sets of the network shown in Figure 2. In Table 4, minimum cut-set  $E_{2,3}$  leads to disconnection of node 7; however, the change in  $AC$  is negative. Thus, the  $\Delta AC$ -based method would fail to identify  $E_{2,3}$ . This phenomenon is termed herein as 'noise'. The reason for the negative  $\Delta AC$  is that the increase of  $AC$  due to disconnection of node 7 is less than the decrease of  $AC$  due to loss of redundancy.

Table 4 | Minimum cut-sets and associated  $AC$  values

Min. cut-set ( $E_{k,pd}$ )	Pipes in the min. cut-set ( $E_{k,pd,1pd,1j}; \dots; E_{k,pd,1pd,kj}$ )	AC	$\Delta AC$
$E_{1,1}$	5–6	0.5188	0.3210
$E_{1,2}$	8–9	0.3512	0.1534
$E_{1,3}$	9–10	0.2662	0.0684
$E_{2,1}$	1–3; 3–4	0.2230	0.0252
$E_{2,2}$	6–7; 6–8	0.5858	0.3880
$E_{2,3}$	6–7; 7–8	0.1823	-0.0155
$E_{2,4}$	6–8; 7–8	0.4384	0.2406
$E_{3,1}$	1–2; 1–3; 1–5	0.2051	0.0073
$E_{3,2}$	1–2; 1–5; 3–4	0.2509	0.0531
$E_{3,3}$	1–2; 2–4; 2–5	0.2104	0.0126
$E_{3,4}$	1–3; 2–4; 4–5	0.2509	0.0531
$E_{3,5}$	1–5; 2–5; 4–5	0.4131	0.2153
$E_{3,6}$	2–4; 3–4; 4–5	0.2051	0.0073
$E_{4,1}$	1–2; 1–3; 2–5; 4–5	0.3004	0.1026
$E_{4,2}$	1–2; 1–5; 2–4; 4–5	0.3004	0.1026
$E_{4,3}$	1–2; 2–5; 3–4; 4–5	0.2137	0.0159
$E_{4,4}$	1–3; 1–5; 2–4; 2–5	0.2137	0.0159
$E_{4,5}$	1–5; 2–4; 2–5; 3–4	0.3004	0.1026

Note that all of the pipes within the set of a minimum cut-set have to break simultaneously to cause disconnection of the network or to increase the *AC*. Removal of only one pipe from the set does not lead to disconnection or an increase in *AC*. Thus, the *AC* of a network with removal of a pipe from the set of a minimum cut-set is always lower than the *AC* of the network with removal of all pipes in the minimum cut-set. This characteristic can be used to eliminate the noise, as discussed below.

To implement this, when  $\Delta AC$  decreases with a set of *b* breaks, the *AC* of the network is calculated with removal of each of the pipes within the set, i.e.,  $AC_u^*$  for removing  $u^{th}$  pipe ( $u = [1:b]$ ). If the largest  $AC_u^*$  is less than the *AC* with removal of the full set of pipes, then disconnection occurs. For example, removal of the set of pipes 6–7 and 6–8 causes a negative  $\Delta AC$  value in Table 4, with an *AC* of 0.1823. The *AC* of the network with the break of pipes 6–8 and 6–7 ( $AC_1^*$  and  $AC_2^*$ ) are 0.1341 and 0.1672, respectively. These two values are lower than the *AC* of the event  $E_{2,3}$  (0.1823). Therefore, the set of pipes 6–7 and 7–8 is a minimum cut-set that leads to disconnection. In this manner, the noise removal is applied to the network (Figure 2) and the minimum cut-sets of the network based on the change in *AC* are successfully determined and tabulated in Table 4.

## RELIABILITY ASSESSMENT USING AC

The exact solution of the network disconnection probability can be obtained theoretically by rigorously examining the minimum cut-sets using FTA, discussed above. However, this process requires tedious and complex computations for a large-scale network. Therefore, a practical approach is required, using assumptions to overcome the difficulties. This section develops a framework to estimate network disconnection probability based on the change of *AC*. It was revealed earlier in this paper that a single event-based failure probability provides a reasonable approximation of the system failure probability (i.e.,  $P(E) \approx P(E_1)$ ). Both single event-based and multiple event-based failure probabilities are considered here for the *AC*-based reliability assessment. For the failure probability of the pipe components, a constant failure rate-based model, a failure function (i.e., exponential function) and Monte Carlo simulation (MCS) are considered.

## Single event-based approximate failure probability

When the pipe breaking is assumed as a Poisson process with a constant failure rate, which depends on the length of the components, the mean pipe failure rate of a component can be related to the length of the component using Equation (9). It can be expressed for the minimum cut-set containing one pipe as:

$$-\lambda_{1,pd} = -\frac{Br}{l_{total}} l_{1,pd} \quad (19)$$

where  $l_{1,pd}$  is the length of pipe corresponding to the  $pd^{th}$  minimum cut-set containing only one pipe.

