Title
Resilience Analysis for Water Distribution Networks

Author
Wu, Renjie

Publication Date
2022

Peer reviewed|Thesis/dissertation

Resilience Analysis for Water Distribution Networks

by

Renjie Wu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kenichi Soga, Chair
Associate Professor Dimitrios Zekkos
Associate Professor Adityanand Guntuboyina

Spring 2022

Resilience Analysis for Water Distribution Networks

Abstract

Resilience Analysis for Water Distribution Networks

by

Renjie Wu

Doctor of Philosophy in Engineering - Civil and Environmental Engineering

University of California, Berkeley

Professor Kenichi Soga, Chair

Cities and people who live in them require clean water to thrive. However, the conditions of water infrastructure in the United States are concerning. Many water distribution networks (WDNs) are serving beyond their intended design life. The aging water infrastructure is reflected by the frequent water main break events across the country in recent years. Using data mining and natural language processing (NLP) techniques, this study validates the commonly-hold opinions that water main breaks cause severe societal troubles, including repair cost, local traffic disturbance, and water quality related health issues.

Hazard events often damage the already in-risk aging WDNs. To estimate the hazard impact (e.g. earthquake) to a WDN, this study developed a WDN hydraulics simulator, HydrauSim, which can quantify the WDN hydraulics (e.g., flow rate, etc.) under normal or damaged states. HydrauSim is highly optimized to be computationally efficient, making it feasible for large-scale networks and tasks that require repeated simulation runs (e.g., Monte Carlo simulation). Using HydrauSim, the post-earthquake response of East Bay Municipal Utility District's (EBMUD) main gravity feed zone in the San Francisco Bay area is simulated. Around 200–800 pipes were estimated to break during the simulated earthquake events. On average, 25% of demand nodes may experience insufficient water pressure levels, which can rise to 78% for the worst-case scenario.

In real-life situations, failed pipes need to be isolated from the main network by closing the corresponding isolation valves to prevent the effects of individual events from spreading throughout the system. However, most utilities do not have sufficient valves installed, and the installed ones may malfunction at the time of usage. This study proposed an analysis framework for WDN pipe isolation risk considering valve condition uncertainties. It is found that the magnitude of the risk depends on the mean and variance of isolation segment sizes and demand distribution across the network.

Using dynamic programming, an optimal valve placement algorithm is developed to find the

best place to install isolation valves to minimize the system risk. The proposed method is tested on two real-life WDNs. Comparing to the existing valve placement configuration, the proposed configuration significantly reduces the pipe isolation risk of the system. Furthermore, the proposed valve placement strategy produces a more robust network than the original one regarding valve failure scenarios. Pipe isolation risks are significantly reduced at all tested failure rates, and the risk-increasing trend (as the valve failure rate increases) is effectively restrained.

Due to resource constraints, it is impractical for water utilities to maintain all the isolation valves in a system. This study proposes a method to rank the isolation valves based on their potential failure consequences. The valve ranking algorithm utilizes network analysis methods and machine learning techniques to label valve maintenance priorities automatically. Simulation on real-life WDNs shows that applying the proposed valve maintenance strategy effectively reduces both the direct and indirect risk for the tested networks, especially under high valve failure rate cases.

To my family and friends,

Robustness is a personal trait, but resilience comes from the support of loved ones. I survived because of you. Thank you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction: Aging Water Distribution Systems

1.1 Background

Modern cities rely on complex infrastructure systems. For example, power systems(e.g., electric grid) supply essential energy, transportation systems (e.g., road networks) support material exchanges, and information systems (e.g., cell towers) provide communication channels for a city. Among all infrastructure systems, water supporting systems are particularly important since both life and industry depend on water. Modern cities use Water Distribution Networks (WDNs) to distribute water across the territory, which is composed of water sources (reservoirs, water treatment plants, and water tanks), water transmission lanes (water mains), and auxiliary equipment such as pumps and valves.

There are about 155,693 public water systems in the United States. Among them, 52,110 (33.5%) are community water system (CWS) supplies water to the same population year-round. The rest 103,583 (66.5%) systems are non-community water systems. 84,744 systems are transient, which provide water to different groups of people (25 or more) for at least 60 days/year. 18,839 systems are non-transient non-community water systems that provide water at least six months per year for the same group of people [133]. These systems are composed of around 2.2 million miles of underground pipes that supply clean water to millions of people [71].

Aging WDNs

The conditions of WDNs in the United States are concerning. In 2021, the American Society of Civil Engineers issued a USA Infrastructure Report Card and gave a C- to drinking water and wastewater infrastructure [9]. Overall, the systems are aging and underfunded. Many of the water systems were constructed during three main periods: 1800s, 1900-1945, and post 1945, corresponding to the period of rapid population growth in the country[45]. As a result, many water infrastructure systems are serving beyond their intended design life, leading

to an undesirable consequence of components failure and posing reliability and emergency preparedness challenges to society. It is estimated that the cumulative cost of degraded water/wastewater infrastructure on households could be as high as 59 billion dollars from 2013 to 2020. The economic effect on businesses could be as high as 147 billion dollars [146]. Nevertheless, things have been improving in recent years. Thanks to federal financing programs and raised rates, water utilities start to receive enough funds to reinvest in their networks [102]. In 2020, more than 12,000 miles of water pipes were planned to be replaced nationwide [54].

Components fail in aging WDNs. The most frequent failed components are underground water pipes [9, 46, 45]. In the United States, there is a pipe failure (water main break) event every two minutes [45]. The resulting water loss is estimated at around 6 billion gallons per day [29]. As discussed earlier, most pipes were installed many decades ago. Constrained by the limited knowledge and available technology, WDN asset owners had few design choices when they built the systems, which led to inadequate design and poor installation. Long time exposure to the underground environment also leads to corrosion and material fatigue for the buried pipes, increasing the failure chance [99]. Moreover, since most pipes are buried underground, it is difficult and costly to assess the conditions of the installed pipes, leading to inadequate maintenance [40]. The rising pipe failure rate and lack of maintenance cause the frequent pipe break (water main break) events nationwide.

It is reported that there are between 250,000 and 300,000 pipe (water main) failures in the U.S. annually [45, 9]. As the ages of WDNs continues to increase, pipe main break events are happening more and more often in recent years. Between 2012 and 2018, overall water main break rates increased by 27% from 11.0 to 14.0 breaks/(100 miles)/year [45]. Many factors may cause a pipe to break [53]:

- Material degradation over time and subsequent loss of structural integrity (e.g., internal corrosion)
- Design defects or construction errors that weaken the system over time (e.g., weak joints)
- Fatigue loading and subsequent localized structural damage
- Adverse environments the system is exposed to (e.g., corrosive soil cover for pipelines)
- External environment impacts (e.g., freezing, construction or digging, earthquake)
- Improper operating schemes (e.g., changes in system pressure, water hammer, and air entrapment)
- Lack of proper maintenance

Utah State University has conducted two comprehensive surveys on water main conditions in North America. In 2012, they surveyed 188 utilities with 117,603 miles of pipes [46]. An

even more thorough survey was conducted six years later, which covered 308 utilities with a total of 197,866 miles of pipe were reported [45]. The major findings are summarized below as the current knowledge base about water main breaks in North America.

- Pipes with different materials tend to have different failure rates. Among commonly used pipe materials in North America, it is reported that Cast Iron (CI) pipes are the most common material for WDN pipes. However, the break rates of CI pipes have increased significantly by 43% from 2012 to 2018 due to their old age (82% of all CI pipes are over 50 years old). For other pipe materials, the break rate of asbestos cement (AC) pipe has increased by 46% in the same 6-year period. Polyvinyl chloride (PVC) pipes had the lowest overall failure rate among all the popular pipe materials due to their resistance to corrosion.

- Pipes can break in different ways. The surveys conclude that the most common failure mode is a circumferential crack, followed by corrosion. These two failure modes account for about 84% of pipe failure reasons. Longitudinal crack is the next common failure mode, and other modes are relatively rare. Pipe failure induced leakages accounts for at least 10% of overall water loss in surveyed utilities.

- The survey found that the average expected life of installed pipe is 84 years. In the 2018 study, 16% of water pipes is estimated to operate beyond their useful lives, nearly doubling the less comprehensive 2012 study estimation. However, only about 0.8% of aging pipes are replaced each year, equivalent to a 125-year replacement schedule. Compared to the estimated 84 years pipe life, such a slow replacement rate is insufficient.

Impacts of water main breaks

As discussed in the previous section, pipe failure events are common in our aging WDNS. As the likelihood of component failures for aging WDN increases, it is crucial to understand the consequences of pipeline failure. However, the knowledge of the overall failure consequence of WDN pipes is limited in the current state [105]. Past literature is summarized as follows. Cromwell et al. developed a Grand Central Model (GCM) to qualitatively describe the water main break consequences [33]. They conceptually divide costs into repair costs, service outage and mitigation costs, utility emergency response costs, costs of lost water, police and emergency costs, and administrative and legal costs. Gaewksi and Blaha studied 30 water main break cases to conclude that the geometric mean of water main break cost is around 500,000 dollars [47]. Yerri et al. studied 20 large diameter water main failures in the US and estimated the overall cost ranges from 1 million to 85 million dollars [146]. In their study, the overall cost of the failure depends on pipeline size, relative elevation, population density, utility response and repair time, and WDN operation pressure.

Most literature classified the cost of water main failure into two categories [146, 47, 105]. First is the direct costs, including water loss, regional supply suspension, and repair

costs, which are related to WDN properties (demand distribution, regional water pressure, pipe cost, etc.). Since the responsible water utilities conduct most repair activities, the direct cost is often certain and well-documented [146]. The other is the indirect social costs. Examples are the cost of property damage (flooded buildings/roads), traffic disturbance (due to flooding or repair works), and public health issues (water contamination). The indirect social costs are often significantly larger than the direct costs. For the 30 cases analyzed by Gaewksi and Blaha, the overall cost can be as low as 6,000 dollars when the social impact of the water main break event is small. However, the cost can rise to an astonishing 8.5 million dollars, which mostly comes from indirect social costs [47]. Similar findings can be found in Yerri's work. Among the 20 studied cases, 15 of them have higher indirect costs than direct costs [146]. Matthews reported that the traffic costs alone could account for 50% of the total costs of repair work [91].

Although social impacts of water main breaks are important, there is no public available large dataset about them [53, 90]. Previous studies tend to deliberately choose certain events to analyze [47, 146]. Therefore, social impacts of water main break events are often concluded through subjective deductive reasoning using a-prior assumptions during the analysis process. To fill such a knowledge gap, this study attempts to summarize pipe break events information across the United States in an objective way. A dataset containing news articles related to water main break events is created using text data mining techniques over internet news sources. Natural Language Processing (NLP) techniques are performed on the collected data to extract the potential social impacts of the events. The objectively extracted information helps examine the commonly held subjective assumptions, facilitating a more profound understanding of the extent and severity of social implications of water main break events.

Text mining for water main break events

Text mining is a process that extracts patterns or knowledge from unstructured text documents [129]. Text mining is used in various tasks, including document retrieval and representation, email surveillance and filtering, and anomaly detection [16]. Text mining is commonly modeled as a four-step process: online data-gathering, data preprocessing, data indexing/storage, and analysis [129]. The first step in the text mining process is to collect unstructured and semi-structured data from internet data sources like microblogs and news web pages. Next, the collected raw data is cleaned up by removing irrelevant information, and the extracted information is converted into a structured format. The cleaned data is then stored in a local disk or cloud storage. During the analysis phase, underlying patterns and meanings of the collected text can be explored using NLP models. This study utilizes the four-step text mining process to collect and analyze water main break data, with details explained in the following sections.

Data collection

In this study, online news articles are used as the data source for water main break information collection. Other text data such as Twitter posts may contain water main break information besides news articles. However, Twitter posts often tend to be short, unstructured, and ambiguous, which may not contain complete information about the event. Hence, only the online news articles are mined in this study.

The naive way of obtaining news data is to parse related news directly from the news provider websites. However, water main break news only accounts for a minor portion of the daily news. Considering the enormous number of online news sources, the web-crawling, filtering, and parsing process can be highly challenging. Luckily, the web-scraping process has already been performed by search engine companies such as Google and Microsoft over the entire internet. This study uses Microsoft Bing News Search API v7 [94] to collect water main break news articles online. The queried news is organized and stored on a local computer. See Appendix A.1 for details of the data collection process.

The data collection process is carried from 2021-01-04 to 2022-01-04. Over the 365 days data collection period, 7674 news articles are collected from the internet. As shown by Figure 1.1, the number of collected news varies day to day. The max number of collected news is 73, and some days have 0 related news. The standard deviation is about 14. On average, around 21 news articles are recorded every day.

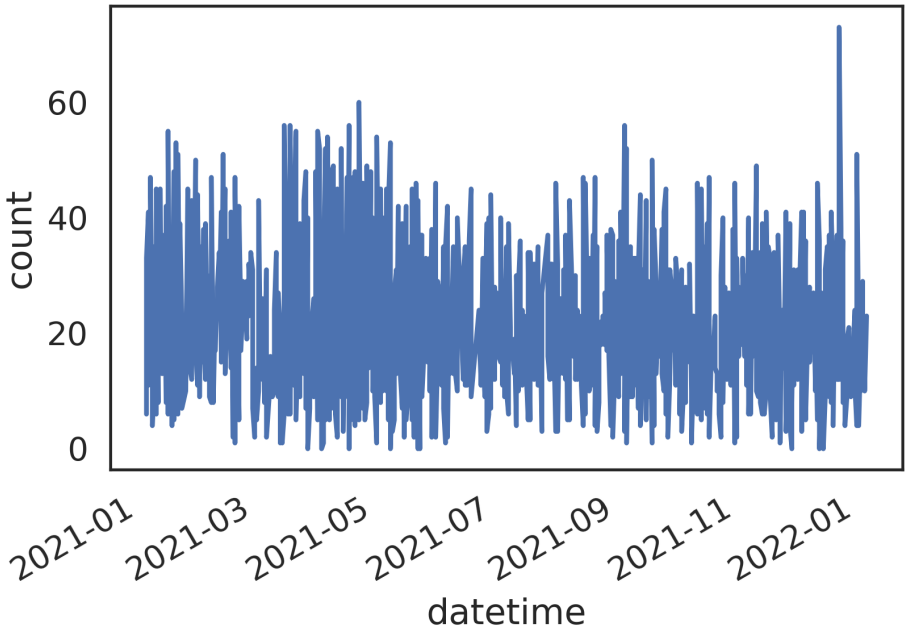


Figure 1.1: Number of collected news across the sampling period

Data preprocessing

Although the Bing News API helps us filter out much irrelevant information, it is not perfect. Reading through the collected news, articles with topics related to society's major concerns at the time of data collection, such as COVID-19 pandemic and 2021 North American winter storm, are accidentally included inside the dataset. A rule-based filtering method is used in this study to remove such irrelevant news, which uses two types of keywords to filter data. The first type is "topic words" (nouns) that are related to the theme we want to explore, such as "water main" and "water lane". The other type includes "action words" (adjectives and verbs) that describe the condition of the water main, such as "break" and "repair". Articles that contain both topic and action words are considered relevant to water main break events. After applying the filtering process to the dataset, only 5405 articles remain relevant to water main break events. Manual inspection of the cleaned dataset validates that most irrelevant articles have been removed.

Figure 1.2 shows top frequent words in the cleaned text data. Words with bigger sizes mean higher appearance frequency in the corpus than small ones. Note that only the top 50 words are shown for visualization purposes. Figure 1.2 provides some clues for the potential social impact of water main break events. The impact of community health can be observed due to the high frequency of sentences like "boil water advisory". The other important observation is that words like street, road, close also have high frequencies, implying the potential traffic disturbance of the events.

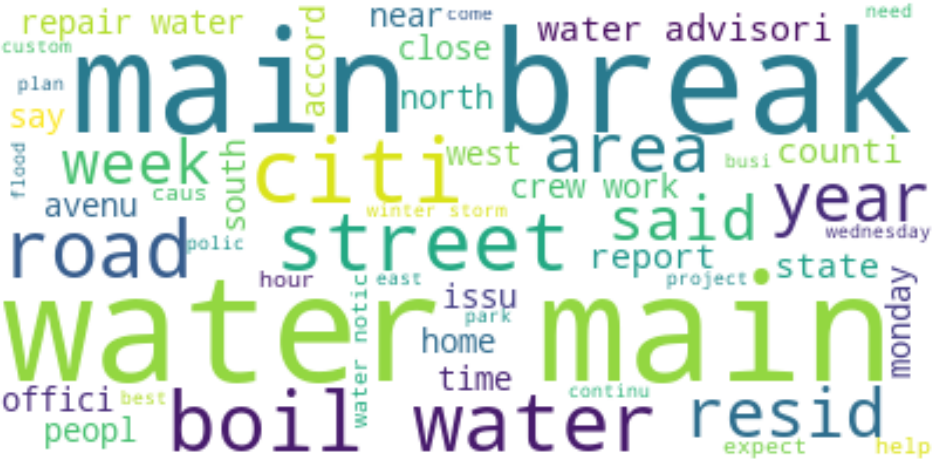


Figure 1.2: Wordcloud visualization of the collected text data

The collected data is further processed. The corpus is constructed by concatenating the article title with the corresponding description. Corpus is then cleaned and prepossessed (e.g., tokenized) and converted to a digital dictionary. Terms (words) in the preprocessed text data are weighted using the frequency-inverse document frequency (TF-IDF) method [107]. See Appendix A.2 for details.

Topic mining models

Since the amount of collected data is large, it is impractical to examine the content of the data manually. Instead, the meaning and patterns of the collected data are analyzed using natural language processing (NLP) techniques. Among NLP algorithms, topic modeling (TM) techniques automatically discover semantic topics from text data [4]. TM techniques have proven to be successful in summarizing organized text documents like news, articles, and books [4]. After experimenting with several state-of-arts TM algorithms, this study adopted the Latent Dirichlet Allocation (LDA) model as the TM method to find the common patterns of the collected water main break news articles. See Appendix A.3 for the details of model choosing.

Results

LDA model fitting requires two inputs. One is the text data, which is the preprocessed TF-IDF corpus in this case. The other is the number of topics that need to be extracted from the input data. This study follows the standard procedure of determining the most fitted number of topics for the data, the trial and error approach with two criteria [4]. Specifically, models with different numbers of topics are tested. The one that yields the largest Jensen–Shannon divergence (measures semantic differences between topics) and the smallest u-mass coherence score (measures coherence inside a topic) is chosen to be the best model. This study finds that the LDA model with four topics produces the best result. The inter-topic map of the selected model is shown by Figure 1.3. All topics are well separate in the projected plane, implying a large Jensen–Shannon divergence. Note that the inter-topic distance is projected into the first two principal components for visualization purposes. The u-mass coherence score is also small, around -3.6 in this case.

Figure 1.4 shows the keywords for each topic using the Wordcloud visualization. The size of a word corresponds to its weight to the topic. In other words, big size words can be viewed as the signature words for the topic (they represent the semantic meaning of the topic). Topic one is related to the traffic impact of the event as it contains keywords such as "road" and "close." Topic two can be interpreted as property damage of the water main break event. Although traffic keywords such as "street", "close" are also considered essential for this topic, keywords such as "flood" and "rescue" have larger weights. Topic three is related to the health impact of the event as water quality-related keywords such as "boil water advisory" or "boil water notice" are important in this topic. Lastly, topic four contains information about the repair/replacement works for the main break events.