Approximating  $P(E) \approx P(E_1)$ , as described in Equation (13), the failure probability of the system can be defined as:

$$\begin{aligned} P(E) \approx P(E_1) &= 1 - \prod_{pd=1}^{p1} (1 - P(E_{1,pd})) \\ &= 1 - \prod_{pd=1}^{p1} (e^{-\lambda_{1,pd} \times t}) = 1 - e^{t \sum_{pd=1}^{p1} (-\lambda_{1,pd})} \end{aligned} \quad (20)$$

Substituting Equation (19) with Equation (20):

$$P(E, t) \approx 1 - e^{(-\frac{Br \times t}{l_{total}} \sum_{pd=1}^{p1} (l_{1,pd}))} \quad (21)$$

For a group of pipes with minimum cut-set, the total length of the group is given by:

$$l_{g1} = \sum_{pd=1}^{p1} (l_{1,pd}) \quad (22)$$

Using the length of the group of pipes, the failure probability of the system can be defined as:

$$P(E, t) \approx 1 - e^{(-\frac{l_{g1} \times Br \times t}{l_{total}})} \quad (23)$$

However, the pipe breaking is not necessarily a Poisson process with or without a constant failure rate. Particularly, the WDN is a system containing a mixture of pipes ranging from newly installed segments to wear-out segments. The age-dependent degradation of WDN pipe segments is an unavoidable mechanism, because these structures are exposed directly to a hazardous environment, leading to corrosion and cracking. Once a pipe is at its wear-out stage, the assumption of a constant break rate may no longer be adequate. The degrading pipe segments would need a more precise failure function with

sufficient input data to describe the failure process with an increasing failure probability. For any other component failure model, the failure probability of the system can be defined based on the assumption of  $P(E) \approx P(E_1)$  as:

$$P(E, t) \approx P(E_1, t) = \left[ 1 - \prod_{pd=1}^{pk} (1 - P(E_{1,pd}, [pd,1], t)) \right] \quad (24)$$

A detailed failure probability estimation procedure using either Poisson process-based failure functions or any other failure functions is illustrated in Figure 4.

### Multiple event-based failure probability

Figure 5 illustrates the detailed procedure of calculating system failure probability considering rigorous multiple events (i.e., minimum cut-sets). As discussed earlier, ‘noise’ may appear in the failure probability assessment for minimum cut-set containing multiple pipes. Thus, there can be some disconnection events with negative  $\Delta AC$ . A sub-algorithm (Figure 5) discussed in the section ‘AC for multiple pipe breaks’ can be applied to avoid the noise problem. For

calculation of failure probability of the components, application of a Monte Carlo (MC) simulation is used.

The methods discussed above are applied to the network presented in Figure 2 through development of codes using Matlab<sup>R</sup> software. The calculated failure probabilities are compared in Figure 6. For the failure function, an equivalent exponential model is considered for illustration. The failure probability based on rigorous investigation of the network using FTA (without the use of AC) is also calculated and included in the figure. Among AC-based methods, the approximate failure probabilities calculated using a Poisson process-based and exponential component failure model are the same in Figure 6, since equivalent reliability models are used for the components. In the developed framework, different reliability models can be used (e.g., Weibull or lognormal distribution) instead of an exponential model, if the distribution is known.

The MC simulation conducted using 10,000 trials each time provides multiple event-based failure probabilities, which is very close to the failure probabilities calculated using single event-based approximation. Note that the MC simulation without the ‘noise’ removal sub-algorithm fails to predict the failure probability (Figure 6).