As a statistical model, the LDA model assigns the probability of having each topic for the given text data. Using the fitted model, each collected text is labeled by its dominated LDA topic (the one with a significantly higher probability than others, 50% in this case). If the predicted chances of all topics are similar, the text is labeled as topic five (mixed topic). All the collected texts are classified into the five LDA topics using this approach. Table 1.1 lists top keywords and sample articles for each LDA topic group.

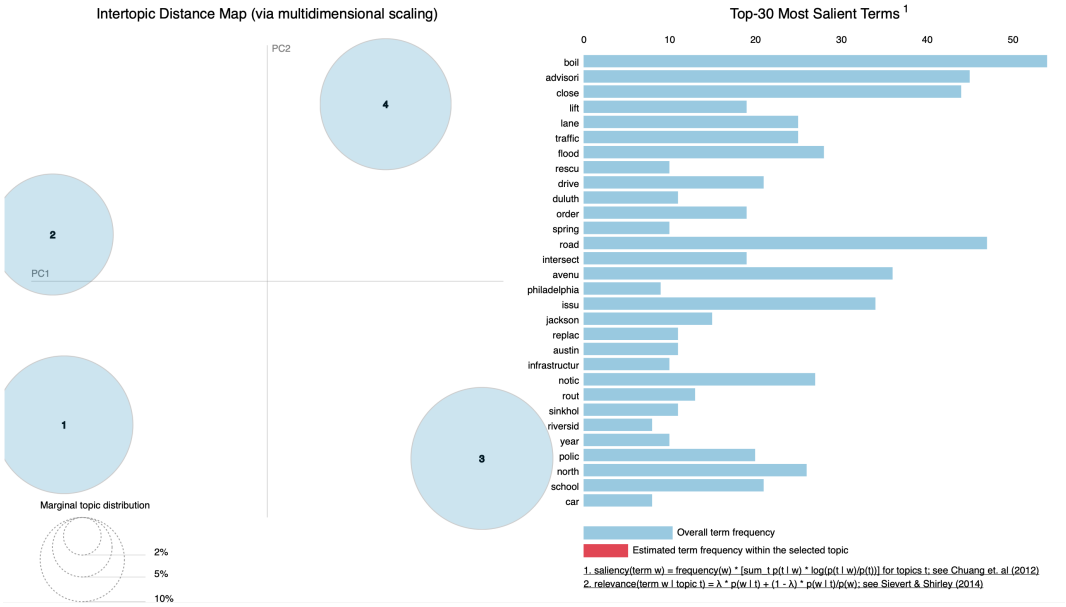
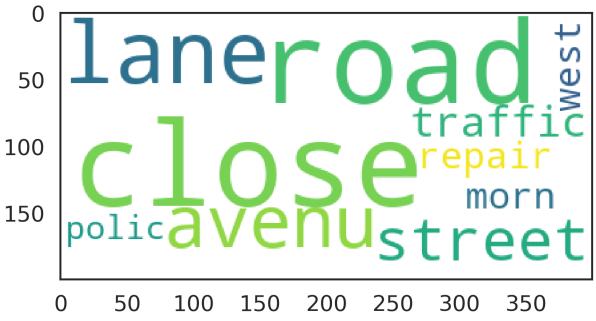
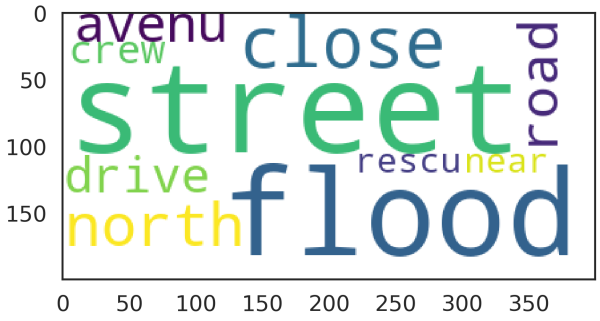


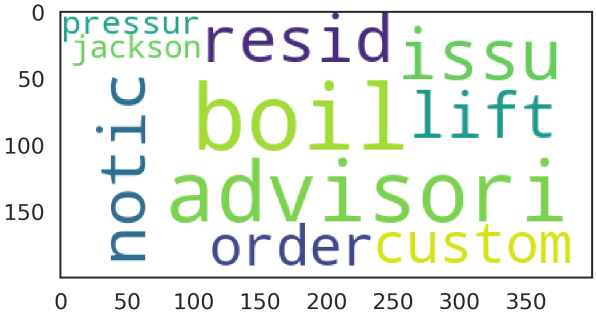
Figure 1.3: Inter-topic map of topics generated by the LDA model



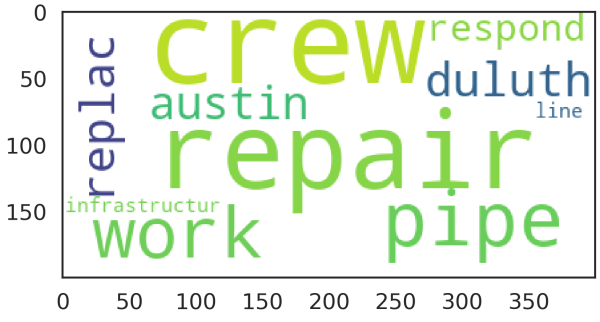
(a) Keywords of LDA topic1



(b) Keywords of LDA topic2



(c) Keywords of LDA topic3



(d) Keywords of LDA topic4

Figure 1.4: Keywords for the extracted LDA topics

Table 1.1: Example articles for the extract LDA topics

Topics	Top Keywords	Example
Topic 1	Road Lane Street Close Police Traffic	Water main break closes Battleground Avenue lanes Part of a busy Greensboro intersection was closed Wednesday morning due to a water main break. Greensboro police said southbound lanes of Battleground Avenue are closed at Green Valley Road because of the break.
Topic 2	Flood Street Road Close Crew Rescue	Massive sinkhole appears in South LA neighborhood after water main break floods streets SOUTH LOS ANGELES (KABC) – A large sinkhole opened up in a South Los Angeles neighborhood after a water main break flooded several streets early Wednesday morning.
Topic 3	Boil Advisory Issue Notice Order Residence	Boil water advisory after break in Kalamazoo The City of Kalamazoo is repairing a water main break on Lake Street that resulted in a temporary loss of pressure.
Topic 4	Pipe Crew Repair Respond Work Replace	Finding a needle in a haystack: Mayor Guajardo address search for a cause of water main break "Moving forward you can bet that this leadership is going to change that as well," said Mayor Guajardo referring to the city upgrading technology.
Topic 5	Mixed topics	"I cannot move it"; Detroit water main break freezes resident's car to the street Travel on Burnette Street in Detroit is a struggle these days as a water main break that spilled water onto the road completely froze the pavement. One woman's car was literally stuck to the ground.

Figure 1.5 shows the topic distribution of the collected data. The first thing to notice is that no topic dominates the collected data. Nevertheless, topic three (health) and one (traffic) are the most common topics in all news. These validate the commonly held assumption that the most severe social damages from water main break events are water quality deterioration (topic 3) and local traffic disturbance (topic 1) [146, 47, 105]. Utility responses on the water main break events, the progress of pipe repair work, and discussions on future replacement strategy (topic 4) are the next common themes for news articles. Among the summarized topics, topic 2, related to property damage due to the break event, has the least amount. The relatively small frequency of this topic might due to the fact that only major break events can cause significant damage to the surrounding properties, which are relatively rare. Lastly, many articles are considered to have a mixture of topics. Considering all the topics are related to water main break events, articles with a mixture of topics are unsurprising.

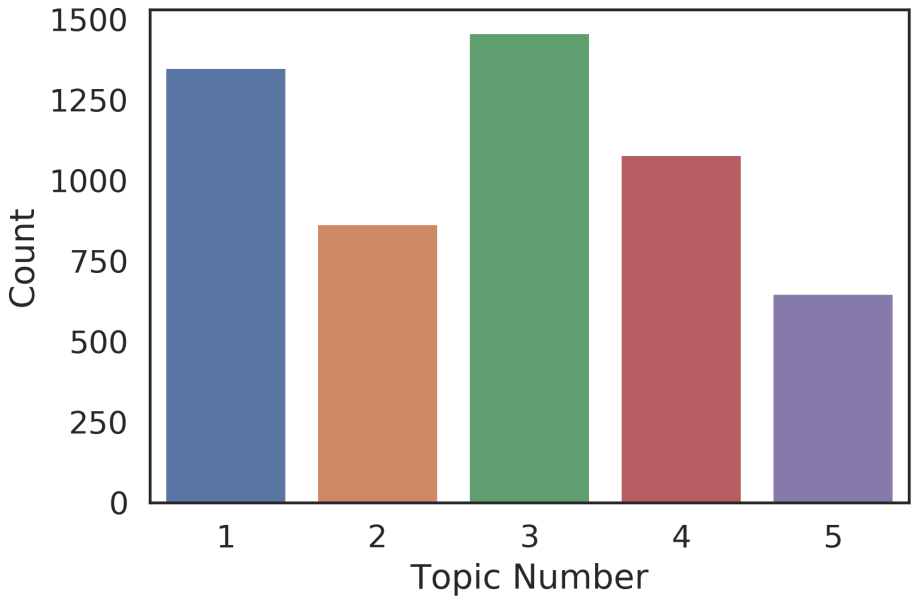


Figure 1.5: Count plot for LDA topics of the collected dataset

Other aspects of water main breaks

Although age-related water main break events cause significant economic and social impacts on our society, their occurrence is relatively sparsely distributed across time and space. In other words, few pipes will break in a short period under normal conditions (for individual WDN), which permits a sufficient response time for the corresponding water utilities to repair. The consequences are devastating when many pipes fail simultaneously inside a

WDN due to external perturbations, such as earthquake events. Since water pipes are buried underground, earthquake-triggered ground movements may damage pipe integrity, resulting in water leakage or even complete water main breaks [122]. Due to limited resources, rapid repair is challenging under such circumstances for water utilities, causing economic and societal consequences to the serving communities. Historical earthquakes events have demonstrated the vulnerability of water systems to ground motions.

The magnitude 7.8 1906 San Francisco earthquake almost destroyed the city's WDN. The devastating earthquake-induced fire was not effectively contained partially due to the lack of water pressure on the damaged WDN [106, 117]. The magnitude 6.9 1989 Loma Prieta earthquake caused at least 761 breaks and leaks to water pipes across the San Francisco Bay area [87]. At least four buildings were destroyed, and seven structures were damaged due to the lack of water pressure for firefighting [116]. The 1994 Northridge earthquake also created great damage to the WDN. It is reported that at least 1087 locations of pipes were repaired during the event [70, 59]. At least twenty-three locations of the main transmission lines were damaged in the magnitude 7.2 1995 Kobe earthquake. 18 out of 21 service reservoirs were closed following the earthquake. It is estimated that indoor and outdoor supply pipe leakage sites will number 50000, out of a total city water supply region of 650,000 homes [27]. Water outages were one of the factors that caused the evacuation of 15 million people of the city [147]. In a more recent case, about 250 pipes were damaged in the City of Napa in the magnitude 6.0 2014 South Napa earthquake, causing water outages to several communities in the region[106].

In addition to water pipes, other WDN components can also fail due to aging or hazard events. One of the critical but often overlooked WDN components that can fail are isolation valves. Isolation valves are indispensable instruments for WDN maintenance activities. They are used to isolate pipe breaks, confine contaminant intrusion, and help unidirectional flushing of distribution systems [37]. Repair of water main breaks always start with the closure of corresponding isolation valves. Thus, conditions of isolation valves influence the range and duration of customer service interruptions [72]. Understanding the impact of having potential malfunctioning valves in a system helps water utilities evaluate their system resilience in a more comprehensive way.

Studies on the valve conditions in the US are limited but reveal worrying facts. American Water Works Association (AWWA) conducted a questionnaire survey of valve management in water utilities [37]. They found that the frequency of encountering missing valves under emergence situations was about 10% per year. The frequency of having malfunctioning valves at the time of usage was about 9% per year. About 81 times per year, utilities found a valve was needed where none exists. Accessibility of the existing valves was also an issue. Nearly 11% valves had moderate accessibility, and 4% had poor accessibility. In a more recent study, Baird reported that 40% of valves would become inoperable if they are not tested and repaired/replaced every five years [11].

1.2 Objectives and Scope

This study focuses on quantifying the hydraulic impacts on failures of water pipes and isolation valves of a WDN. Specifically, it aims to address the following research questions. First, what are the hydraulic consequences of multiple, simultaneous pipe failure events (e.g., under earthquake scenarios) to a large-scale WDN? In addition to pipe failures, will malfunctioning isolation valves cause trouble to the operations of WDNs? If so, how can we quantify the risk while uncertain about the exact status of the valves? What are the key factors that influence the magnitude of the risk? How can we control these factors to a) find the best location to place valves; b) maintain valves with priority?

The next chapter reviews past research on the subject in detail. Critical findings are summarized, and knowledge gaps are delineated, which will be filled by works described in later chapters.

Chapter 3 introduces the design and implementation of an efficient, multi-threaded C++ WDN hydraulic simulator (HydrauSim). In addition to normal operation condition simulation, HydrauSim allows disruptive incidents (e.g., pipe leaks/breaks) to be modeled using the modified pressure-demand-driven (PDD) simulation. It also provides tools to read, configure, and analyze the impact of isolation segments concerning different isolation valve configurations and valve failure scenarios. The developed fast and comprehensive hydraulic simulator is essential for investigating the proposed research questions, which all depend on the hydraulic properties of the WDN.

Chapter 4 utilizes the developed hydraulic simulator to understand the consequences of earthquakes on a large WDN. Impacts of the hypothetical M7.05 Hayward fault rupture event in the San Francisco Bay Area are simulated on East Bay Municipal Utility District's (EBMUD) main gravity feed zone. Instead of focusing on a single scenario, a Monte Carlo (MC) simulation was used to deal with uncertainties across different ground-motion scenarios. The simulation results match the findings from previous state-of-art research, validating the developed simulator's correctness.

Chapter 5 presents a framework to understand the impact of having malfunctioning isolation valves for a WDN. It starts with the mathematical definition of system valve failure risk. Influencing factors for the risk are extracted and explored analytically and through simulation. Lastly, the mechanism of risk change is explained, and the relationship between the system risk and the rate of failed valves is determined. Chapter 6 validates the proposed theory on two real-life WDNs.

Chapter 7 presents an algorithm that automatically determines the best locations to place isolation valves to minimize system valve failure risk. After proving the optimality and feasibility of the algorithm, the algorithm is tested on real-life WDNs. In addition, the proposed method is compared with standard empirical valve placement strategies. The proposed method yields better results and clarifies why some empirical strategies work (or do not work).

Chapter 8 describes an efficient valve maintenance strategy that minimizes valve failure risk under valve condition uncertainties. It guides the water utilities to choose a small portion

of valves in the system to be maintained in priority. Maintenance of the selective valves help reduce system valve failure risk while valve conditions deteriorate. The effectiveness of the proposed maintenance method is validated on real-life WDNs.

The final chapter summarizes the research findings. It presents the conclusions of the research and gives recommendations for future studies.

Chapter 2

Literature Reviews

2.1 WDN hydraulic simulation

Research interests in understanding the hydraulic behavior of water distribution networks (WDNs) through simulation have appeared since computers' invention and wide application. Martin and Peters formulated the WDN hydraulic problem using nodal equations with heads as unknown (the head method) [89]. Their proposed algorithm was the first computer algorithm to solve the hydraulic head at each junction node simultaneously. The core of their approach is to solve the linearized system (through Taylor expansion) iteratively using the Newton Raphson method. Later studies expanded the application range of the head method. Shamir and Howard showed that hydraulics of WDNs with pumps and valves could also be simulated using the head method [119]. Rao et al. extended the method to include the time component (extended period simulation model) [108]. In addition to research interests, the head method also inspired simulation software development. Gessler and Walski created WADISO for the U.S. Army Corps of Engineers [49]. In addition to WDN hydraulic simulation, WADISO also provided tools to help design different components of a WDN. Another WDN head method based hydraulic simulator, AQUA, was created in the same period by Sarikelle and Cesario [115].

Approaches other than the head method were also proposed for WDN hydraulic simulation. Epp and Fowler applied the Newton-Raphson method to solve loop corrective flows using loop-equations based formulation (the loop method) [41]. Compared to the head method, the loop method requires a smaller computation matrix, thus improving the computation efficiency. Later, Wood and Charles proposed another method (the simultaneous pipe method) [142]. The new method simultaneously solves the flow rate in each pipe using the combination of conservation of mass and energy equations for each loop. Wood and Rayes reported that both the loop method and simultaneous pipe method yielded a better convergence rate than the head method [143]. Multiple commercial software was developed based on the two methods, including WATER [24], WOODNET and KYPIPE [141].

Based on the simultaneous pipe method, Todini and Pilati proposed the Global Gradient

Algorithm (the simultaneous network method) [130]. Instead of formulating conservation of mass and energy equations for each loop, the simultaneous network method formulated them on each pipe segment. Therefore, it does not require identifying a fundamental set of loops, which is needed for both the loop and simultaneous pipe method. Using the simultaneous network method, hydraulic heads at each junction node are solved directly along with the flows, simplifying the simulation procedure. Todini and Pilati proved that the simultaneous network method is equivalent to the loop method regarding simulation convergence speed. Specifically, the problem space of the simultaneous network method can be converted to the problem space of the loop method through linear transformations. In more recent work, Todini and Rossman proposed a unified framework for deriving simultaneous equation algorithms for WDN hydraulic simulation [131]. After comparing various algorithms, they concluded that algorithms formulated by the simultaneous network method are the best regarding several computational factors.

The simultaneous network method has advantages over both the head and loop method. Without the need for loop identification, it is as flexible as the head method, allowing easy model expansion to include other system components such as valves and pumps. At the same time, it is computationally efficient, having the same convergence rate as the loop method. Due to these benefits, the United States Environmental Protection Agency (USEPA) adopted the simultaneous network method to the development of a WDN hydraulics (and water quality) simulator, the EPANET [113]. In 2000, USEPA upgraded the software to EPANET 2.0, which included a Programmer's Toolkit, which allowed developers to customize the program to their need [111]. EPANET2 is considered the reference (most popular) software package for WDN hydraulics and water quality simulation [7].

Despite numerous advantages, EPANET2 has its limitations. One of the most significant limitations is the relatively long computation time. Although it can solve regular size WDNs in seconds to minutes, the computation time for extremely large size WDNs is considerable [23]. In addition, many tasks (e.g., network optimizations, long-term effect simulations, rehabilitation planning) require multiple runs of the hydraulic simulation. In such cases, the computation inefficiency of EPANET gets multiplied, resulting in significant waiting time to solve such problems [38, 127].

Many studies have been conducted to improve the computation efficiency of EPANET2. One way of improvement is to utilize advanced computation techniques. Alonso et al. used distributed parallel computing techniques to speed up the matrix factorization part of the EPANET2 [6]. Crous et al. explored the possibility of using graphics processor unit (GPU) for WDN hydraulic simulation [34]. They concluded that the GPU based simulator is faster than the conventional central processing unit (CPU) based simulator on large-scale WDN models. Guidolin et al. investigated the use of high-performance computing (HPC) techniques to accelerate WDN hydraulic solvers [56]. After modifying EPANET2 to be compatible with HPC libraries, they observed a moderate speedup in performance. Alvarez et al. found that hydraulic solver speedup can be achieved on multicore systems using OpenMP [8]. However, Burger et al. argued that using modern multicore capable solvers alone could not lead to significant speedups on EPANET2 [23]. New algorithms have also been proposed

to improve the computation efficiency of the simulator. Alvarez et al. proposed an efficient procedure to select the network loops that result in sparse computation matrix, improving the efficiency of the loop method [7]. Diao et al. utilized the domain decomposition method, the Schur complement domain decomposition, to accelerate WDN hydraulics simulation [39]. The effectiveness of the proposed method was validated on a large-scale real-world system.