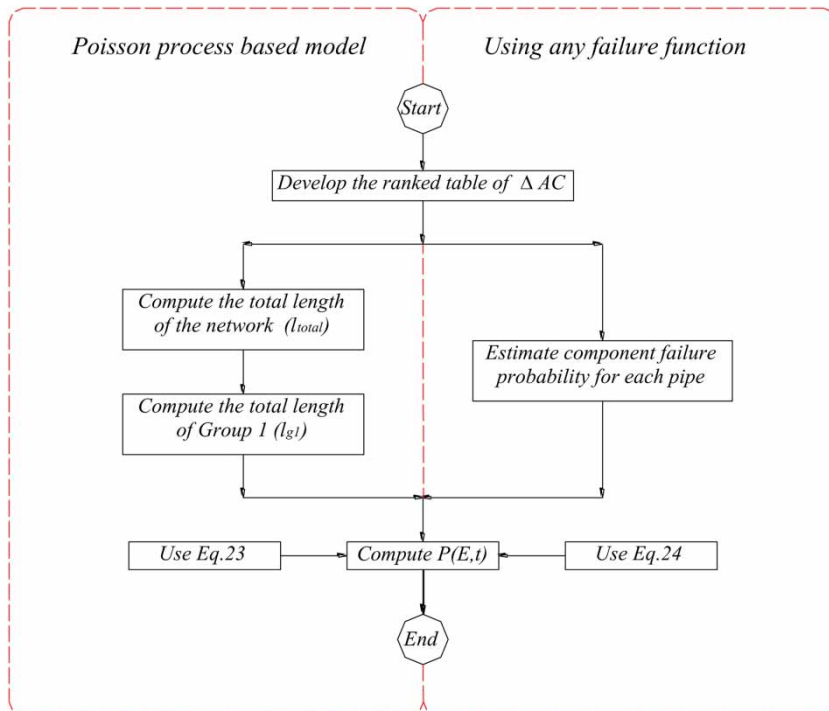
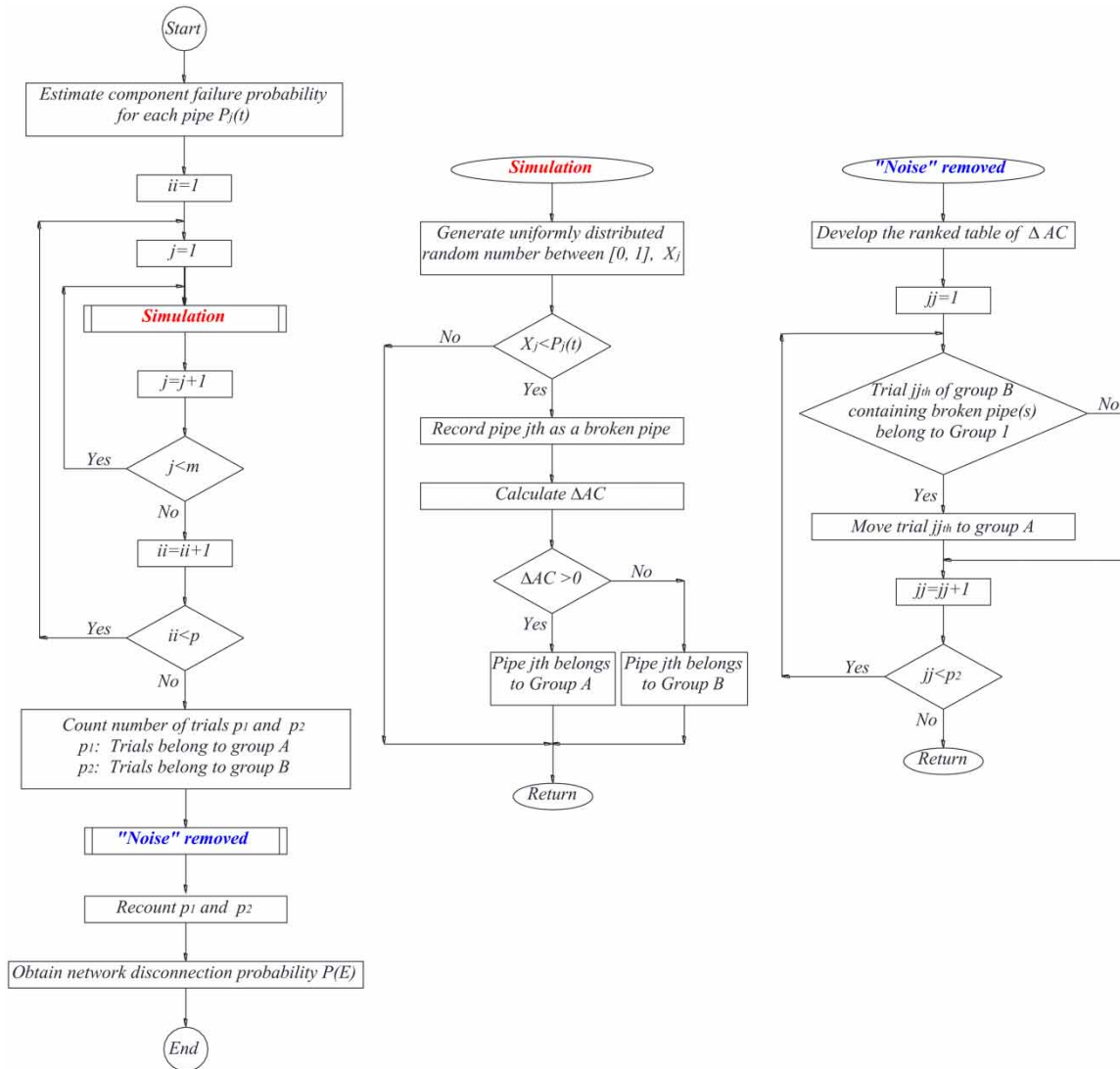


Figure 4 | Flow chart for single event-based approximate failure probability ( $P(E) \approx P(E_1)$ ).



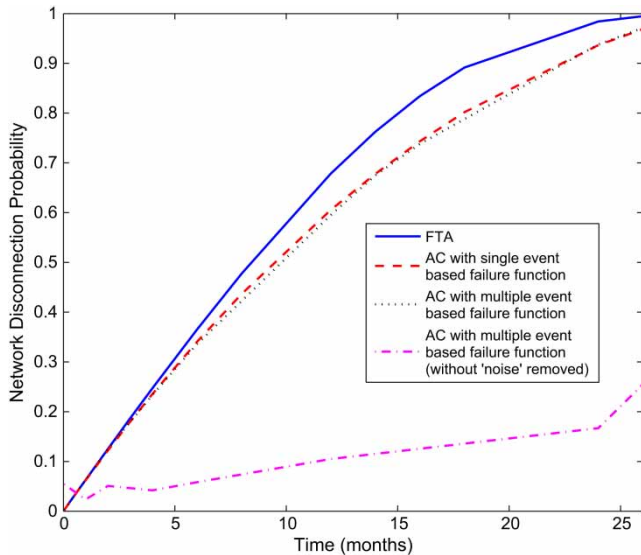
**Figure 5** | Flow chart for multiple event-based failure probability calculation.

In [Figure 6](#), the failure probability calculated using the FTA and AC methods is not significantly different. The network disconnection probability reaches close to 100% after 26 months, according to the rigorous FTA solution. The AC-based method with reliability algorithms predicts the failure probability as 97% (after 26 months).