2.2 WDN seismic impact modeling

In addition to analyzing the hydraulic behavior of WDNs under normal operational conditions, numerous researches have been conducted to quantify WDN responses during disruptive incidents, especially in the case of seismic events, which can be grouped into two categories. The first group consists of studies that use empirical relationships for WDN damage estimations. The essence of this approach is to estimate damages (with uncertainty) of the hazard incident (e.g., earthquakes) using the pre-established vulnerability (fragility) functions [106]. Developed by Federal Emergency Management Agency (FEMA), HAZUS-MH [42] is a GIS(geographic information system)-based tool that estimates the damage of infrastructure (including WDNs) in terms of social and economic values under multi-hazard scenarios. HAZUS-MH uses the fragility function proposed by O'Rourke and Ayala [101], which describes a relationship between the median rates of repairs of pipeline and seismic wave intensity (peak ground velocity) based on observations of data on four U.S. and two Mexican earthquakes. The Mid-America Earthquake Center developed MAEviz [98] for infrastructure earthquake risk evaluation. MAEViz can generate damage estimates, test mitigation strategies, and support social consequence modeling for various infrastructure systems, including WDNs. Unlike HAZUS-MH, which assumes water main exists under each street with a fixed ratio of material types, MAEViz allows users to specify the location and characteristics of each pipe segment for the study. Based on MAEviz, a platform called Interdependent Networked Community Resilience Modeling Environment (INCORE) [48] was developed as an open-source multi-hazard assessment, response, and planning tool for community resilience planning. In addition to WDN hazard simulation, INCORE provides tools to evaluate the interactions between different infrastructure systems under hazard conditions. Recently, Porter proposed a stochastic simulation model for risk analyses of earthquake damage on WDNs [106]. Using empirical models, it quantifies damages and restoration of WDN over the entire earthquake sequence with account for interactions with different lifeline systems.

A major limitation of empirical model based tools is the lack of precision as that they do not explicitly model the hydraulic behavior of damaged WDNs. Since the release of EPANET, research has been conducted to evaluate WDN seismic reliability through hydraulic analysis. GIRAFFE (Graphical Iterative Response Analysis for Flow Following Earthquakes) [121] employs the iterative approach for the treatment of negative pressures in the damaged water supply system using the EPANET hydraulic solver. The WDN reliability analysis results are linked to GIS (geographic information system) functions to provide clear

visualizations. Based on GIRAFFE, Yoo et al. proposed REVAS.NET (Reliability Evaluation model of Seismic hazard for water supply NETwork) [147], which includes seismic events generation into the simulation process for a comprehensive WDN seismic resilience estimation. One drawback of adopting EPANET hydraulic simulator (such as GIRAFFE) is that it does not model network damage explicitly. Thus advanced hydraulic simulation techniques—such as pressure-driven hydraulic simulation—are not included, limiting the hazard analysis capability of the extended tools. Recently, an open-source python-based WDN resilience software, WNTR [78], has been developed, which is a comprehensive tool developed by EPA and Sandia National Laboratories that aims at WDN resilience analysis. However, the WNTR hydraulic simulator simulation speed is not scalable– an issue when considering the size of WDNs– thus making it unsuitable for resilience analysis of large-scale WDNs or optimization tasks requiring repeated hydraulic simulations.

Earthquake impacts on real-world WDNs have been explored using the developed tools. Ballantyne et al. evaluated earthquake damage to the Seattle water supply network, and proposed a method to estimate system serviceability using hydraulic network analysis [13]. Shinozuka et al. evaluated the seismic performance of the water supply system operated by Memphis Light, Water and Gas (MLWG) in Memphis and Shelby County, TN [125]. Lee et al. conducted seismic reliability analysis on Korea's A industrial water system [80]. They concluded that the earthquake resistance strategies should be based on considering the reliability of the water supply, the demand for water, and the cost of earthquake resistance. Yoo et al. tested their seismic reliability evaluation model (REVAS.NET) on Korea's J City network [147]. They found that performing earthquake impact analysis on skeletonized transmission networks may overestimate seismic reliability than using the entire system.

Due to the high frequency of earthquakes, seismic impacts on WDNs in California have been widely studied. Multiple studies focused on evaluating the seismic reliability of San Francisco's Auxiliary Water Supply System (AWSS), with a general conclusion that the seismic serviceability of the AWSS is very sensitive to pipe breaks [77, 88, 117]. The United States Geological Survey (USGS) recently led a series of research on quantifying engineering and social impacts of a hypothetical magnitude 7.0 earthquake (the HayWired scenario) on the Hayward Fault in the San Fransisco Bay area [64]. The HayWired scenario analyzed the earthquake vulnerability of two WDNs in the region, the San Jose Water Company WDN and East Bay Municipal Utility District (EBMUD) WDN. Based on the analysis results, they proposed several preparation and restoration plans to improve WDN resilience [106]. Another widely studied WDN in California is the Los Angeles Department of Water and Power (LADWP) WDN, a large complex system sitting on top of several active earthquake faults. Shi et al. validated their WDN earthquake impact analysis software, GIRAFFE, using the damage data of the 1994 Northridge earthquake on the LADWP network [121]. Romero et al. investigated the potential seismic hazards on the network from a hypothetical 7.8 MW earthquake (the ShakeOut Scenario) [110]. They found that nearly 2,700 locations of pipeline may be damaged, resulting in a 66% decrease in water service. Tabucchi et al. described a discrete event simulation model of post-earthquake restoration for LADWP WDN [128]. The model was validated using the 1994 Northridge earthquake data, proving

its capacity to estimate the restoration's time and spatial sequence.

2.3 Role of isolation valves in a WDN

As a complex engineering system, a WDN consists of various components. Besides water pipes, previous researchers have emphasized the importance of valves in WDNs. Ysusi summarized the four primary purposes of WDN valves: a) subsystems separation; b) pressure and flow control; c)stagnant water drainage; d) air release and vacuum prevention [92]. Recognizing the significance of WDN valves, Whittaker and Arscott described potential problems in identifying, selecting, operating, monitoring, and recording valves [139]. They proposed applying data logger capsules on valve heads to preserve valve information. Bouchart and Goulter conducted the failure analysis of WDN valves [22]. They concluded that valve malfunctioning, which is inevitable for systems with many valves installed, can degrade the hydraulic performance and water quality of a WDN. Later, Goulter et al. reviewed the reliability analysis on WDN valves. They concluded that valves should be installed at all critical points (such as the end of each city block) to prevent significant failure consequences [92].

Walski first developed a systematical way of modeling the failure consequence of isolation valves in a WDN [137, 138]. He formulated the concept of isolation segment, which is the smallest set of pipes that can be isolated after closing all the surrounding valves. Failures of valves result in increasing of pipes in isolation segments, implying deteriorating WDN conditions [137]. Based on the concept of isolation segment, Walski further suggested representing the WDN using node as segments and edge as valves [138]. In this way, valve failure consequences can be analyzed using network analysis techniques. Inspired by Walski's work, Jun and Loganathan proposed an algorithm automatically identifying isolation segments from a WDN [73]. Their method tracks the unintended isolation, which are subnetworks that lose connection to the water source due to the isolation of other system parts. Later, Giustolisi and Savic proposed a method that utilizes the topological incidence matrices of network hydraulic models to identify segments from a WDN [52]. They argued that the proposed model could be incorporated into the commonly used WDN hydraulics simulator (e.g., EPANET) due to the lack of reliance on network analysis. Recently, Huzsvár et al. analyzed the topological features of the segment-valve graph for some WDNs [66]. They found that the segment-valve graphs of the nine tested WDNs follow the same degree distribution, and structural properties (e.g., clustering coefficient) of these graphs are similar to a random graph.

Graph theory and network analysis have been used in WDNs to achieve various tasks. Torres et al. explored the topological effects on WDN performance using graph theory and statistical models [132]. The study suggested that WDN performance may have percolation-like transitions. Perelman et al. used unsupervised learning algorithms from graph theory to automatically partition a WDN into smaller independent sub-systems [104]. Three different algorithms: global clustering, community structure, and graph partitioning, were tested in the study and demonstrated promising results. Recently, Santonastaso et al. included

actual positions of isolation valves to the design of WDN partitioning algorithm [114]. A generic partitioning algorithm is applied to the dual topology of a WDN, the segment-valve graph, which produces WDN partitioning solution that requires no additional isolation valve installations. Liu et al. applied spectral clustering algorithm and graph theory to facilitate district metered areas (DMAs) design [85]. Their proposed method can identify different DMA solutions effectively under various performance criteria for a real WDN. In addition to clustering methods, modularity-based approaches are also widely adopted in literature for DMAs partitioning [28, 50, 31].

2.4 Valve vulnerability analysis

Various researchers have conducted vulnerability analysis of WDNs considering valve conditions. Hernandez and Ormsbee considered valves' existence when evaluating WDN performance in response to pipe failures [61]. Instead of only removing the failed pipe from the system, the corresponding isolation segment is removed based on the valve locations. The hydraulic impact of the isolation is then quantified using EPANET. Li and Kao developed a method to identify critical segments whose isolation significantly impacts other parts of a WDN using hydraulic simulation [83]. They argued that utilities should pay more attention to the pipes inside the identified segments due to the significant failure consequences. Choi and Kang considered the valve layout on WDN serviceability modeling during seismic events [25]. Incorporating valve positions leads to a more realistic identification of water suspension areas, improving the efficiency of potential restoration strategies.

Some researchers analyzed the consequence of simultaneous failures of multiple components, including valves, in a WDN. Liu et al. used three indicators, the number of valves needed for segment isolation, the size of the segment, and demand shortage, to quantify the impacts of isolation valve failure of a WDN [84]. After analyzing several case studies, they concluded that a higher density of isolation valves leads to less impact of valve failure. Berardi et al. used multi-objective optimization to search for the most disruptive combinations of multiple asset failures for a WDN [15]. They concluded that many failure-induced WDN topological modifications (e.g., detachment of reservoirs) might cause severe WDN service disruptions. Shuang et al. modeled the failure propagation time of a WDN experiencing cascading failure considering network topology, water demand distribution, demand multiplier, and pipe break isolation [126]. They also proposed a method to identify critical pipes to install valves, which can effectively reduce the consequence cascading failures for a system.

2.5 Valve placement strategy

Since valves are essential to a WDN, research has been conducted to find the best way to place valves in a system. Early studies focused on developing rule-of-thumb valve placement strategies. Ysusi suggested that the ideal approach is to install two valves at both ends

of every pipe in the system (the n-rule) [92]. However, if the n-rule can not be met due to budget constraints, installing one less valve than the number of pipes at a junction is also acceptable(the (n-1) rule). In addition, valves should be installed at selected intervals for large diameter transmission pipelines. Ozger and Mays recommended another rule that a pipe isolation should require less than four valve closures [103]. Walski suggested that valves should be installed on connections to large mains to prevent isolation from small lines' failures [136].

In addition to empirical rules, quantitive methods (e.g., simulation, optimization) have also been applied to find the efficient valve configuration for a WDN. Bouchart and Goulter presented a valve placement model to minimize demand volume deficit [22]. They concluded that WDN reliability could be improved by increasing the number of valves in a system. Hwang and Lansey found that randomly locating valves in distribution pipes while increasing the number of transmission isolation valves using the N or N-1 rule is effective at improving system availability [67]. Giustolisi and Savic formulated the valve placement problem as a multi-objective optimization problem [52]. They then demonstrated the feasibility of using genetic algorithms to solve the problem on an Apulian network. Similarly, Creaco et al. presented an optimal valve placement method using a multi-objective genetic algorithm [32]. After testing different objective functions for the algorithm, they concluded that the most appropriate objectives to be minimized are the total cost of the valves and the system-level water demand deficit. Recently, Meng et al. proposed a valve placement framework based on system resilience assessment [93]. Unlike most previous methods, their work considers the likelihood of failures for the installed valves. The proposed method is validated on a real-life WDN.

Rather than reconfiguring the valve placement for a WDN, some studies focus on improving system reliability by adding new valves into an existing system. Gupta et al. developed an iterative procedure to improve WDN reliability [57]. Their method considers the installation of additional valves as one of the ways to improve system resilience during the iteration process. Hernandez et al. developed a graph based iterative heuristic to add isolation valves within an existing WDN to decrease the magnitude of service interruptions in the system [62]. Yang et al. proposed a multi-objective optimization model for adding optimally located isolation valves to existing WDNs, considering both economy and reliability [145].

2.6 Valve maintenance strategy

Various researchers have addressed the importance of maintaining existing valves in a system. Karjalainen stated that valve maintenance and exercising programs are indispensable to ensure the overall reliability of a WDN [75]. The research also found that manual exercising programs tend to be expensive, thus automated monitoring systems are recommended for cost reduction. Regarding the maintenance procedure, Wilson and Garbark reported that a proper valve maintenance program should include locating, global positioning, marking, exercising, correcting problems, recording information, and mapping valves [140]. Shea

reported the Boston Water and Sewer Commission's valve maintenance program, which includes valve exercising, leaks repair, valves replacement, and temporary by-pass piping [120]. Hoff proposed a procedure to select appropriate valves to maintain based on specific purposes of the utility [63].

Some research attempts to incorporate valve information into the WDN management system to facilitate the valve maintenance program. Coate developed a water main shutdown analysis program using the geographic information system (GIS) [30]. Using the location information of both pipes and valves in a WDN, the developed program can locate valves that need to be closed and generate notification reports. The American Water Works Association Research Foundation (AwwaRF) conducted a series of research on providing the criteria for valve location and system reliability [37]. It ended up with developing a strategic valving tool, which helps water utilities balance distribution system investments between the numbers and locations of valves and the overall system reliability. Recently, Morosini et al. proposed a method to manage WDN in pressure-driven analysis (PDA) conditions with isolation valves [96]. It utilized network analysis with objective functions to choose the optimal number of valves to close for system control.

2.7 Summary

Numerous studies have been conducted to simulate the hydraulic behavior of a WDN. Early studies focus on the mathematical formulation of WDN hydraulic model [89, 119, 108, 41, 142, 130]. With a clear understanding of the mathematical properties, later studies aim at building faster, more comprehensive hydraulic simulators [113, 111, 6, 34, 56, 8, 23, 39]. In addition to describe hydraulic behavior of a WDN under normal operation conditions, hydraulic simulation is used to quantify the losses due to hazard events (e.g., earthquakes) [121, 147, 78, 48]. However, hydraulic simulators that can conduct damaged network analysis are relatively slow [23], which makes them unsuitable for computationally intensive tasks. Unfortunately, since hazard events are often uncertain, methods that require many runs (e.g., Monte Carlo simulation) are required to capture the uncertainty [122]. Moreover, large-scale WDN hydraulic analysis is often needed to capture the regional impact of the hazard events. This study fills the research gap by developing an efficient program that can quantify the hydraulic behavior of WDNs before and after a disruptive hazard event such as an earthquake. By synthesizing the past simulator speed-up experiences [23, 8, 56], the simulation speed of the developed program is greatly improved on large scale networks compared to other state-of-art simulators [78, 111]. Details of the developed simulator are described in Chapter 3, with the application demonstrated in Chapter 4.

In addition to pipes, isolation valves can also fail in a WDN. Previous studies found that malfunctioning valves degrade the hydraulic performance of a WDN [22, 84, 15]. However, only a few studies focus on developing methods to understand the failure consequences of isolation valves [137, 138, 52, 66]. This study extends previous works to illustrate how the systematic risk varies as multiple valves fail simultaneously. Factors that control the

magnitude of the risk are identified and verified. See Chapter 5 and 6.

Since isolation valves are important to a WDN, research has been conducted to find the best location to install valves in a system. However, previous studies are either oversimplified, which relies on empirical rules [92, 103, 136, 139], or too complicated, which relies on complex optimization algorithms [52, 32, 93, 145]. This study develops a novel valve placement algorithm using dynamic programming. Compared to optimization-based methods, the proposed algorithm is intuitive and straightforward, yet the solution is both optimal and computationally feasible. Moreover, many widely used empirical valve placement strategies can be successfully interpreted and reproduced using the proposed method. See Chapter 7.

Compared to installing new valves into the system, maintaining the existing valves is more cost-effective. Importance of valve maintenance is stressed by previous research [75, 140] and some experience-based maintenance strategy is reported [120, 63]. Very few research focus on developing quantitative methods to assign maintenance priorities for isolation valves [37, 96, 84]. Chapter 8 fills the knowledge gap by proposing a valve maintenance priority assignment method using network analysis and machine learning techniques. Only the high-rank valves, which help reduce the total system pipe isolation risk while valve conditions deteriorate, are recommended to be maintained under resource constraints.

Chapter 3

Hydraulic Simulation for Water Distribution Networks

3.1 Overview

Water distribution networks are considered part of the critical infrastructure of every modern city by providing clean drinking water for industrial and residential usage. With the growth of population and the physical expansion of cities, WDNs in major cities tend to be large and complex. Considering the overall aging status of the ever-growing water distribution systems across the country [46, 45], concerns over the resilience of the WDNs have risen in recent years [97, 26, 124]. Recent experience from natural and man-made water-related disasters [147] suggests that current water infrastructure systems cannot protect against or prevent all disruptive events and may perform unreliably because of high uncertainty of disturbances, complicated inter-dependency of infrastructure systems, and stochastic failures resulting from unpredictable events [10, 124]. The National Infrastructure Advisory Council (NIAC) states that the effectiveness of a resilient infrastructure or enterprise depends upon its ability to anticipate, absorb, adapt to, and/or rapidly recover from a potentially disruptive event. Based on this criterion, understanding how drinking water systems will perform during disruptive incidents is essential in characterizing the resilience of the WDNs.

This study developed a computer simulation package, HydrauSim, to understand the hydraulic behavior of WDNs before and after a disruptive hazard event such as an earthquake. HydrauSim is an efficient C++ program designed to analyze the robustness of WDNs. Robustness is defined herein as a WDNs capacity to absorb and adapt to a hazard. Similar to EPANET [111] and WNTR [78], HydrauSim is capable of performing the following tasks:

1. Generate water network models from scratch or from existing EPANET formatted water network model input (EPANET INP) files;

2. Modify network structure by adding/removing components and changing component characteristics;

3. Modify network operation by changing initial conditions and component settings; and

4. Simulate network hydraulics (steady-state) using either a demand-driven mode (DD) or a pressure-driven mode (PDD).

Unlike EPANet (v2.0), which targets simulating the normal operating conditions of WDNs, and WNTR, which deals with the resilience of WDNs in general, HydrauSim focuses on assessing the ability of WDNs to resist hazards on a whole system scale by extending the original EPANet .inp file to include disruptive incidents (i.e., pipe leaks) to the network. The status of leakages can be quantified through a modified pressure-driven model (PDM) simulation, as described in Section 3.2. In addition to hydraulic simulation, HydrauSim also includes functionalities to analyze a WDN's robustness from the point of view of isolated segments, which are defined as the smallest component to isolate a pipe from the remaining system [73]. Compared to WNTR (v0.3.0), which provides limited functionalities in regards to isolated segments, HydrauSim provides tools to read, configure, and analyze the impact of isolation segments concerning different isolation valves configurations and valve failure scenarios (see Section 3.2). Moreover, the hydraulic simulator in HydrauSim is highly optimized to simulate extremely large-scale WDNs (see Section 3.4). Hence, it is more suitable for computatively intensive tasks such as large-scale WDN hydraulic analysis or optimization procedures that require many simulations. Table 3.1 summarizes the differences between WDN hydraulic simulators. Note: the current version of HydrauSim does not include extended-time hydraulic simulation nor water-quality simulation. HydrauSim can be downloaded from `https://github.com/cb-cities/pipe-network`.