## A CASE STUDY OF THE CITY OF MOUNT PEARL

The developed method is applied to the network in the city of Mount Pearl, which is a community with a population of

approximately 25,000. The water main system contains 130 km of pipe network (see schematic shown in [Figure 7](#)). Thus, on average, 192 people are served by each kilometre of water main, which is comparable to common WDNs in the USA and Canada, which serve 164 people per kilometre ([Folkman 2012](#)). Diameters of the pipes in the city range from 150 to 600 mm. Water is bought from St John's – a neighbour city – to two water tanks located at the north (S1) and south-west (S2) parts of the city. The WDS to support the entire city is based on gravity flow. The locations of the sources are ignored in this study, for simplification. However, they could be useful for further



**Figure 6** | Comparison of network disconnection probability with different approaches.

study, which defines system failure as the disconnection of the source to nodes or reachability of the network. A sample of a raw database for the network is presented in Table 5. The table is included to show the typical WDN data that can be used for the AC-based method. After reorganizing the database, it is found that the WDN of the city of Mount Pearl contains 4,848 nodes and 5,046 pipes (Table 5).

This pipe system generally comprises ductile iron (91%), cast iron pipe (5%) and other materials (4%). Based on the annual water main break records of the city, the annual break rate is 6.931 breaks/year/100 km. This break rate is comparable to the average break rate of 6.77 breaks/year/100 km reported in Folkman (2012). It is understood that the break rate based on limited records is not sufficient for failure assessment, but is acceptable for the purpose of illustration.

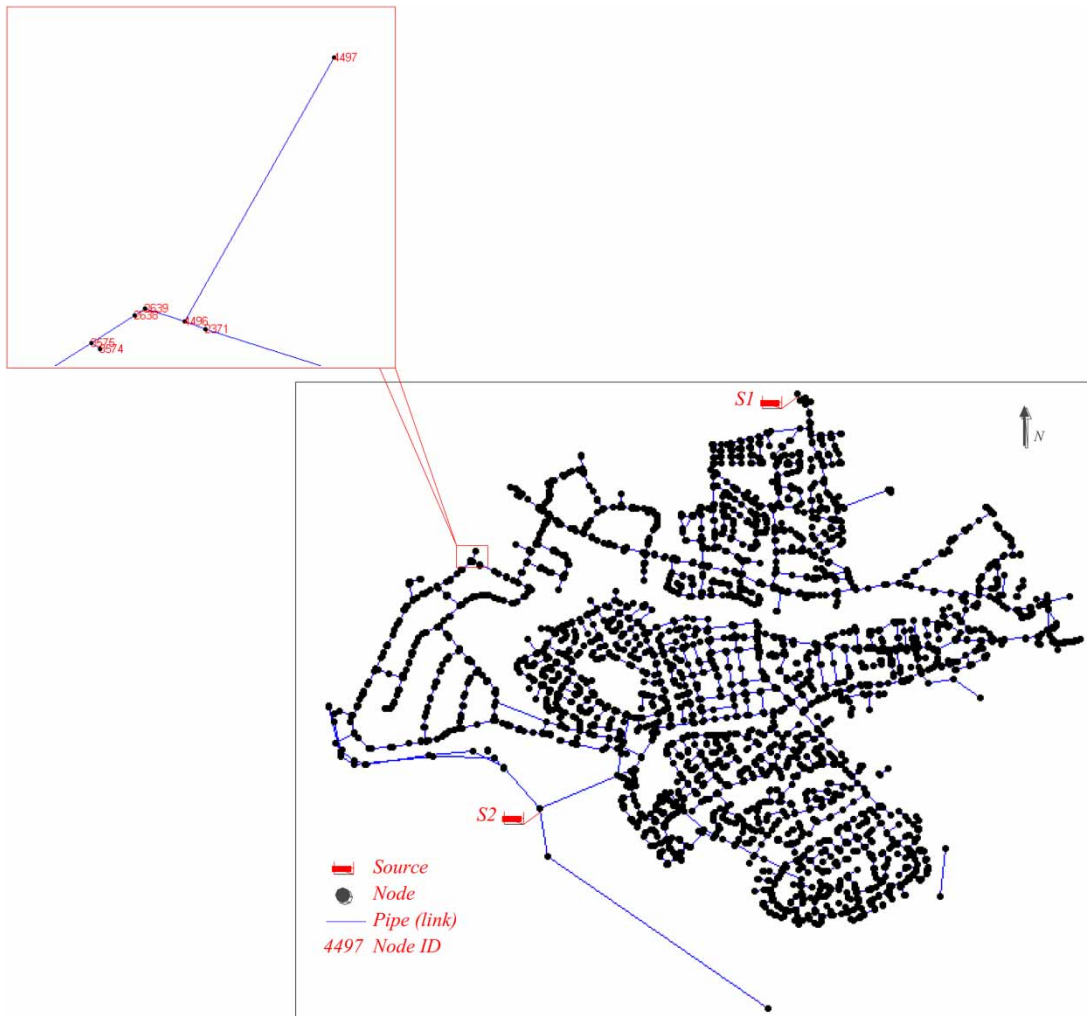
Phan et al. (2017) calculated different parameters used in complex network analysis to assess the robustness and redundancy of the network for the WDN of the city. Yazdani & Jeffrey (2011a) provided a summary of the parameters used to assess the redundancy and robustness of networks. The parameters calculated for the network in the city of Mount Pearl are provided in Table 6, and are compared with values for different other cities, as reported in Yazdani & Jeffrey (2011a). In Table 6, the parameters for the city of Mount Pearl are mostly within the ranges of those for the other cities, indicating that the redundancy, connectivity and robustness of the city network are similar to those for other