Table 3.1: Comparison between WDN hydraulic simulators

Features	EPANET (v2.0)	WNTR (v0.3.0)	HydrauSim
Modify Network Structure	YES	YES	YES
Add Disruptive Incidents	NO	YES	YES
Pressure Demand Driven Simulation (PDD)	NO	YES	YES
City-scale Simulation	NO	NO	YES
Isolation Segment Analysis	NO	LIMITED	YES

3.2 Model Description

Network components

HydrauSim models a water distribution system as a collection with nodes connected by links; it follows the same approach as [78] and [111]. The nodes represent junctions and reservoirs. The links represent pipes, pumps, and valves. In addition, to control valves, isolation valves are also included for isolated segments of damage. General descriptions for the system components are listed below. Readers can refer [111] for detailed descriptions.

1. Junctions: Junctions are connections points for links in the network. Water paths split or merge on junctions.

2. Reservoirs: Reservoirs are external sources (with unlimited supply) of water to the network.

3. Pipes: Pipes are the media that conduct water from one junction to another junction in the network.

4. Pumps: Pumps raise the hydraulic head of the water across connecting junctions by injecting external energy into the passing water flow. The head gain from the pump is determined by their pump curves.

5. Valves: Valves act as flow regulators for the network by restricting the pressure or flow at the locations where they are installed.

Four different valve types are modeled in HydrauSim:

1. Pressure Reducing Valve (PRV): PRVs are used to limit the pressure at the locations where they are installed.

2. Flow Control Valve (FCV): FCVs are used to limit the passing flow to a desired amount.

3. Throttle Control Valve (TCV): TCVs are used to simulate partially closed valves.

4. Isolation Valve: Isolation valves are used to isolate damaged sub-networks from the main network to prevent the effects of individual events from spreading throughout the system.

A typical water distribution network is illustrated by Figure 3.1. Water demand is supplied from junctions, which are connected to water sources (reservoir and tank) via pipes. Pumps provide additional energy other than gravity to drive fluid motion, and valves are used to control fluid behavior or to isolate damaged part from the system.

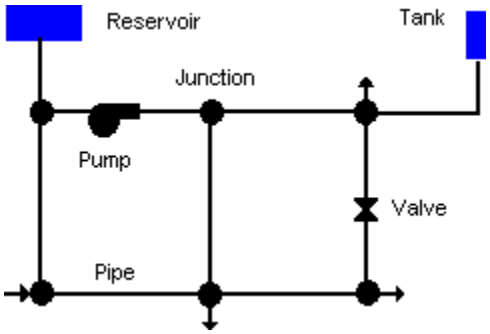


Figure 3.1: Physical components for a water distribution network (from [111]).

Steady-state hydraulic analysis

Basic concepts

Hydraulic simulation of WDNs aims at simulating water distribution status for a given network. Two important quantities are simulated to represent the water distribution status at a given time. One is the flow rate inside pipes, Q, and the other is the hydraulic heads on junctions, H. Although a fast flow rate may damage pipes, a stagnant flow may lead to water-quality problems. Junction pressure is related to hydraulic heads by Equation (3.1). High pressure may break junctions, causing water leakage, whereas low pressure may produce insufficient water supply to end-users.

$$P = H - E \tag{3.1}$$

where P is the pressure, H is the hydraulic head, and E is the elevation for junctions of a WDN.

A large number of different flow conditions may occur in water pipeline systems, such as quasi-steady flow, surge, and water hammer [92]. This study only considers the steady-flow condition, where flow conditions inside the network do not change dramatically over a short period of time. This is because under normal conditions with a relatively large simulation time step, flows inside the network can be approximated as steady state [92]; the steady-state assumption is held by both [78] and [111] as well. Two sets of governing equations are used for the steady-state flow simulation of WDNs [130]. The first governing equation raises from the law of conservation of energy across each pipe k :

$$H_i - H_j - \Phi(Q_k) = 0 \tag{3.2}$$

where i and j represent the nodes at the extremes of pipe k, and $\Phi(Q_k)$ describes the head loss due to friction as a function of flow.

The other governing equation arises from the law of conservation of mass on nodes, which states that total inflow must equal to total outflow at the junctions in the network.

$$\sum_{k=1}^{n_i} Q_{k_{i,j}} + q_i = 0 \tag{3.3}$$

where $Q_{k_{i,j}}$ is the flow in the pipe $k_{i,j}$ connected to junction i; n_i is the number of connected links on junction i; and q_i represents the demand or withdrawal rate on the junction i.

There are many formulas to describe the energy loss (i.e., head loss) across pipes, $\Phi(Q_k)$, in the energy conservation equation [Equation (3.2)]. Following [78], the Hazen-Williams headloss equation is adopted in this work due to its simplicity and universality. The Hazen-Williams headloss formula is defined as:

$$h_k = 10.667 C^{-1.852} d^{-4.871} L q^{1.852} \tag{3.4}$$

where h_L is the energy loss (head loss) in the pipe k; C is the Hazen-Williams roughness coefficient; d represent the diameter of pipe k; L is the length of the pipe, and q is the flow rate of water in the pipe.

In summary, reservoirs/tanks in the network act as sources, which provide water to the end-users (sinks), who withdraw water to meet their demand from the end-service nodes. As water flows from the water sources to the end-users, energy is lost due to the friction of the pipe. To accommodate such energy loss (head loss), external energy sources (reservoirs and/or pumps) must be included to inject external energy into the system. Since system-component conditions, such as reservoir head and pump curves are known and fixed for a given WDN, the behavior of water distribution across the network will be determined by user demand (the sink rate) using the following equations: Equations (3.2), (3.3), and (3.4), which will be discussed below.

Demand driven simulation

In a demand-driven simulation, the distribution state of water, node pressure H, and link flow Q depends on the node demands, which are assumed to be known and must be met for the simulation to run properly. Such an assumption holds under normal operating conditions.

To solve the unknown H and Q, we combine and rewrite Equation (3.3) and Equation (3.2), which gives the following matrix notation, known as the global algorithm representation:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} \begin{bmatrix} Q \\ H \end{bmatrix} = \begin{bmatrix} -A_{10} h_0 \\ -q \end{bmatrix} \tag{3.5}$$

where h_0 is the known nodal heads (from reservoirs); q is the known nodal demands (user demand); $A_{1,0}$ is a matrix relating the pipes to the fixed head nodes; $A_{1,1}$ is a diagonal matrix relating flow to headloss; $A_{1,2}$ is a matrix relating the pipes to the unknown head nodes; and $A_{2,1}$ is a matrix relating flows to known demand. See [130] for detailed explanations.

Equation (3.5) is not linear because of the non-linearity of the headloss equation of Equation (3.4), which may be addressed using the Newton-Ralphson (NR) method:

$$J \cdot dx = R \tag{3.6}$$

where J is the Jacobian matrix; dx is the NR variable updating vector; and R is the residual vector. At each NR iteration, J and R need to be computed from the current state of unknown variables (H and Q) using the governing equations. The variable vector is updated by subtracting dx, which is updated by solving the linear system of Equation (3.6). This iterative process ends when the system residual reaches a small number ($< 1e-8$ as default). Note that unlike the EPANet's implementation [111], which decomposes Equation (3.5) to solve H first and then updates Q, the same procedure in WNTR [78] is followed to solve the linear system in Equation (3.6) directly. Such simple implementation makes it easy to extend model functionalities when considering more governing equations, as will be shown below.

Pressure depends demand simulation

Because WDNs may be damaged by natural hazards (earthquakes, etc.) or man-made hazards (power outages, etc.), in such cases, some of the nodes may experience low-pressure conditions due to pipe leaks, pump failures, or other broken components. Customers with low-pressure supply nodes do not always receive their requested demand. Hence, the basic assumption for the demand-driven simulation is no longer suitable, and the pressure depends demand (PDD) simulation needs to be used instead.

In the PDD simulation, the demand pressure relation is formulated with the following equation [134]:

$$d = \begin{cases} 0 & p \leq P_0 \\ D_f(\frac{p-P_0}{P_f-P_0})^{0.5} & P_0 \leq p \leq P_f \\ D^f & p \leq P_f \end{cases} \tag{3.7}$$

where d is the actual demand; D_f is the desired demand; p is the pressure; P_f is the nominal pressure, which is the pressure threshold that the consumer will receive the requested demand if it is met; and P_0 is the minimum pressure, which is the pressure threshold that the consumer will not receive any water if it cannot be met. Equation (3.5) can be easily extended to account for Equation (3.7). In addition to node head H and link flow Q, actual demand d will be included as unknown variables as well, and the pressure demand relation will be encoded by extending the A part [leftmost part in Equation (3.5)]. Note that nodal pressure is related to nodal head by Equation (3.1). The extended system equations can be directly solved using NR without further mathematical derivations, as described earlier.

Leak model

One of the most common consequences of a damaged WDN is water leakage. For modeling purposes, pipe leaks are aggregated to their closest nodes. The leak is modeled using a general form of the equation by [35]:

$$d_{leak} = C_d A p^{\alpha} \sqrt{\frac{2}{\rho}} \tag{3.8}$$

where d_{leak} is the leak discharge; C_d is the discharge coefficient; A is the area of the broken hole; p is the gauge pressure inside the pipe; α is the discharge coefficient; and the ρ is the density of the fluid. Similar to the PDD model in the previous section, Equation (3.5) can be extended to account for the leak model. The NR algorithm is still effective in solving the extended system of equations, which include the unknown leak discharges in addition to the nodal head, nodal demand, and link flow as part of the variables. Information on leaks can be configured through the extended .inp file; see Appendix B.1.

Isolation segments

When WDNs are damaged, the damaged sections need to be isolated from the main network by closing the corresponding isolation valves to prevent the effects of individual events from spreading throughout the system [73]. For instance, damaged pipes need to be isolated from the system by closing at least two isolation valves before being repaired. Note, WDNs may have missing or inoperable valves in real-world scenarios due to human/environmental factors such as equipment aging. As a result, isolating the desired component may include other components of the system due to malfunctioning or missing isolation valves. For instance, consider the situation in Figure 3.2. If $P1$ is damaged and $V2$ is missing or inoperable, then isolating $P1$ will result in the isolation of $N2$ and $P2$. Moreover, when a segment is closed, there may be other parts of the network that become disconnected from the sources, creating unintended isolation. Therefore, understanding the properties of isolation segments concerning systems configurations and the consequences of missing/malfunctioning isolation valves is indispensable for assessing the resilience or robustness of WDNs. To achieve this goal, HydrauSim provides a "ValveGraph" module to read, model, and analyze the impact of isolation valves and their corresponding isolation segments on the system. Isolation-valve configurations can be imported through a modified version of the .inp file; see Appendix B.2. Efficient valve-segment analysis algorithms are implemented, which allows the user to: (a) find isolation segments inside the system; (b) evaluate the impact of valve failures; and (c) find any unintended isolation due to closure of desired segments.

HydrauSim includes an automatic isolation segment-finding algorithm based on the algorithm proposed by [73]. By replacing the recursion structure with lookup tables, the modified structure is more computationally efficient compared to the original algorithm. The segment-finding algorithm detects the corresponding isolation segment information for

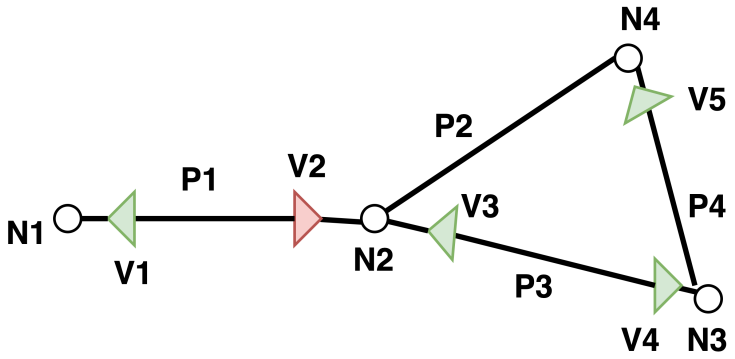


Figure 3.2: System configuration for isolation segment identification.

a given pipe, including isolated pipes, isolated nodes, and isolation valves that need to be closed. The basic structure of the algorithm as follows:

1. For a pipe that needs to be isolated, the corresponding isolated nodes will be found first; they are pipe end nodes that do not have a valve installed on the searching pipe.
2. For a found isolated node i, its corresponding valve lacking pipes is found, which are pipes that do not have valves installed close to the node i.
3. Perform step 1 for the newly found valve-lacking pipes.
4. Stop when there is no pipe and node to explore.
5. Add all the valves that are related to isolated pipes and nodes as the valves that need to be closed for this segment. Program ends.

The above algorithm is illustrated by considering the system configuration in Figure 3.2 as an example. Assume pipe $P1$ need to be isolated, but valve $V2$ is broken: the segment finding algorithm works as follows:

1. Searches for isolated nodes of pipe $P1$, which is node $N2$.
2. Searches for valve-deficient pipes from node $N2$, which is pipe $P2$ ($P1$ has been explored).
3. Searches for isolated nodes of pipe $P2$, which is $N4$ ($N2$ has been explored).
4. Searches for valve-deficient pipes from node $N4$, of which there are none ($P4$ has $V5$).
5. If there are no new pipes or nodes to explore, add all related valves, which are $V1$, $V3$, $V5$. Isolated pipe $P1$ will result the isolation of $P1$, $P2$, and $N2$,$N4$ by closing $V1$, $V3$, $V5$.

The isolation-finding algorithm can be performed for all pipes to determine all potential isolation segments from the system layout. The found isolation segmentation and their corresponding valves can be used to construct a dual graph representation of the WDN, which represents graph nodes with isolation segments and links to those valves that should be closed (see [73]). This segment-valve representation approach is adopted by HydrauSim to analyze the impact of component isolation and valve failures on the system. To isolate a malfunctioning pipe, one needs to find the corresponding isolation segments and close all the corresponding valves by removing its incident edges. Note, removal of one isolation segment may cause unintentional isolation for the system.

Consider the isolation of pipe $P2$ in Figure 3.1. The corresponding segment-valve representation is shown by Figure 3.3(a). Since $P2$ corresponds to segment $S2$, isolating $P2$ means the removal of edges that link to $S2$, which are $V2$, $V3$, and $V5$ in the graph. Removal of these links results in three mutually disjoint components in the graph; see Figure 3.3c). If we assume water comes from $N1$, then isolating pipe $P2$ will unintentionally cause the isolation of pipe $P3$ and $P4$ as there is no available connection to the water source for these two pipes. Representing the system with a segment-valve graph makes it easy to analyze the impact of valve failures. Failure of valves simply means the removal of such valves (edges) from the segment-valve graph and merging of the linked segments. For example, failure of valve v2 will result in the removal of edge v2 from the Figure 3.3(a) and the combination of segments $S1$ and $S2$. The isolation valves module of HydrauSim can automatically detect both intended and unintended isolation segments.

3.3 Program Design Architecture

The architecture of HydrauSim is presented below for users to better understand, utilize, and potentially contribute to the program. The essential design principle for HydrauSim is the decoupling different types of functionalities as much as possible. In this way, the program can be relatively easy to extend with new features without much impact on the existing system. Figure 3.4 illustrates the simplified program design architecture of the program. The system is composed of six major parts:

1. Model Settings: parameter settings for the system, such as the initial value of flow rate of the hydraulic simulation, normal pressure for sufficient water supply on demand nodes, etc.

2. Nodes: nodal objects in the network, including junctions, tanks, and reservoirs.

3. Links: link objects in the network, including different types of pipes, valves, and pumps.

4. Mesh: an object that represents the network, which contains all the nodal objects, link objects, and network properties, such as links/nodes that are disconnected from water sources. Isolating segments resulting from closing isolation valves are also included in this category.

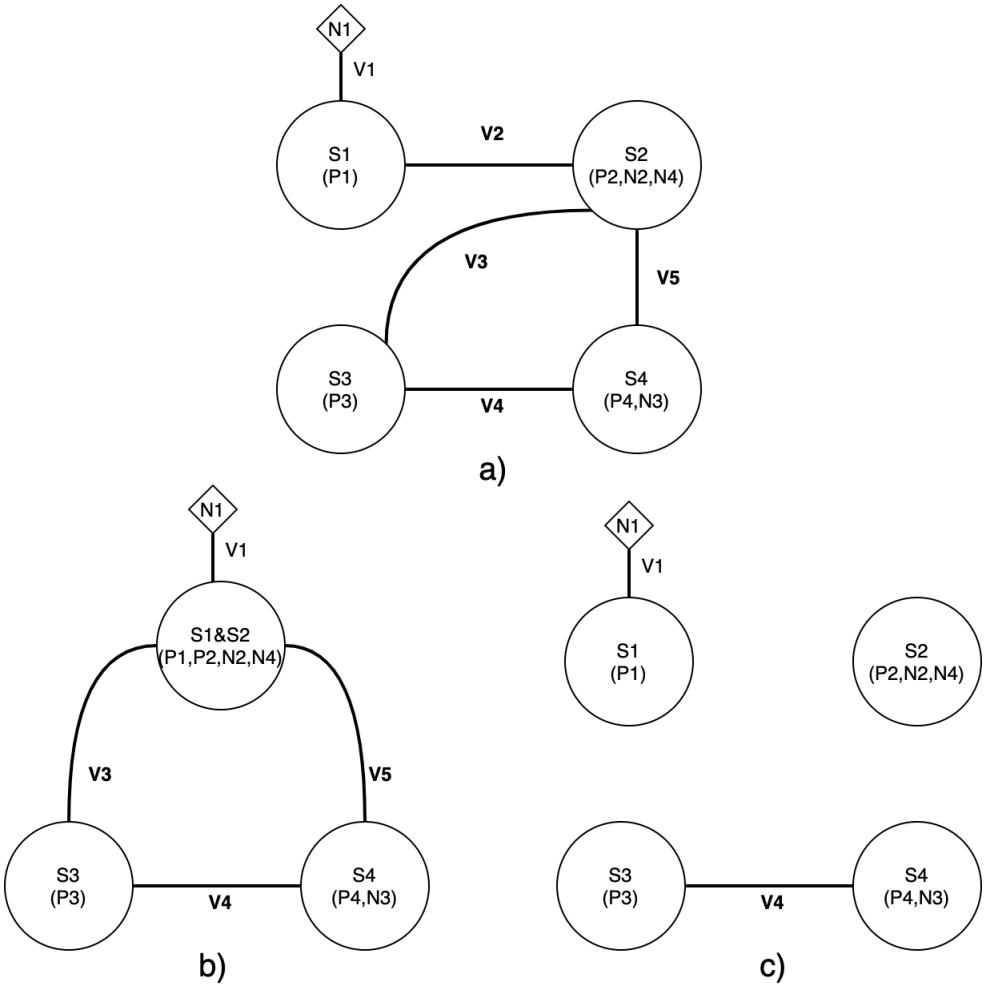


Figure 3.3: Segment-valve representation for a WDN shown by Figure 3.2: (a) all the isolation valves function properly; (b) valve $V2$ fails; and (c) unintended isolation due to the isolation of pipe $P2$.