cities reported in Yazdani & Jeffrey (2011a). The redundancy, connectivity and robustness parameters of the Mount Pearl network range from the average to lower bound values of the parameters in the networks reported in Yazdani & Jeffrey (2011a). The low well-connectivity can be observed since the network contains high betweenness pipes, the removal of which could lead to the separation of the network.

### Identification of the critical pipes

For identification of critical pipes using AC, the pipes can be ranked according to the change in AC ( $\Delta AC$ ) with and without removal of a pipe from the network. Sample results of ranking of the pipes in the city of Mount Pearl's WDN are presented in Table 7. The histogram of the  $\Delta AC$  is presented in Figure 8. According to Table 7, there are 2,061 out of 5,046 pipes belonging to Group 1 that directly lead to disconnection, and the remaining pipes belonging to Group 2 cause a reduction of redundancy. From Figure 8, it can be seen that the  $\Delta AC$  generally ranges from  $-2.5 \times 10^{-6}$  to  $2.5 \times 10^{-6}$  with some outliers that have significantly high positive values of up to  $40 \times 10^{-5}$ . This sudden jump in  $\Delta AC$  marks the most critical pipes, the failures of which would lead to an overall structural change in the network (i.e., disconnection on a large scale). Group 2 does not contain such outliers. Even though Group 2 has no outliers, it is helpful to understand the reduction of redundancy on a macro scale. To investigate the groups of pipes further, a set of the top 50 ranked pipes from Group 1 is analysed in more detail. The most important pipes of Group 2 (first 100 ranked pipes) have also been analysed, as illustrated in Figure 9.

The critical pipes in the network are also identified using the shortest path (SP) method for comparison. In the SP method, the shortest path to all pairs of nodes within the network is investigated. The method is thus computationally more demanding than the AC-based method. A pipe component appearing in the highest number of the shortest path is considered to be most critical, since water is expected to flow through this pipe component to a maximum number of destinations. The pipes are therefore ranked according to the number of their appearances on the shortest paths. The pipes with higher rank are considered to be more critical. Figure 10 shows the top 100 highest ranked pipes (thick solid lines) in the city of Mount Pearl's network.



**Figure 7** | Schematic of water main network in city of Mount Pearl.

**Table 5** | Water main database for city of Mount Pearl

Node ID	Shape	Pipe ID	Material	Diameter (mm)	POINT_X	POINT_Y
1	Point	1	Ductile iron	150	320991.62230	5264475.70150
2	Point	1	Ductile iron	150	321007.09150	5264498.97460
3	Point	2	Ductile iron	200	321048.28280	5265709.24110
...	...	...	...	...	...	...
8564	Point	9474	DI Private (approx.)	150	320423.51620	5263016.32110

As seen in the figure, the critical pipes connect the northern and southern communities. The highest ranked pipe is the one laid between the Southern community, 1, and Northern community, 2 (Figure 10). Note that the sub-communities are not disconnected with removal of the pipe, since there

are several other paths (redundancy) to connect the two sub-communities. An alternative path is shown as bold black lines in Figure 10. Thus, the SP method provides information for the important pipes (critical pipes), which may or may not lead to the separation of part(s) of the network. The

**Table 6** | Complex network system parameters for the city of Mount Pearl (Phan et al. 2017)

Parameter	Equation	Calculated value	Reference values <sup>a</sup>
Total length	$l_{total} = \sum_{i=1}^n \sum_{j=1}^n l_{ij}$	129.837 (km)	
Number of pipe	–	5.046	769–3,065
Number of node	–	4.848	755–2,799
Link-per-node	$e = \frac{m}{n}$	1.041	1.01–1.10
Average node degree	$k_a = \frac{2m}{n}$	2.082	2.04–2.23
Link density	$q = \frac{2m}{n(n-1)}$	$4.29 \times 10^{-4}$	$7.83 \times 10^{-4}$ – $2.7 \times 10^{-5}$
Independent loop	$f = m - n + 1$	197	–
Meshness coefficient	$r_m = \frac{f}{2n-5}$	0.0203	$9.97 \times 10^{-3}$ – $5.86 \times 10^{-2}$
Threshold for random removal of node	$f_c = 1 - \frac{1}{\frac{\text{mean}(k^2)}{k_a} - 1}$	0.1770	0.22–0.42
Route factor	$g = \frac{1}{1-n} \sum_{i=1}^{n-1} \frac{s_i}{s_i}$	1.7560	1.45–1.67
Characteristic path length	$l = \frac{1}{1-n} \sum_{i \neq j} d_{ij}$	88.378	25.94–51.44
Algebraic connectivity	–	$1.5625 \times 10^{-5}$	$6.09 \times 10^{-5}$ – $2.43 \times 10^{-4}$
Spectra gap	–	$2.55 \times 10^{-2}$	$9.08 \times 10^{-3}$ – $7.27 \times 10^{-2}$

<sup>a</sup>From Yazdani & Jeffrey (2011a).

**Table 7** | The changes of AC for the city of Mount Pearl WDN

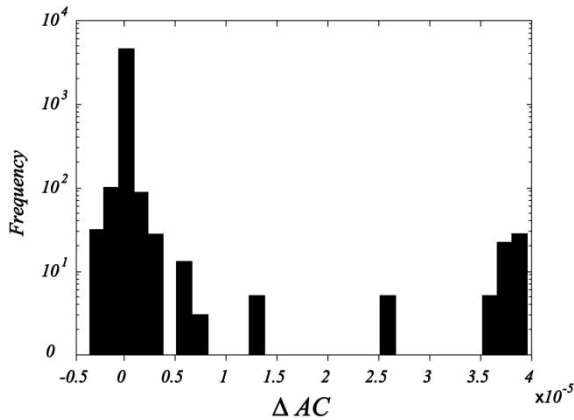
Removed pipe ID	AC	$\Delta AC = AC - AC_{intact}$	Type of network structural change	Grouped as
1734–1735	$5.52 \times 10^{-5}$	$3.96 \times 10^{-5}$	Disconnection	Group 1
1734–3518	$5.52 \times 10^{-5}$	$3.96 \times 10^{-5}$	Disconnection	Group 1
1733–3518	$5.52 \times 10^{-5}$	$3.95 \times 10^{-5}$	Disconnection	Group 1
...	...	...	...	...
2373–2631	$1.21 \times 10^{-5}$	$-3.56 \times 10^{-6}$	Less redundancy	Group 2
1655–2631	$1.21 \times 10^{-5}$	$-3.57 \times 10^{-6}$	Less redundancy	Group 2
1655–1656	$1.21 \times 10^{-5}$	$-3.57 \times 10^{-6}$	Less redundancy	Group 2

method cannot be applied to identify the pipes that may lead to network disconnections.

It can be seen that the top ranked pipes based on the  $\Delta AC$  (Figure 9) matched well with the shortest path results presented in Figure 10. Thus, the  $\Delta AC$  could effectively classify the critical pipes in Group 1 and Group 2. The ranked  $\Delta AC$  table can also be used to identify the alternate path of connecting two points, which cannot be obtained from the shortest path method.

### Estimating the system disconnection probability

Disconnection probability of the city of the Mount Pearl WDN is analysed using the methods discussed above. The results are presented in Table 8 and Figure 11. Time periods of 1, 2, 6 and 12 months are used in the analysis. From Table 8, it can be seen that the  $P_{Br,av}$  of the network is significantly low compared with the example in Figure 2. Thus, the calculated failure probabilities using the single



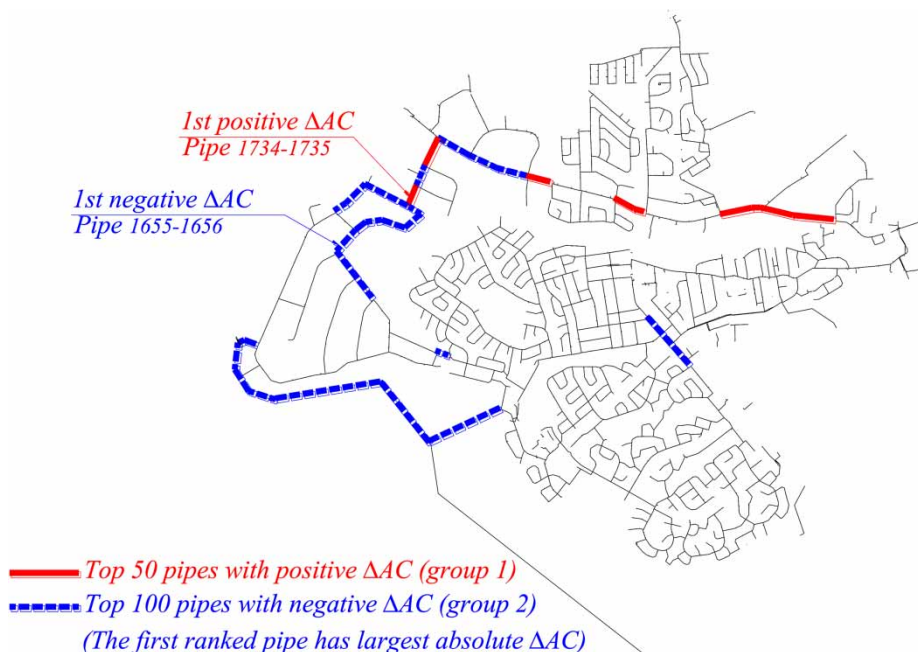
**Figure 8** | Histogram of the  $\Delta AC$  with single break from Table 7.

event-based approach and the multiple event-based approach are almost the same in Figure 11.