5. Matrix Assembler: an object that contains methods to compute, update, and store required matrices (Jacobian matrix, variables, and residuals) for hydraulic simulation from the network (the Mesh).

6. Simulator: an object that performs required hydraulic simulation (DD or PDD) for the network using matrices assembled by the Matrix Assembler. The linear system from NR method is solved either by a sequential solver (Eigen LU Solver) for a small network or a parallel solver (MKL Pardiso Solver) for a large network. Details on solver choices will be discussed in the next section.

7. IO: an object that contains protocols for reading .inp files and writing the simulation result back to a file.

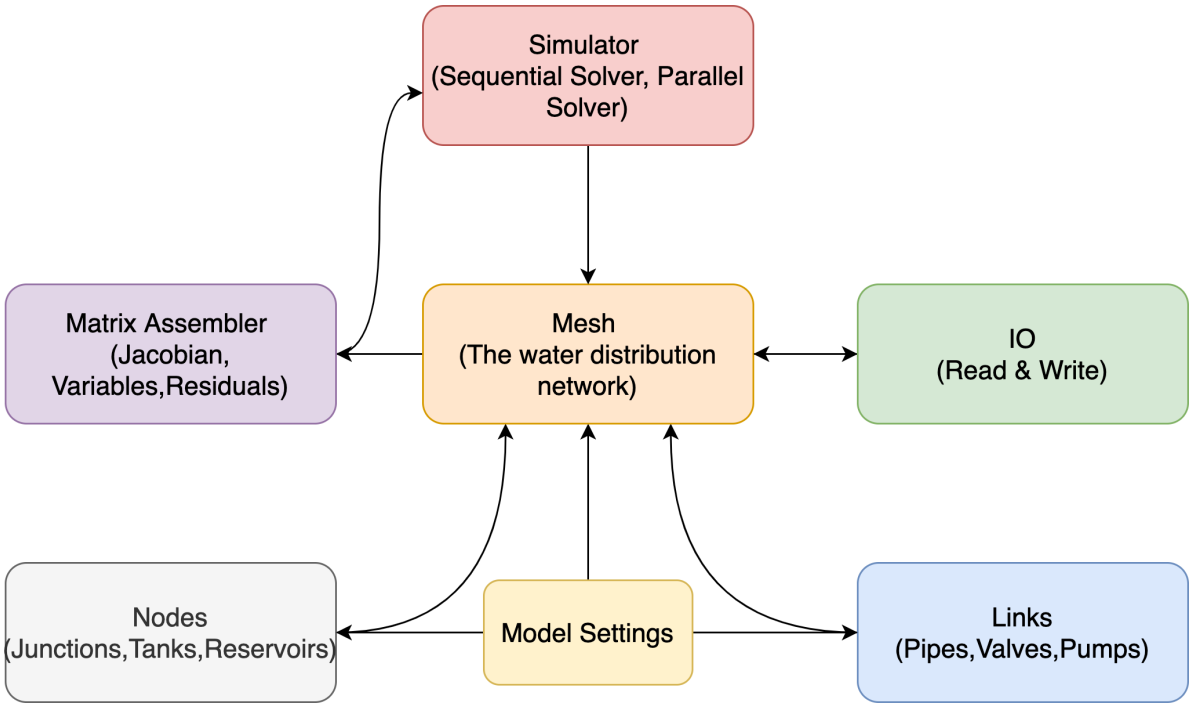


Figure 3.4: Software design architecture, simplified.

A typical workflow for the system is as follows: The IO module reads the .inp file from the user. The parsed information is then used to create the Mesh, which contains Nodes and Links. Network analysis—such as finding isolation segments—can be performed by the Mesh. To perform the steady-state hydraulic simulation, network information contained in the Mesh is assembled by the Matrix Assembler for use by the Simulator. The Simulator performs the desired simulation (e.g., DD/PD and leak quantification) and updates the Mesh with simulation results. Finally, the IO object can be called to export the analysis result back to the user.

3.4 Program Profiling

One key feature of HydrauSim is its ability to perform steady-state hydraulic simulation for large-scale water distribution networks. Specifically, the simulation time for HydrauSim scales linearly to the network size, whereas both WNTR and EPANet exhibit nonlinear behavior for large networks. As shown in Figure 3.5, the simulation times for both WNTR and EPANet rise sharply for large-scale networks ($n > 10000$) compared to HydroSim. To achieve such computation efficiency, both SIMD (single instruction, multiple data) and MIMD (multiple instructions, multiple data) parallelism concepts were incorporated into the design of the hydraulic simulator. For SIMD, all simulation variables were vectorized using C++ package Eigen3 [55] with Intel AVX (advanced vector extensions) enabled for Intel processors. In this way, the benefit of vector-focused modern computer architecture can be fully utilized for variable updating and residual computation during the simulation.

In addition, an adaptive approach was adopted to solve the linear system; see (Equation 3.6). Specifically, the Pardiso [36] unsymmetrical sparse linear system solver was used for large networks. Currently, one of the fastest modern linear system solvers, Pardiso parallels its computation over available cores of the host machine through MIMD. That said, multi-threading does have drawbacks. One major drawback is that multi-threading often requires a fixed amount of time as "threading overhead", which may slow down the whole system when the actual computation time is small. Therefore, for small networks, the highly-efficient, sequential SuperLU solver from the Eigen3 package was used to solve the sparse linear system. The default network size setting for solver switching is 10,000 nodes, but this can be changed by the user for other considerations.

Based on the design configuration above, HydrauSim was constructed and profiled; see Figures 3.5 and 3.6. The testing system used was Ubuntu 18.043, Python 3.7; C++14, GCC 7.4 with Intel i7-9800X CPU, 64 Gb memory. WNTR V0.2.2 and EANETTools v1.0.0 were used for benchmark purposes. Note that profiling was performed using synthetic networks to ensure fixed network properties among networks with different sizes. In other words, only the network size—not network features inside the network—was the variable for the profiling procedure. The details of how to generate realistic synthetic networks is included in Appendix B.3.

As shown in Figure 3.5, there was no significant difference regarding simulation time

for all simulators on small networks; however, for those networks that exceeded 10k, the advantage gained by using HydrauSim becomes obvious. For a network with 100k nodes, the simulation time for HydrauSim is around 5 sec; it is around 20 sec for EPANET and over 30 sec for WNTR. Profiling of the program also shows promising results. Figure 3.6(b) shows the majority of HydrauSim's computation time (94%) is due to the linearSystem solver part, implying efficient or even optimal computation is performed by other parts of system; the linear system solving time is rather fixed as advanced external library was used. Any endeavor to improve the performance is incumbent on the other part of the system.

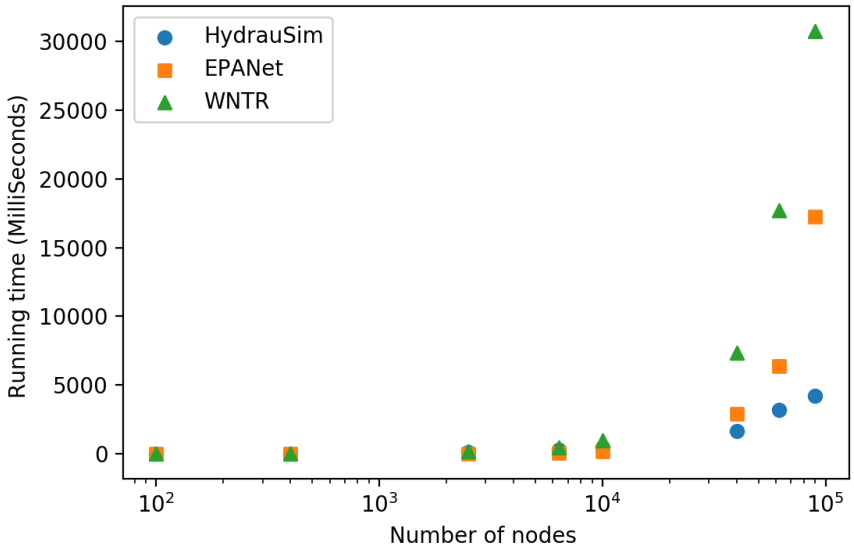


Figure 3.5: Hydraulic simulation speed for different WDN analysis software.

3.5 Summary

An efficient, multi-threaded C++ WDN hydraulic simulator (HydrauSim) was developed to quantify the hydraulic behavior of water distribution networks (WDNs) after a disruptive hazard event such as an earthquake. HydrauSim includes functions to include disruptive incidents (pipe leaks/breaks) in the WDN hydraulic analysis. A modified pressure-driven model (PDD) simulation was used to quantify the pipe-failure-induced water leakages. Moreover, HydrauSim includes tools to understand the relationship between isolation valve conditions and WDN system risks: specifics include the ability of reading, configuring, and analyzing the impact of isolation valves on the network. Most importantly, HydrauSim is capable of performing efficient simulations on extremely large-sized WDNs. It is considerably faster than existing WDN hydraulic simulators on networks with over 10,000 nodes and edges.

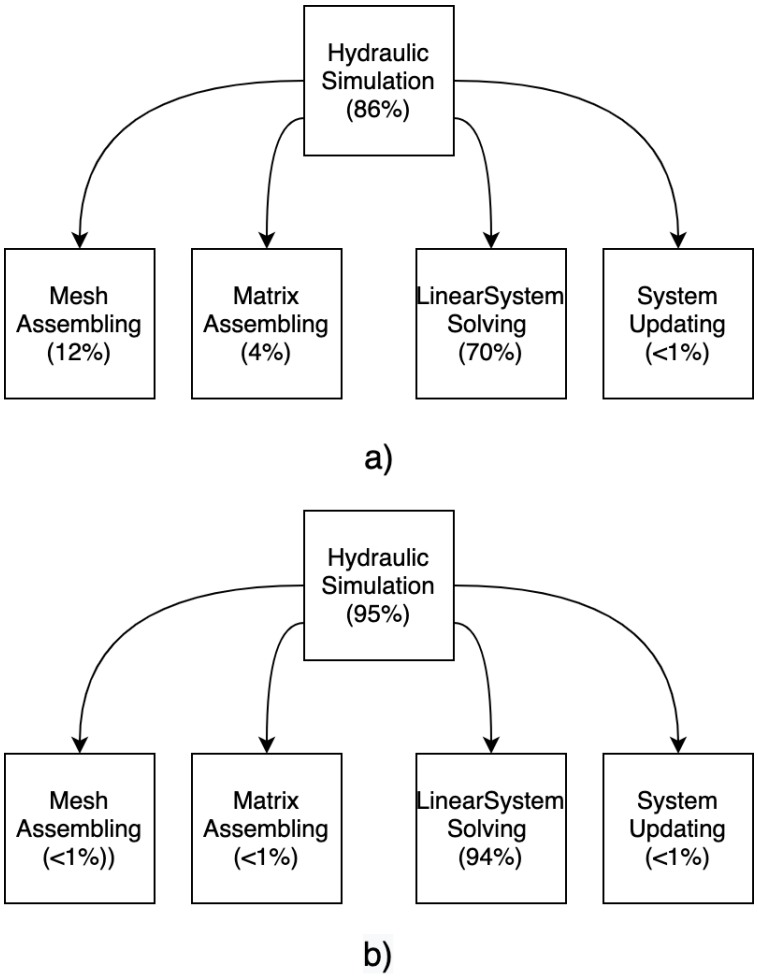


Figure 3.6: Profiling of hydraulic simulation systems: (a) EPANet and (b) HydrauSim.

Chapter 4

Earthquake Impacts on a EBMUD WDN

4.1 Overview

Past earthquakes, such as the Northridge earthquake (1994) [59, 70] and the Kobe earthquake (1995) [27], caused significant damage to WDNs [147]. Clearly, understanding and quantifying the potential impact of a major earthquake on regional WDNs is critical for city hazard resilience. In collaboration with the East Bay Municipal Utility District (EBMUD), this study assessed the potential earthquake damage of the WDN in the east San Francisco Bay Area using the newly developed hydraulic simulator, HydrauSim. Specifically, the earthquake impact to EBMUD's Central Pressure Zone—a major distribution zone—was considered. This zone covers many major cities in the East Bay (Richmond, Berkeley, Oakland, Alameda Island, San Leandro, and part of Hayward); see Figure 4.7.

A probabilistic seismic hazard analysis (PSHA) was adopted herein as the general approach to simulate earthquake risk for the study area, the San Francisco Bay Area. The considered earthquake events are similar to the USGS HayWired case study [3]. Specifically, multiple M7.05, Hayward-Rodgers Creek HN+HS earthquake rupture events realizations from UCERF2 [144] were simulated using the ground-motion prediction equations (GMPEs) from [19]. Since the study focused on earthquake risk assessment of spatially distributed infrastructure system, spatial correlation of intensity measures (IMs) (PGA and PGV) between different sites were modeled using the method developed by [69].

The potential earthquake impact on a WDN was modeled as follows: (1) using the simulated earthquake ground-motion IMs (PGV), the probability of failure for each network component was estimated using the pipeline fragility curves; (2) based on the estimated failure probability, certain components were sampled as the failed component with corresponding damaging degrees; and (3) a pressure-dependent (PDD) hydraulic simulation was performed on the damaged network to estimate the water supply shortage across the selected study region. A Monte Carlo (MC) simulation was used to deal with un-

certainties from different ground-motion scenarios. The newly developed WDN hydraulics program, HydrauSim, was used to perform hydraulic simulations on damaged networks. Code for the earthquake impacts analysis of the case study networks can be found in `https://github.com/rewu1993/EQ-WDN-Analysis`

4.2 Earthquake Ground-Motion Intensity Generation Model

The first step in an earthquake hazard analysis is to acquire the earthquake-induced ground-motion intensity data for the study area. Since this study was aimed at hypothetical earthquake hazard scenario testing, simulated ground-motion intensities were used instead of real data. This research adopted PSHA as the general approach on simulating earthquake risk for the San Francisco Bay Area to quantify uncertainties of earthquake events and combine them to produce an explicit description of the distribution of future shaking that may occur at a site. Hence, PSHA provides more comprehensive risk assessments than a simple deterministic approach [12]. The principle equation for PSHA is presented by Equation 4.1.

$$P(IM) = \sum_{i=1}^{n_{sources}} P(M_i > m_{min}) \sum_{j=i}^{n_M} \sum_{k=1}^{n_g} P(IM|m_j, r_k) P(m_j) P(r_k) \tag{4.1}$$

where $P(IM)$ is the probability distribution of an earthquake IM; $P(M_i > m_{min})$ is the probability of a given earthquake source capable of producing rupture magnitude larger than a threshold (m_{min}); $P(IM|m_j, r_k)$ is the probability distribution of an IM value given the rupture magnitude and the distance from the epicenter; and $P(m_j)$ and $P(r_k)$ represent the probability distribution for rupture magnitude and distance for an earthquake, respectively.

In Equation (4.1), the probability distribution of an IM for a site or location, $P(IM)$ is calculated by summing up uncertainties on earthquake source ($\sum_{i=1}^{n_{sources}}$); uncertainties on the magnitudes of an earthquake ($\sum_{j=1}^{n_M}$); and uncertainties on the locations of an earthquake ($\sum_{k=1}^{n_g}$). This study selected a set of specific earthquake rupture events with a given magnitude that followed assumptions similar to the U.S. Geological Survey (USGS) Haywired case study [65]: i.e., the uncertainties on earthquake sources and earthquake magnitude were not considered in this case and only the distance uncertainty remains. Therefore, Equation 4.1 is reduced to Equation 4.2.

$$P(IM) = \sum_{k=1}^{n_g} P(IM|r_k) P(r_k) \tag{4.2}$$

The region of interest covers most of the major populated areas in the San Francisco Bay Area; see Figure 4.1. Specifically, a bounding box area (-123.52544, -121.21856, 36.904645, 38.8581786) was generated that contained the San Francisco Bay Area, with a 2 km by 2 km grid of simulation points following [3]. Grid points that fall into the sea or non-populated

areas were removed for computational efficiency. In total, 5709 grid points remained for ground-motion simulations. Although this case study shares many similarities to the Hay-Wired case (study area, rupture fault, rupture magnitude, etc.), there are several ways in which they are different:

1. The rupture configurations are different. The HayWired case study considered an M7.05 earthquake scenario (HS+HN G04 HypoO) from a suite of 39 scenario earthquakes for the Hayward fault developed by [1]. The study case herein used an M7.05, Hayward-Rodgers Creek HN+HS earthquake rupture event from UCERF2 [144].

2. The current case does not fix the epicenter of the earthquake, whereas the HayWired case study has its epicenter in Oakland. Having variances on the earthquake epicenter location allows exploring infrastructure systems responses under different spatially distributed ground motions.

3. The ground-motion simulation models are fundamentally different. The HayWired case uses a three-dimensional (3D) numerical simulation to solve the wave equation using the USGS Bay Area Seismic Velocity Model 08.3.0 [2] for the properties in the 3D volume. Herein, the PSHA computation framework, OpenSHA [44] is used, with spatial correlation considerations for earthquake IMs generation. Although PSHA lacks simulation precision compared to the 3D numerical modeling, it requires fewer computation resources and input information, e.g., slip distribution, hypocenter, and the geological model of the study area. This makes feasible the exploration of many earthquake scenarios with different epicenter locations.

OpenSHA is an open-source framework that provides a flexible community-modeling environment for the development and testing of a seismic hazard analysis (SHA) algorithm [44]. The fundamental framework for OpenSHA is shown by Figure 4.2. It requires the selection of earthquake IMs, study sites list, the IM relationship (i.e., GMPEs), and an earthquake-rupture forecast model. The earthquake IMs used in this study were PGA and PGV. As mentioned previously, 5709 grid points were simulated for this study; see Figure 4.1. The earthquake-rupture forecast model, UCERF2 [144], was used to generate the earthquake rupture events (M7.05, Hayward-Rodgers Creek HN+HS). The GMPEs from [19] were used for the IM relationship. The time-averaged shear-wave velocity from the surface to 30 m (V_{s30}) was assumed to be constant as 760 m/sec.