## CONCLUSIONS

In an existing WDS, pipe breaks causing isolation of part of the system should be avoided, to provide uninterrupted service to the dwellers. Studying the disconnection probability of WDS is therefore important for optimizing repair/maintenance prioritization planning. A pipe or combination of

pipes causing network isolation can be identified using the minimum cut-sets method. In the conventional minimum cut-set method, pipes causing system failure (i.e., isolation) are rigorously examined to determine a combination of the minimum number of components' failure leading to the system failure. The process of finding minimum cut-sets is very complex and requires extensive computation. This paper presents a method of finding minimum cut-sets using AC for complex network analysis. The AC-based approaches reduced the amount of computation required compared to the existing methods, including the graph theory used in Yannopoulos & Spiliotis (2013). The AC is a network parameter which is determined by solving Eigenvalue problems that avoids counting the paths, as described in Yannopoulos & Spiliotis (2013). The decrease of AC due to a pipe break indicates a reduction of redundancy and the increase of AC indicates disconnection (isolation) of parts of the network. For the events with multiple pipe breaks, some breaks might cause an increase in AC while others might cause a decrease in AC, resulting in a 'noisy' situation. A methodology is developed to eliminate the 'noise' for multiple pipe breaks through scrutinizing the pipes in the minimum cut-sets. A failure probability estimation framework is developed using the AC-based minimum cut-sets.



**Figure 9** | Top ranked pipes of Group 1 and Group 2.



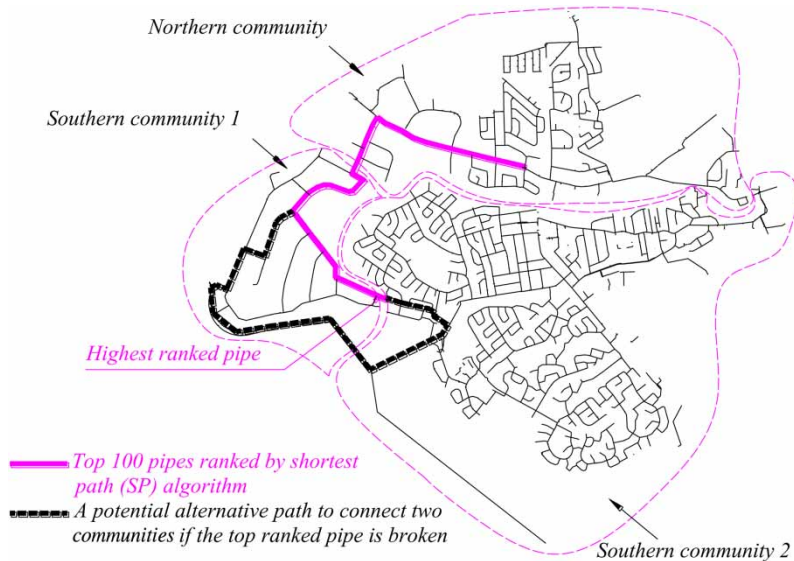


Figure 10 | Results for shortest path analysis.

Table 8 | Comparison of failure probabilities

$\Delta t$ (month)	Average number of breaks	$P_{Br.av}$	$P(E_r, \Delta t)$ single event based	$P(E_r, \Delta t)$ multiple event based
1	0.716	$1.486 \times 10^{-4}$	0.243	0.229
2	1.518	$2.973 \times 10^{-4}$	0.427	0.436
6	4.479	$8.918 \times 10^{-4}$	0.812	0.814
12	8.834	$17.836 \times 10^{-4}$	0.964	0.973

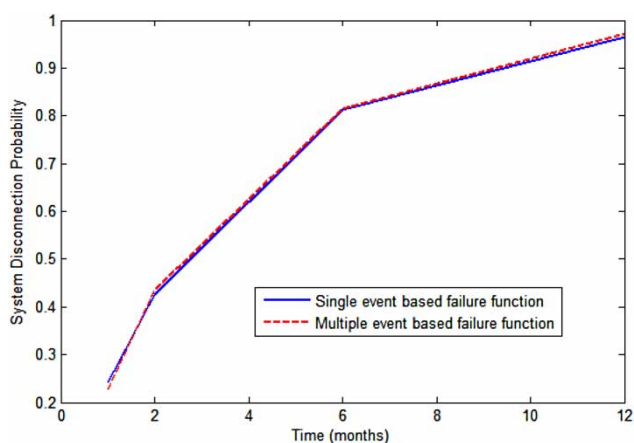


Figure 11 | System disconnection probability.

The proposed method employs removal of pipes one by one from the network and subsequent estimation of the change in  $AC$  ( $\Delta AC$ ) for estimating the network failure probability and ranking pipes within the network according to

their importance level. Based on these steps, the critical pipes in the network can be identified. The suitability of the proposed method is demonstrated through application, first to a simple hypothetical network and then to a medium-sized real WDS. The study reveals that the single event-based approximate failure probability can reasonably be used for system probability assessment, saving extensive computational time. The critical and high betweenness pipes identified using the  $AC$ -based method are validated through comparison with the results of the SP method. The SP method is a distance-related measurement in CNA, which is useful for identifying high betweenness pipes or nodes (Newman 2010). However, the SP method cannot be used for identification of network isolation.

The proposed reliability model considers the probability of disconnection only. Hydraulic availability of the demand nodes is not considered. Thus, lower bound values of the system failure probability are obtained. The developed

model examines constant annual break rate-based component failure functions.

## ACKNOWLEDGEMENTS

The financial support for this research has been provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) through its Collaborative Research and Development Grants and the city of Mount Pearl in the province of Newfoundland and Labrador. The city of Mount Pearl provided necessary information regarding their WDN along with the past water main break records. The comments from anonymous reviewers are appreciated.

## REFERENCES

- Awumah, K., Goulter, I. & Bhatt, S. 1991 [Entropy-based redundancy measures in water distribution network design](#). *Journal of Hydraulic Engineering* **117** (5), 595–614.
- Aydin, N. Y., Mays, L. W. & Schmitt, T. 2014 [Sustainability assessment of urban water distribution system](#). *Water Resources Management* **28**, 4373–4384.
- Capocci, A., Servedio, V. D., Caldarelli, G. & Colaiori, F. 2005 [Detecting communities in large networks](#). *Physica A: Statistical Mechanics and its Applications* **352** (2), 669–676.
- Coppersmith, D. & Winograd, S. 1990 [Matrix multiplication via arithmetic progressions](#). *Journal of Symbolic Computation* **9** (3), 251–280.
- Cullinane, M. J., Lansey, K. E. & Mays, L. W. 1992 [Optimization-availability-based design of water-distribution networks](#). *Journal of Hydraulic Engineering* **118** (3), 420–441.
- Fiedler, M. 1973 Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal* **23** (2), 298–305.
- Flannery, B. P., Press, W. H., Teukolsky, S. A. & Vetterling, W. 1992 *Numerical Recipes in C*. Press Syndicate of the University of Cambridge, New York, p. 24.
- Folkman, S. 2012 *Water Main Break Rates in the USA and Canada: A Comprehensive Study*. Utah State University, Utah, USA.
- Ghosh, A. & Boyd, S. 2006 Growing well-connected graphs. In: *Proceedings of the 45th IEEE Conference on Decision and Control, IEEE*, San Diego, CA, USA, pp. 6605–6611.
- Goulter, I. C. & Coals, A. V. 1986 [Quantitative approaches to reliability assessment in pipe networks](#). *Journal of Transportation Engineering* **112** (3), 287–301.
- Goulter, I. C. & Kazemi, A. 1988 [Spatial and temporal groupings of water main pipe breakage in winnipeg](#). *Canadian Journal of Civil Engineering* **15**, 91–97.
- Kleiner, Y. & Rajani, B. 2010 [I-WARP: Individual water main renewal planner](#). *Drinking Water Engineering and Science* **3** (1), 71–77.
- Nazempour, R., Monfared, M. A. S. & Zio, E. 2016 A complex network theory approach for optimizing contamination warning sensor location in water distribution networks. *arXiv preprint arXiv:1601.02155*.
- Newman, M. E. 2006 [Finding community structure in networks using the eigen-vectors of matrices](#). *Physical Review E* **74** (3), 036104.
- Newman, M. 2010 *Networks: An Introduction*. Oxford University Press, Oxford, UK.
- Phan, H. C., Dhar, A. S. & Sadiq, R. 2017 Complex network analysis for water distribution systems by incorporating the reliability of individual pipes. In: *6th CSCECRC International Construction Specialty Conference, CSCE 2017 Annual General Conference*, Vancouver, BC, May 31–June 3, 2017.
- Quimpo, R. G. & Shamsi, U. M. 1991 [Reliability-based distribution system maintenance](#). *Journal of Water Resources Planning and Management* **117** (5), 321–339.
- Shamir, U. & Howard, C. D. 1981 [Water supply reliability theory](#). *Journal of the American Water Works Association* **73** (7), 379–384.
- Shinstine, D. S., Ahmed, I. & Lansey, K. E. 2002 [Reliability/availability analysis of municipal water distribution networks: case studies](#). *Journal of Water Resources Planning and Management* **128** (2), 140–151.
- Su, Y. C., Mays, L. W., Duan, N. & Lansey, K. E. 1987 [Reliability-based optimization model for water distribution systems](#). *Journal of Hydraulic Engineering* **113** (12), 1539–1556.
- Tung, Y. K. 1985 Evaluation of water distribution network reliability. In: *Hydraulics and Hydrology in the Small Computer Age, Proceedings of the Specialty Conference ASCE*, Lake Buena Vista, FL, pp. 359–364.
- Wagner, J. M., Shamir, U. & Marks, D. H. 1988 [Water distribution reliability: analytical methods](#). *Journal of Water Resources Planning and Management* **114** (3), 253–275.
- Yannopoulos, S. & Spiliotis, M. 2013 [Water distribution system reliability based on minimum cut-set approach and the hydraulic availability](#). *Water Resources Management* **27** (6), 1821–1836.
- Yazdani, A. & Jeffrey, P. 2010 Robustness and vulnerability analysis of water distribution networks using graph theoretic and complex network principles. In: *12th Annual Conference on Water Distribution Systems Analysis*, Tucson, AZ, pp. 933–945.
- Yazdani, A. & Jeffrey, P. 2011a [Complex network analysis of water distribution systems](#). *Chaos: An Interdisciplinary Journal of Nonlinear Science, ASCE* **21** (1), 933–945.
- Yazdani, A. & Jeffrey, P. 2011b [Applying network theory to quantify the redundancy and structural robustness of water distribution systems](#). *Journal of Water Resources Planning and Management* **138** (2), 153–161.

First received 30 May 2017; accepted in revised form 19 January 2018. Available online 7 March 2018