Since this study focused on earthquake risk assessment of spatially distributed infrastructure systems, spatial correlation of PGA and PGV between different sites needed to be explicitly modeled [82]. Hence, the simulated site-specific IMs provided by OpenSHA were corrected with spatial correlations using the method from [69]. Specifically, semi-variogram was used as the geostatistical measure to model the ground-motion correlation with the simulated medians and standard deviations of IMs from the ground-motion model (OpenSHA in this case) as input. Modeling spatial correlations between the IMs also mitigated the

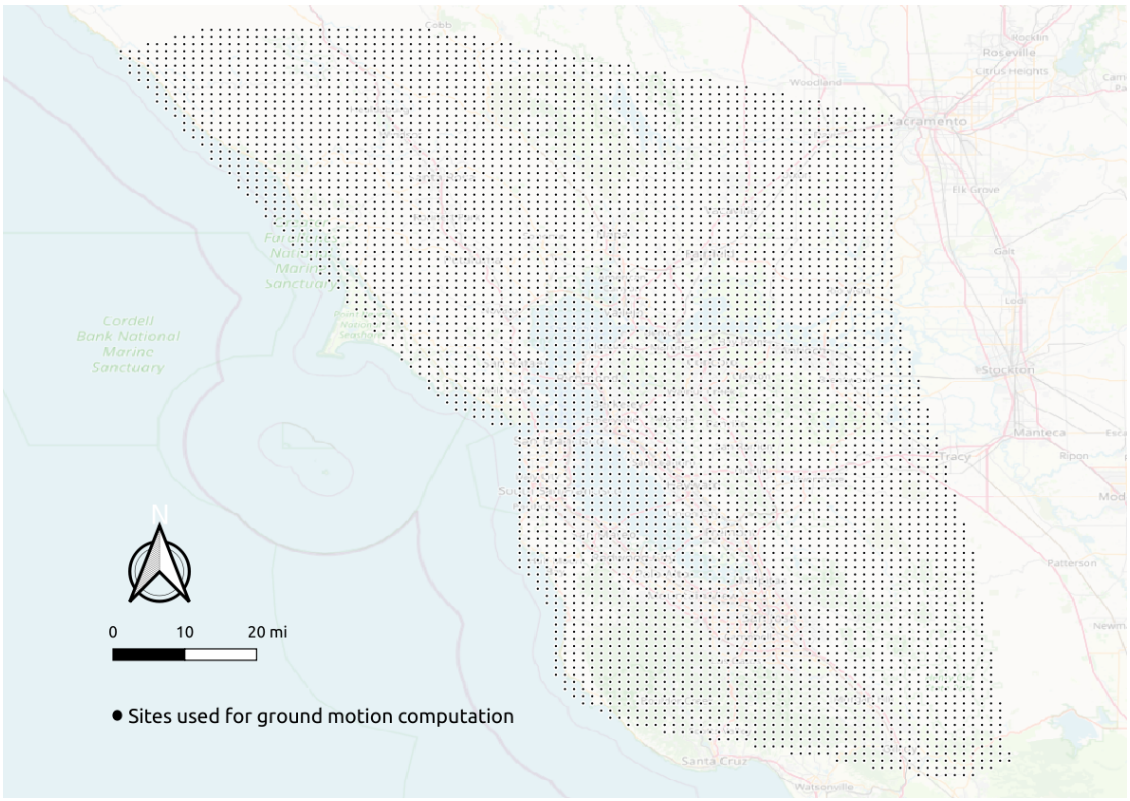


Figure 4.1: Study area for the ground-motion simulation. Black points indicate grid points (sites) for the earthquake ground motion simulation model.

"median map problem" (lacking variances on IMs) of PSHA, which has been criticized in the HayWired case study [65].

Using the method described above, 100 ground-motion scenarios of the M7.05, Hayward-Rodgers Creek HN+HS earthquake rupture events were generated for the study region. Each scenario included both PGA and PGV estimations for all simulated sites with spatial correlation considerations. Figure 4.3 shows the distribution of the average value (across all simulated sites for each scenario) of PGA/PGV for the 100 scenarios. The range for the simulated mean PGV of the study area is from 9 to 22 cm/sec, with a mean value around 13 cm/sec. The range for simulated mean PGA is from 0.05 to 0.28g, with a mean value around 0.14g. According to the Modified Mercalli Intensity table [135], this means that most of the study area will experience degree VI or higher earthquake intensity level for most scenarios, signaling the potential for moderate to high damage to occur. To validate the simulation result, the simulation realizations were compared with the ground-motion map from the HayWired study. When simulated epicenter locations are close to Oakland, the ground-motion simulation results of this study match the ground-motion simulation map produced by the HayWired case study. For these two cases, a comparison between Figure

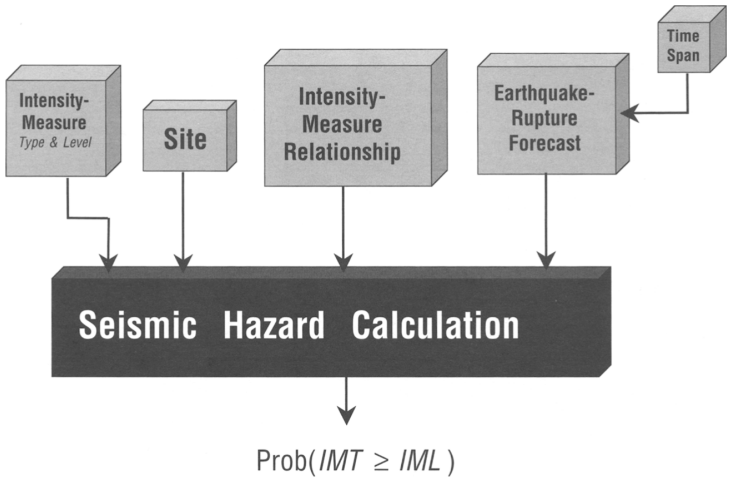


Figure 4.2: The fundamental elements needed for conducting an earthquake hazard calculation in OpenSHA [44].

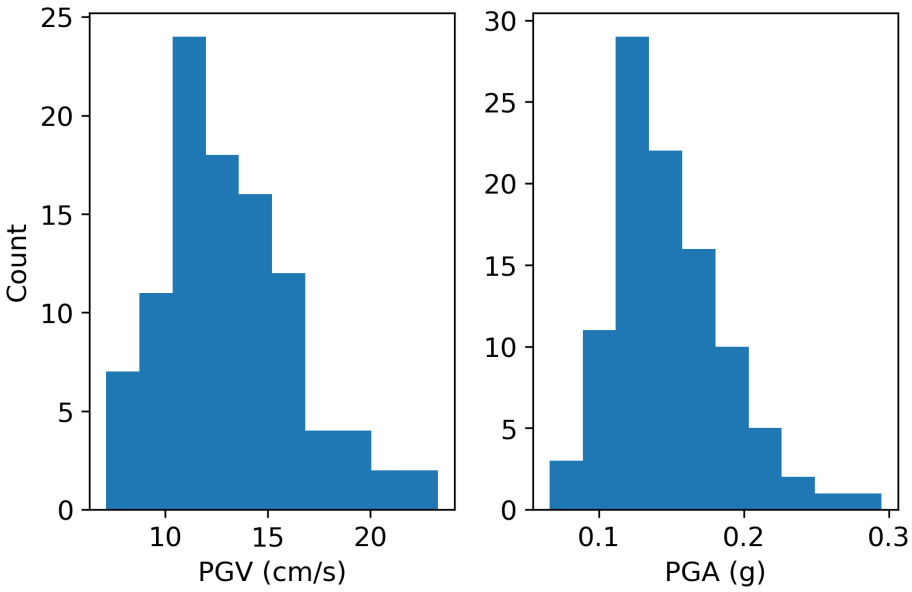


Figure 4.3: Distribution of the mean values of IMs (over all simulated sites for each realization) for 100 earthquake scenarios.

4.4 with Figure 4.5 demonstrates that both the PGV magnitude range and corresponding spatial distribution match closely.

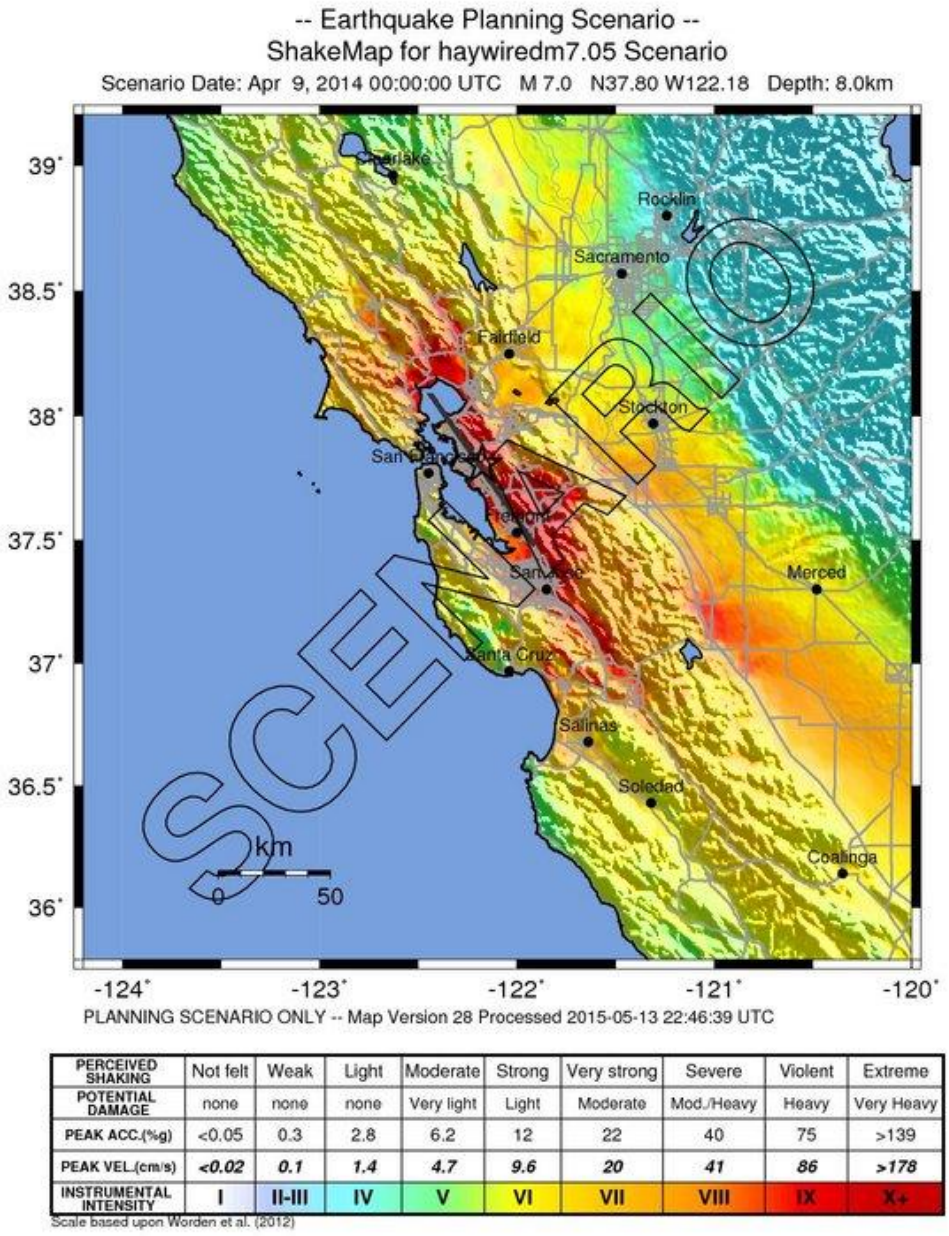


Figure 4.4: Ground-motion intensity map of the HayWired case.

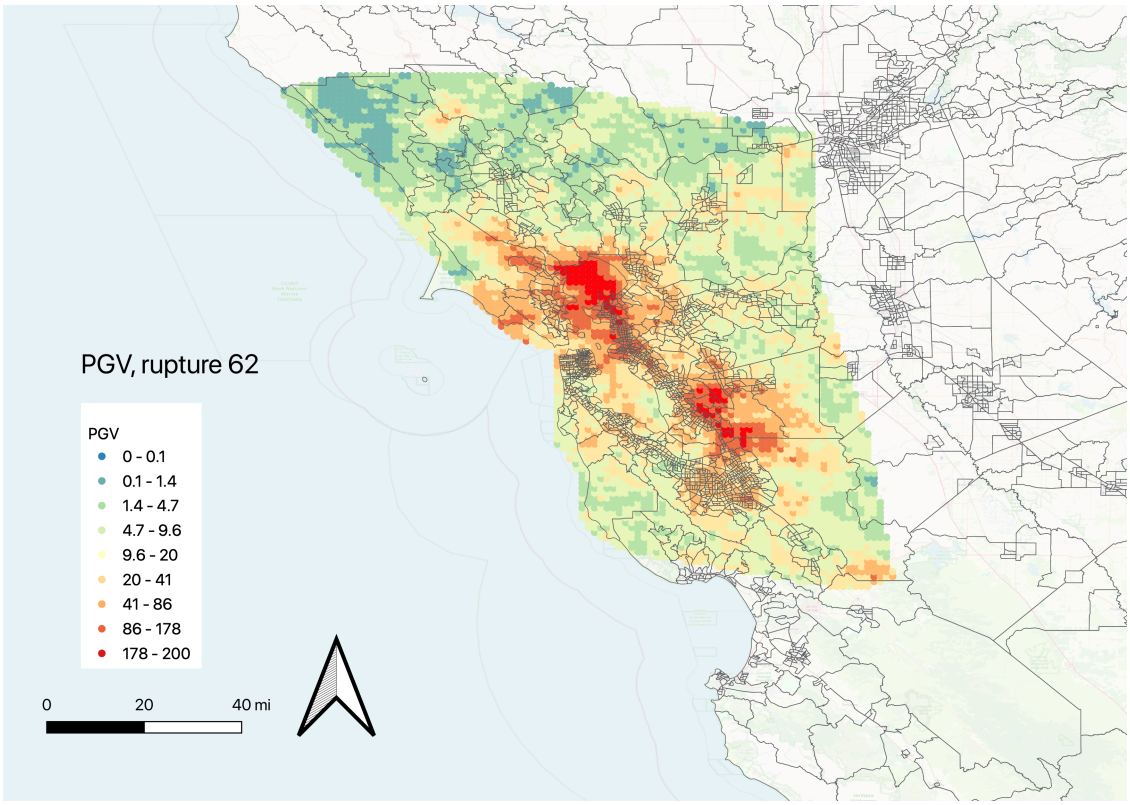


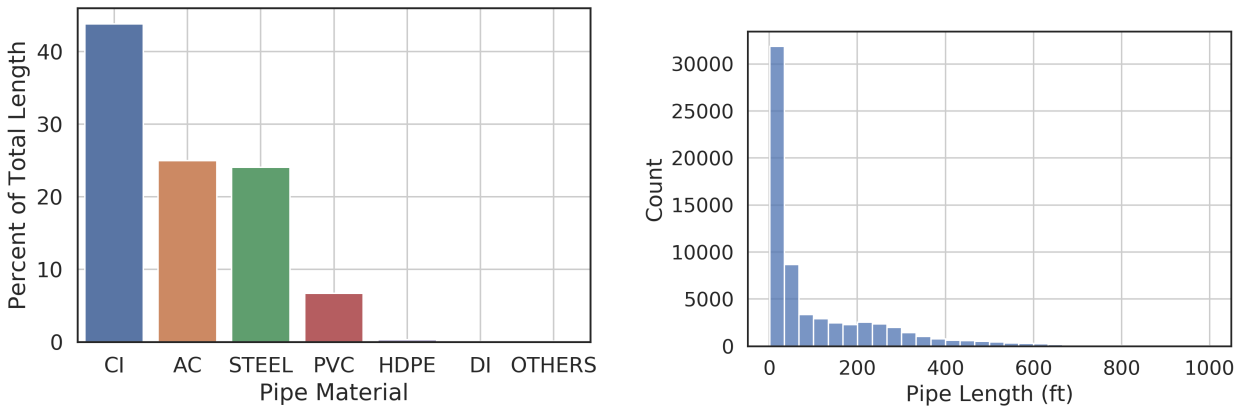
Figure 4.5: Ground-motion intensity map of a selected simulation scenario.

4.3 WDN Attributes

Table 4.1: Material types and their abbreviations.

Abbreviation	Description
CI	Cast Iron
Steel	Steel
AC	Asbestos Cement
PVC	Polyvinyl Chloride
HDPE	High Density Polyethylene
DI	Ductile Iron

The EBMUD gravity-feed zone consists of 65,700 distribution pipes with a total length of 1368 miles (2201 km). The properties of pipes are shown by Figure 4.6. The majority



(a) Pipe material distribution

(b) Pipe length distribution

Figure 4.6: Pipe properties of the Alameda Island WDN: (a) pipe-length distribution of the study area; and (b) pipe-material distribution of the study area.

(98.4%) of the pipes are less than 600 ft long, and the minimum, mean, and maxiumum pipe lengths are 1 ft, 109 ft, and 2711 ft, respectively. Table 4.1 listed the pipe materials used in the EBMUD network. The majority of pipes are made with cast iron (CI) and asbestos cement (AC), which make up around 70% of pipes in the system. Steel and PVC pipes also exist; pipes made with other materials are rare in the system. The roughness coefficient values for the pipes are provided by EBMUD.

The averaged annual day demand for the year 2018 was used for this service area, which is around 50,000 gallons per minute (GPM) in total. The service zone is supplied by seven control stations located at the boundary of the service zone, which are shown as red stars in Figure 4.7. The control stations are located at relatively high elevations (Figure 4.8) to ensure a sufficient water supply for the rest of the system. The hydraulic heads for the sources are set at 150 ft, which satisfies all the demand in this area before the hazard occurs. In this study, pumping operations were not considered since the major driving force for water distribution in this service zone is gravity. Compared to the empirical equation based WDN resilience model by [106] in the HayWired case study, the current model is based on hydraulic simulation of the damaged network using HydrauSim. Moreover, a stochastic simulation process was adopted in this study to evaluate WDN responses considering earthquake ground-motion uncertainties.

4.4 Methodology

The potential earthquake hazard on a WDN is modeled as follows:

1. Using the simulated earthquake ground-motion IMs (PGV), the probability of failure

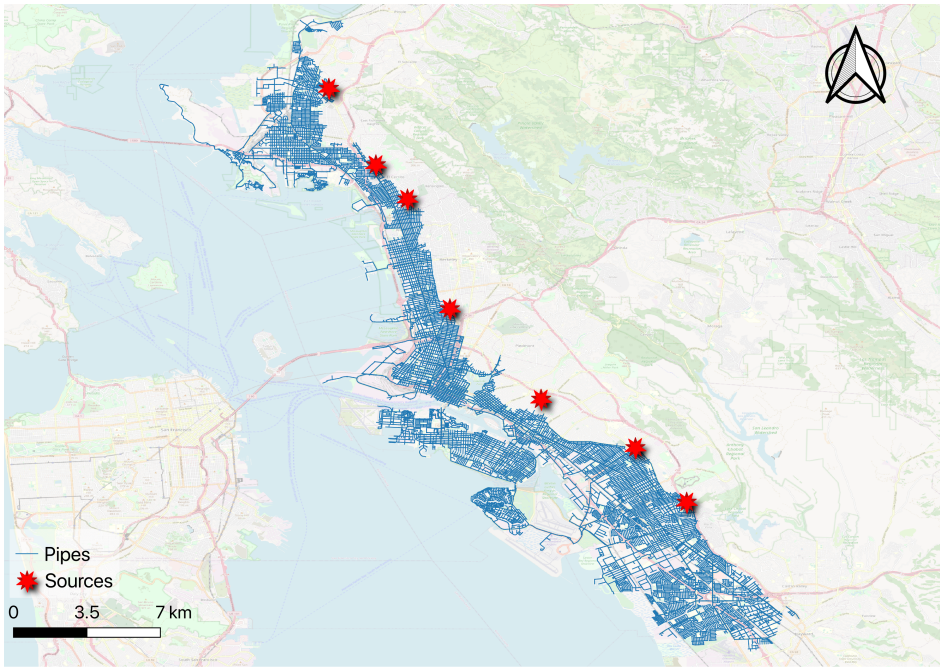


Figure 4.7: Water distribution network used for the study.

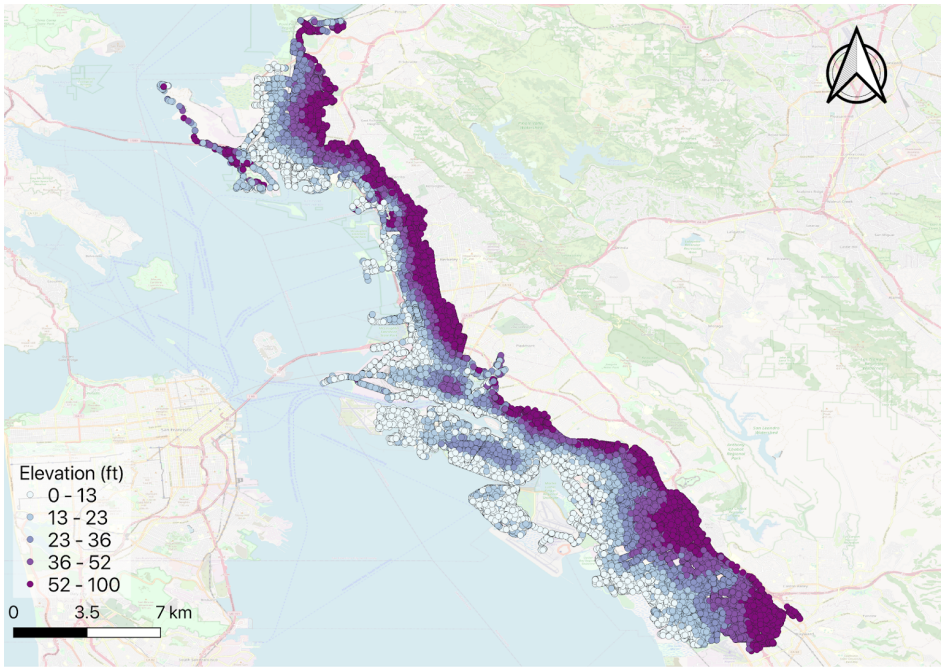


Figure 4.8: Elevation map of the study area.

for each network component is estimated using the pipeline fragility curves.

2. Based on the estimated fail probability, certain components were sampled as the failed component with the corresponding degree of damage.

3. A PDD hydraulic simulation was performed on the damaged network to estimate the degree of water supply shortage anticipated for the study case.

This procedure is summarized in Figure 4.9.

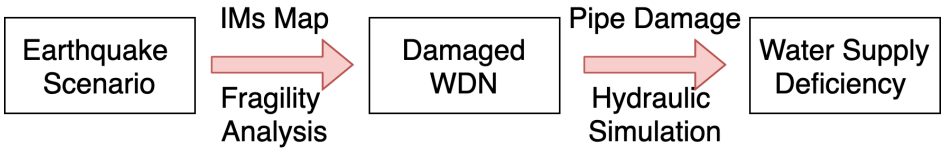


Figure 4.9: Earthquake hazard analysis procedure for a WDN.

Similar to the assumptions made by [106], this study considered only buried distribution pipes as potential fail components. Unlike other critical WDN components such as tanks and reservoirs, not many buried distribution pipes were included in the seismic improvement programs due to economical considerations, which make them especially vulnerable to earthquake impacts. The failures of other vulnerable WDN components—such as pumps—were not considered because the region under study relies on a gravity feed zone where pumps are not required. Following [5], an earthquake-induced buried pipe failure event was modeled as a Poisson process given the selected ground-motion level and pipe properties, which can be expressed by Equation 4.3.

$$P_{f,i} = 1 - e^{-RR_i L_i} \tag{4.3}$$

where $P_{f,i}$ represents the probability of failure for pipe i; L_i is the length of pipe i; and RR_i is the repair rate for pipe i modeled by Equation 4.4 [5].

$$RR_i = k_1(0.00187)PGV_i \tag{4.4}$$

where k_1 is the fragility constant, which is dependent on the pipe material, diameter, and other properties. The list of k_1 values used in this study is presented in Table 4.2. PGV_i is the simulated PGV for site i. Note that the ALA vulnerability function for permanent ground deformation was not considered in this study due to the lack of soil data.

Each PSHA earthquake ground-motion scenario contains the simulated PGV values for every simulation grid point; see Figure 4.1. The PGV value for each pipe is estimated by mapping the center of the pipe to the spatially closest simulation grid point. Since PGV values do not vary significantly over small distances, this closest-point mapping method provides a simple yet reasonable estimation for the PGV values of individual pipes. The

Table 4.2: Fragility constant for different pipes.

Pipe Material	Diameter	K
Cast Iron	small	1.0
Ductile Iron	small	0.5
Asbestos Cement	small	1.0
PVC	small	0.5
Welded Steel	small	0.9
Welded Steel	large	0.15

estimated pipe failure rate for a given pipe can be calculated using Equations 4.3 and 4.4. This study ignored those pipes with an extremely small probability of failure probability ($P_{f,i} < 0.0001$) because the total number of pipes for the study area is at the magnitude of $10,000$. After removing these pipes, group sampling on the remaining pipes was proposed to approximate the spatial distribution of damaged pipes. Specifically, the remaining pipes were grouped into 10 bins according to the simulated fail probability (with non-likely fail pipes removed) to reduce the variances and for computation convenience. Within each group, the expected number of pipes to fail is estimated using Equation 4.5.

$$N_{i,fail} = P_{i,fail} * N_i \tag{4.5}$$

where $N_{i,fail}$ is the expected number of pipe failures in group i; $P_{i,fail}$ is the probability of fail within group i, which is estimated as the mean probability of fail for this group; N_i is the total number of pipes in group i.

Using Equation 4.5, pipes (uniform sampling) are randomly selected to fail for all 10 subgroups. Combining the sampled pipe failures from each group gives which pipes are expected to fail for the whole network. The damaged degree for each pipe failure is modeled by the leak-hole diameter or the equivalent orifice diameter (EOD) for the pipe[121]. According to HAZUS [43], 80% of pipe damage is assumed as leaks, and 20% is assumed as breaks. For leaking pipes, the EOD is randomly assigned between 5% and 25% of the pipe diameter [121]. For broken pipes, the EOD is chosen as 80% of the pipe diameter to approximate heavy water loss due to broken pipes. The damaged network is then summarized into a .inp (EPANet input file format) file, and HydrauSim is used to simulate the steady-state hydraulic of the damaged WDN.

Since both the exact pipe failure locations and the corresponding degree of damage is uncertain, a Monte Carlo simulation was considered for this study. Specifically, for each earthquake scenario, 500 different damaged network scenarios were generated and simulated

using HydrauSim. The final hydraulic simulation result was obtained through averaging the hydraulic measure of the water shortage ratio for every damaged network scenario. Figure 4.10 summarizes the WDN hazard analysis workflow for a given earthquake scenario.

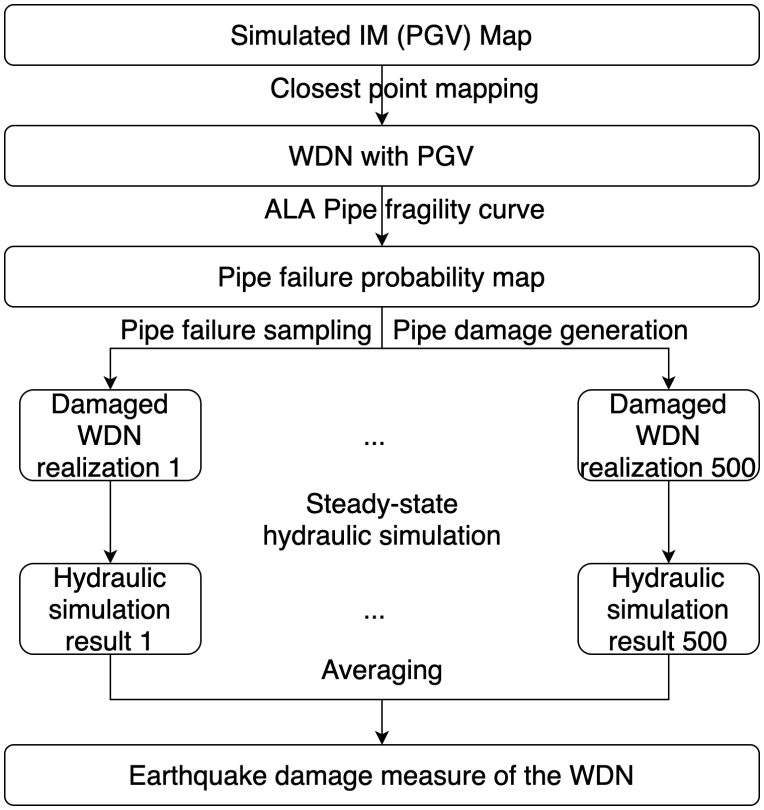


Figure 4.10: WDN hazard analysis workflow for a simulated earthquake scenario.

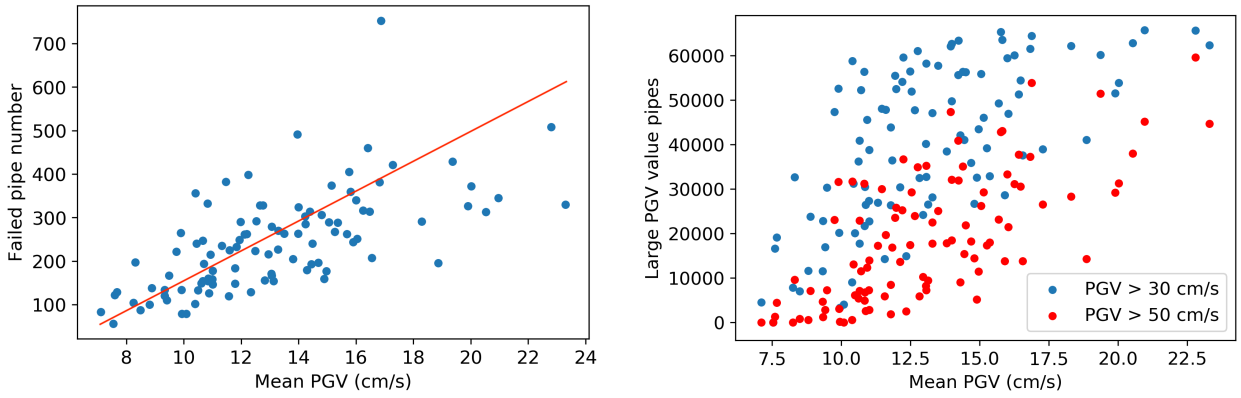
4.5 Results and Discussion

Figure 4.11(a) shows that as the level of ground movement (mean PGV) increases, more pipes tend to break. For the 100 earthquake ground-motion scenarios produced by the PSHA model, the minimum number of expected pipe breaks is 56 (corresponding to mean PGV 7.5 cm/sec), and the maximum number of pipe breaks is 752 (corresponding to mean PGV of 18 cm/sec). On average, 242 number of pipes in the study area were considered broken. The HayWired study [106] reported 2037 pipes might be damaged by the earthquake mainshock wave passage for the entire EBMUD network (6698 km). Converting the damage number to the scale of the EBMUD gravity feed zone, around 600 pipes will be damaged based on the HayWired report estimation, which falls into the range of 56–752 damaged pipes obtained by this study. Comparing the number of pipe breaks in the HayWired study

to the number of pipes obtained in this study (around 65,000), the number of pipe failures is small. As discussed below, even a small number of pipe failures may have a noteworthy impact on demand.

Although Figure 4.11(a) shows a clear correlation between mean PGV value (average across all simulated sites for each scenario) and the number of pipe failures (marked by the red line), variations on the number of broken pipes for a certain mean PGV value can still be observed. This is because the mean PGV value as an aggregated measure does not reflect the spatial variances of the ground-motion intensities. For example, one earthquake scenario may have a larger mean PGV value than another case, but the severe ground shaking (high PGV values) in the region occurred in non-populated areas, resulting in fewer broken pipes, which has a higher ground-motion intensity compared to populated areas.

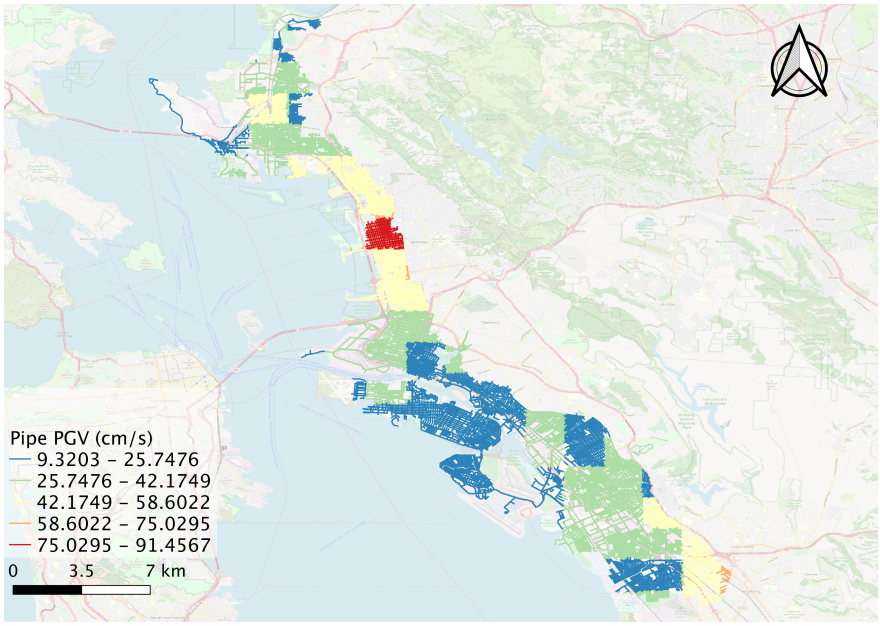
This is illustrated in Figure 4.11(b). As the mean PGV value of the study area increases, the number of pipes with large PGV values increases. Moreover, note the large variances of the numbers of pipe breaks with a given mean PGV range, demonstrating the spatially heterogeneous distribution of PGV values across different realizations. Moreover, differences in material and length across network pipes also contribute to the variations of number of pipe failures since they are important factors in the fragility function; see Equations 4.3 and 4.4). Figure 4.12 shows a map of both the PGV level experienced by the pipes and the resulting probability of pipe failure for the region of study of one earthquake scenario. In general, the high PGV value of the region tends to have a high probability of pipe failure; the region chosen as an example is the city of Berkeley. The variations of the intrinsic properties of the pipes are also reflected by the variations in the probability of pipe failure within the area with similar PGV values.



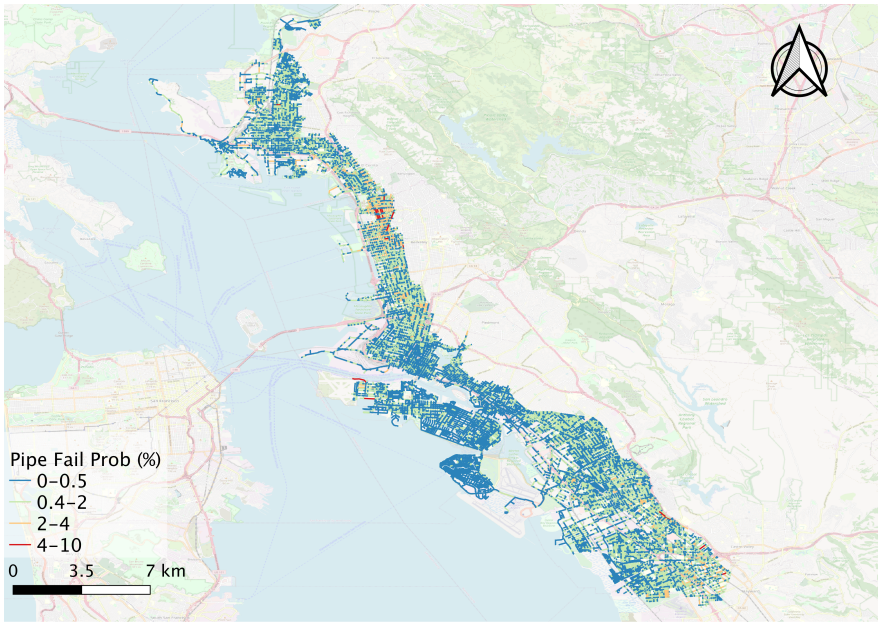
(a) PGV level of pipes

(b) Pipe fail probability

Figure 4.11: WDN response to ground motion inputs; (a) simulated number of pipe failures for all 100 earthquake ground-motion scenarios with respect to mean PGVs; and (b) number of pipes experiencing large PGV values for all 100 earthquake ground-motion scenarios.



(a) PGV level of pipes



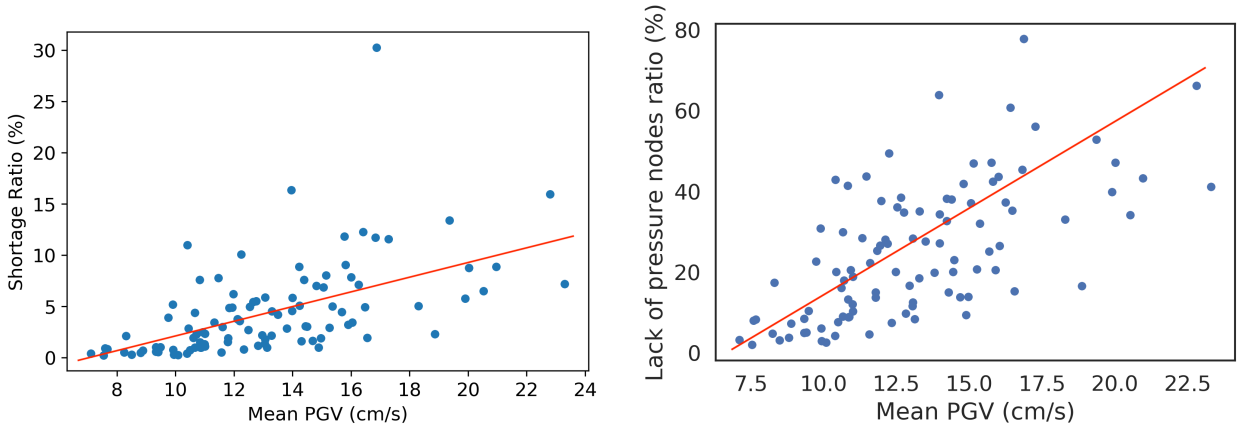
(b) Pipe failure probability

Figure 4.12: Earthquake potential impacts on a WDN: (a) interpolated PGV level for each pipe; and (b) the resulting map of probability of pipe failures using the fragility curve from [5]. The shown case has a mean PGV value of 13.06 cm/sec, and the average number of pipe breaks determined from the simulation is 167.

To quantify the consequences of the earthquake to the WDN, Monte Carlo simulations were performed based on the generated pipe failure map for each earthquake scenario. The result is summarized in Figure 4.13. Each point on the plot represents a single earthquake scenario. Figure 4.13(a) shows the relationship between the mean PGV value of a scenario and the resulting total water shortage ratio, which is defined by Equation (4.6).

$$S_{tot} = \frac{(demand_{tot} - supply_{tot})}{demand_{tot}} * 100 \tag{4.6}$$

where S_{tot} is the total water shortage ratio for the WDN; $demand_{tot}$ is the total demand for the WDN; and $supply_{tot}$ is the simulated total supply for the WDN.



(a) Total water shortage ratio

(b) Lack of demand ratio

Figure 4.13: Simulated hazard consequence for the WDN of 100 earthquake scenarios: (a) shows the total water shortage ratio for each scenario (depth of the hazard); and (b) shows the ratio of the number of demand nodes that are below normal operating pressure (width of the hazard).

A clear trend (see the red line on the figure) is noticeable in Figure 4.13(a) where the degree of total water shortage increases as the overall ground shaking level increases. The variations of the shortage ratio are caused by the variations of the number of pipe failures for a certain PGV value, as discussed previously. The minimum shortage ratio is only 0.234%, implying a limited earthquake impact on the WDN in some cases. On the other hand, the maximum shortage ratio is around 30%. On average, the total water shortage ratio is around 5%. Considering the small number of pipe failures generated by the simulation—around 0.45% from the total pipes—it is reasonable to conclude that the WDN performance is sensitive to strong ground motion. Although the total water shortage ratio measures the total severity of the earthquake impact on a WDN, it does not reflect how this disruption of the water supply is distributed among users. Hence, another measure, the lack of pressure

nodes ratio, is used to capture the width (number of influenced demand nodes) from the earthquake on a WDN. As given by the following equation.

$$r_{lack} = \frac{(N_{tot} - N_{shortage})}{N_{tot}} * 100 \tag{4.7}$$

where r_{lack} is the lack of pressure nodes ratio; N_{tot} is the total number of nodes that experience demand, and $N_{shortage}$ is the number of nodes that experience pressure below normal operating conditions due to the damaged WDN. Insufficient node pressure often leads to water supply deficiency (see Chapter 3 for more detail). In this study, nodes with pressure less than 30 psi were defined as lack of pressure nodes. r_{lack} indicates the spatial impact of the earthquake on water supply for the region. Post-hazard actions, such as isolation valve closure due to inspection/maintenance, might cause water shortages for the impacted region.

Figure 4.13(b) shows a trend (see the red line on the figure) that the lack of pressure nodes ratio increases as the mean PGV value increases. Again, the variations are caused by the variations in the number of pipe failure for a certain PGV value. Compared to the total supply shortage ratio, the lack of pressure nodes ratio shows much more variations. For the simulated 100 ground-motion scenarios, the minimum lack of pressure nodes ratio is only 2%, but the maximum value reaches 77.77%. The average is about 24.98%.

Overall, the studied earthquake scenarios do not create a severe total water shortage when measured against the total supply quantity (5% on average). However, the width of the impact, which is measured as the number of demand nodes with insufficient water pressure, is huge. On average, about 25% of demand nodes may experience water shortages due to the lack of pressure, which can rise to 78% for the worst case scenario. The worst-case scenario estimation matches well with the HayWired report's assessment, which suggests that approximately 75% of services will be impacted after the earthquake event [106].

To further explore the relationship between the WDN damage state and the resulting water supply shortage, the number of pipe failures versus the total supply shortage ratio for all earthquake scenario realizations is plotted in Figure 4.14. One important feature is that the variance of the shortage ratio for a certain degree of pipe failures is small: the relationship between a pipe failure number and the resulting total water shortage ratio is nonlinear. Specifically, the trend is convex, i.e., the total water supply shortage ratio increases faster when the number of pipe failures increases. As the damage state of a WDN increases, users in the less-damaged region may experience severe water shortages due to water path blockages from severely damaged regions.

Two simulated scenarios are plotted to illustrate this phenomenon. Figure 4.15 represents a case with a moderate damaged state. The mean PGV value for this case is around 13 cm/sec, and the averaged simulated pipe break number is 167. A comparison between Figure 4.15(b) and Figure 4.15(a)demonstrate that the water supply shortages mainly occur on the damaged part of the network (top and bottom). Undamaged or lightly damaged parts of WDN do not experience a major supply shortage (middle part). Figure 4.16 shows a heavily damaged WDN case, where the number of simulated pipe breaks is around 750. When the WDN is severely damaged as a whole, all sections of the network may experience severe

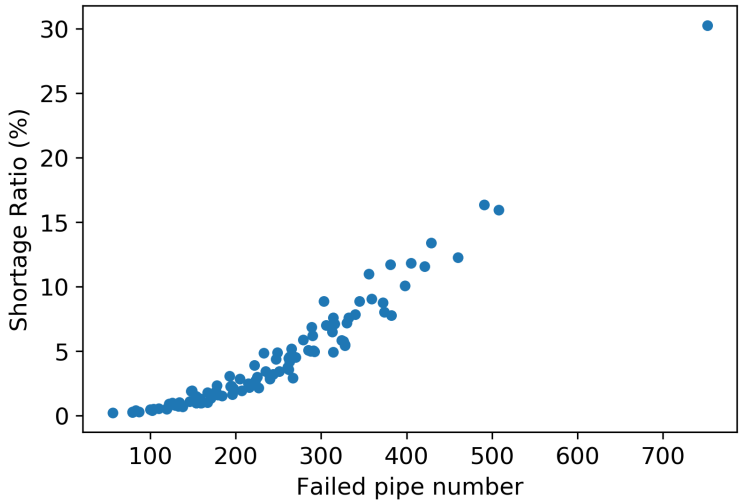
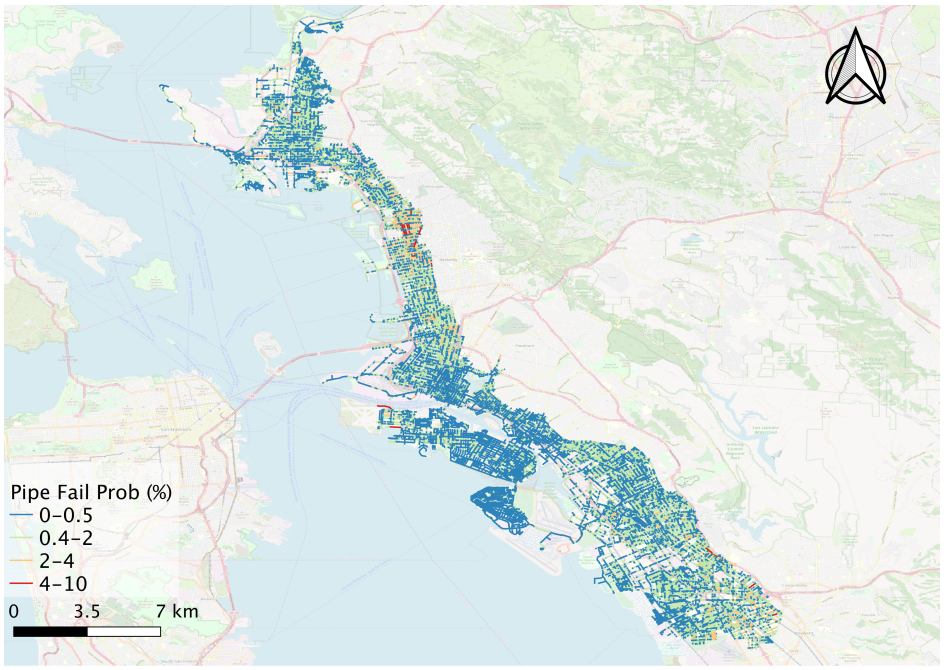


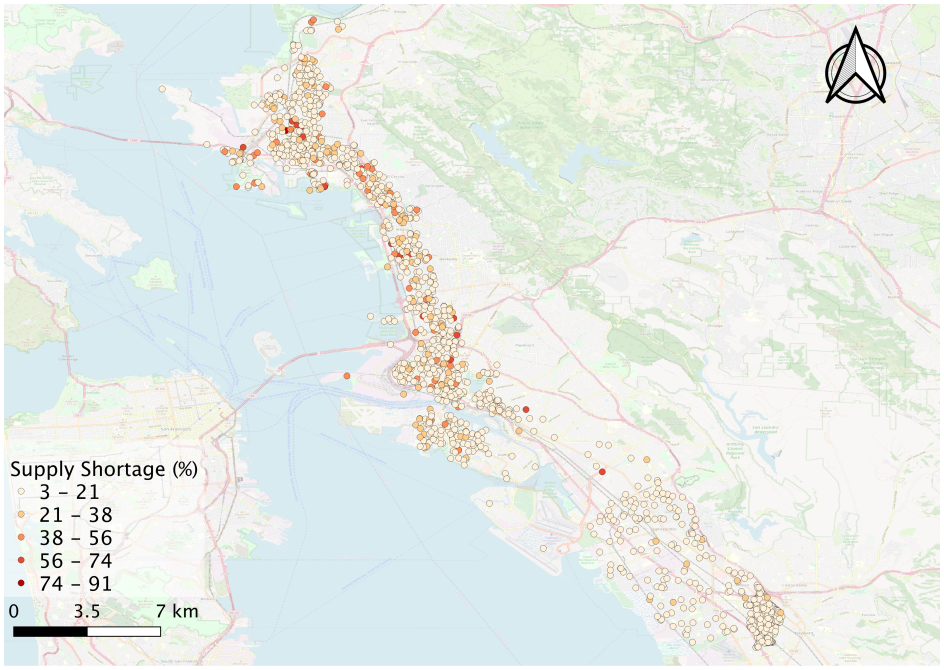
Figure 4.14: The relationship between the number of pipe failures and simulated total water shortage level for 100 earthquake scenarios.

supply shortages regardless of local damage states. Figure 4.16(b) shows no obvious spatial pattern on the water supply shortages, whereas certain areas (bottom right) are expected to have more pipe breaks compared to other areas, as illustrated by Figure 4.16(a).

Figure 4.17 shows the average simulated water shortage ratio across all 100 simulated scenarios. One noticeable feature of the averaged result is that the number of trivial impacted nodes (nodes that experience no or very small water supply loss) is very large, consisting 70% of the overall supply nodes. Note: this number is deceptive as spatial variances of the simulated water supply shortages are large across individual realizations, as previously discussed. Figure 4.17(b) shows that the relative standard deviation (RSD) of the simulated water shortage ratio is high across most of the study area, reaching over 100% (the standard deviation is bigger than the mean water shortage for most of the nodes) for some nodes. Moreover, nodes with small water supply shortage ratios tend to have larger RSD values compared to nodes with larger water supply shortage ratios. Such relationships implies that simulated scenarios agree on the locations of nodes experiencing severe supply shortage but not on the locations of minor impacted nodes. Hence, it is statistically meaningful to draw conclusions on high degrees of water shortage nodes but not on other nodes. Figure 4.17(a) shows that a high degree of water shortage nodes tends to group at certain parts of the network. The elevation map, Figure 4.8, shows that such clusters correspond to areas that are relatively high in elevation (i.e., the lower right corner). One can understand this phenomena by considering the energy change of the damaged WDN. Damaged WDNs tend to lose energy because of water leakages compared to the pre-hazard state. Thus higher elevations are more likely to lose water supply regardless of the actual distribution of the damaged pipes due to energy concerns.

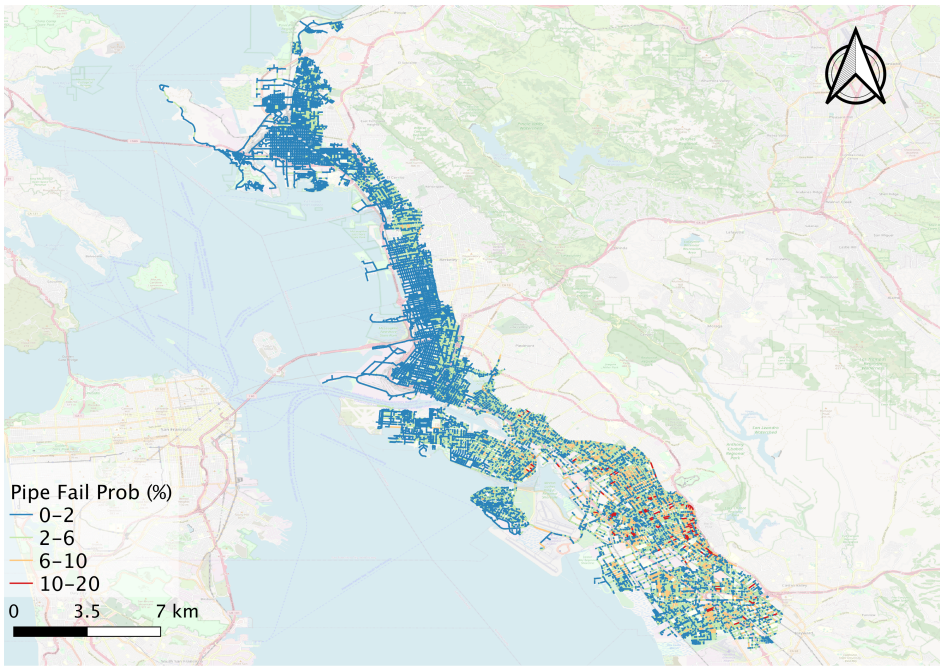


(a) Pipe failure probability

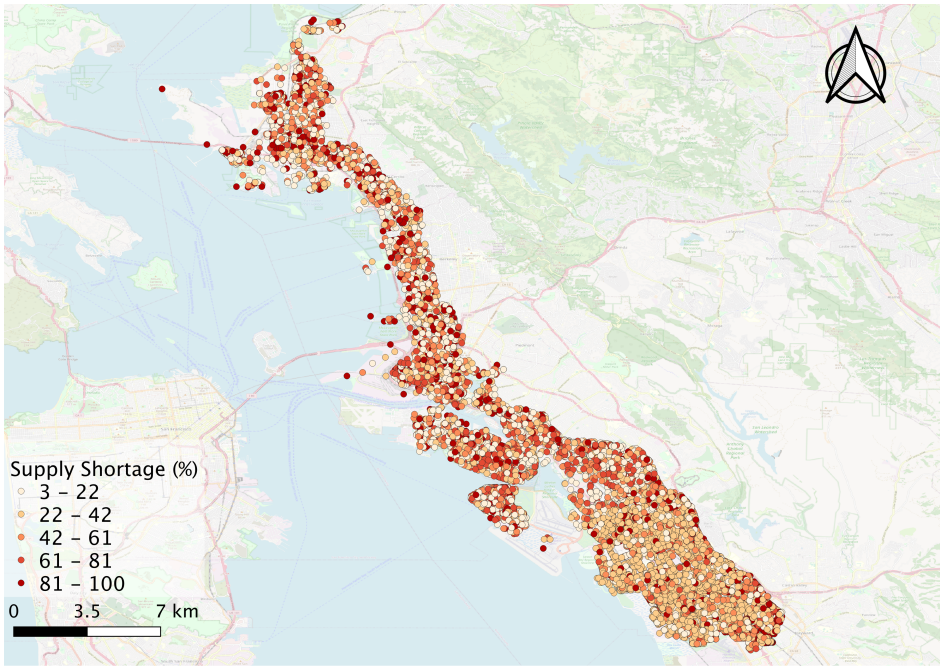


(b) Demand shortage distribution

Figure 4.15: Pipe failure rates and simulated water shortages for a medium damage-level earthquake scenario, which has a mean PGV value 13.06 cm/sec and averaged simulated pipe break number 167.

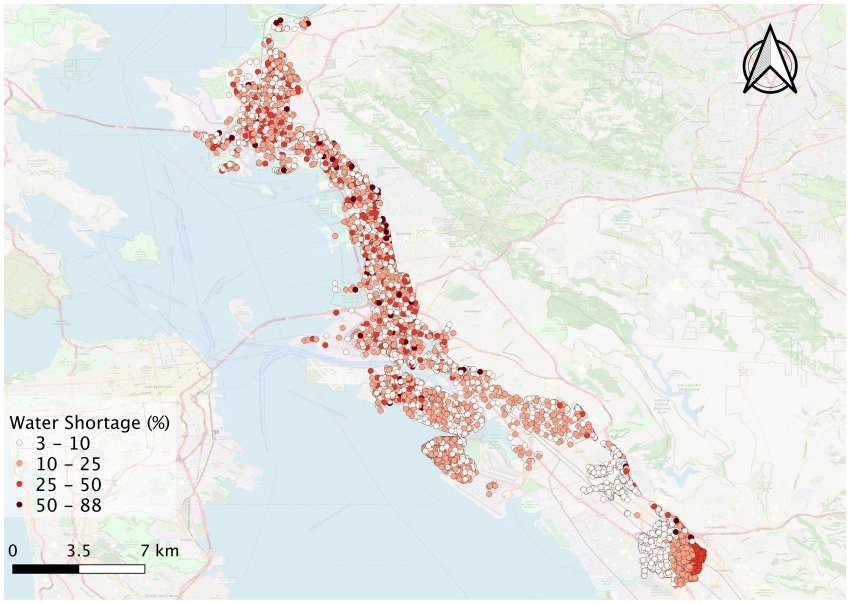


(a) Pipe failure probability

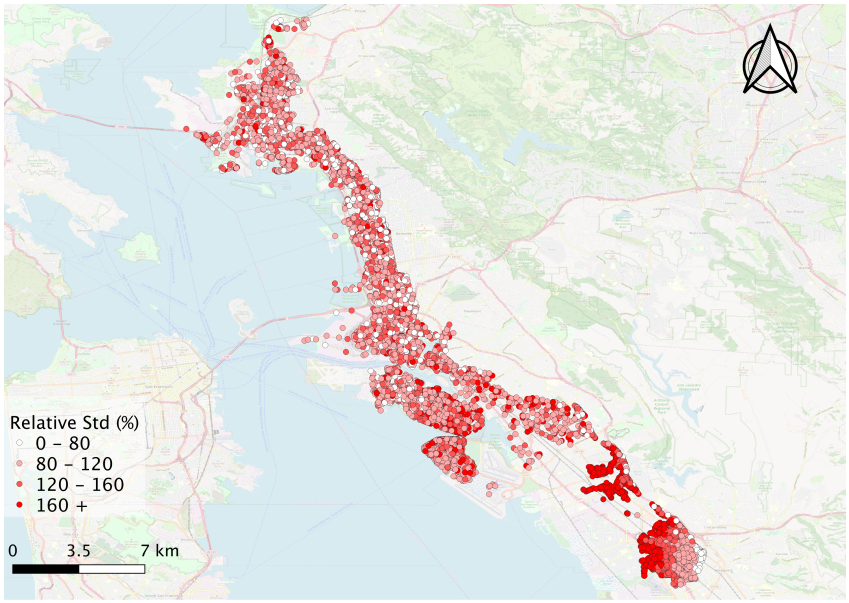


(b) Demand shortage failure

Figure 4.16: Pipe failure rates and simulated water shortages from a large damage-level earthquake scenario, which has a mean PGV value of 16.87 cm/sec and averaged simulated pipe break number 752.



(a) Average simulated water shortage ratio



(b) Relative standard deviation ratio

Figure 4.17: Aggregated results from all earthquake scenarios; (a) Average simulated water shortage ratio across 100 scenarios simulations; and (b) relative standard deviation ratio of simulated water shortage ratio across simulations. Note that supply nodes with trivial water supply shortage (smaller than 3%) are removed from the graph for visualization purposes.

4.6 Summary

Previous studies focus on evaluating earthquake responses for either a small or simplified WDN under multiple scenarios [13, 125, 80], or a large network under single hazard scenario [122, 106]. Using the newly developed efficient WDN hydraulic simulator, this study explores the earthquake response of a large-scale WDN (65,700 pipes) under various ground-motion scenarios (100 realizations).

This study analyzes the impact of M7.05 Hayward Fault rupture to a regional WDN in the San Francisco Bay Area. The ground-motion scenarios were generated using a probabilistic seismic hazard analysis (PSHA) approach. Specifically, A set of 100 rupture events from UCERF2 [144] for a M7.05, Hayward-Rodgers Creek HN+HS earthquake were simulated with spatial correlated earthquake intensity measures (IMs).

The earthquake's impact on the water distribution network was estimated using the generated IMs. First, the failure probability of pipes was calculated using pipeline fragility curves with the generated earthquake ground motion IMs (PGV) as inputs. Then, leaked/broken pipes were sampled based on the estimated failure probability, and the earthquake-induced regional water shortage was quantified using the developed hydraulic simulator, HydrauSim. Note, Monte Carlo simulation was used to deal with uncertainties of failure type and locations.

The simulation was performed on East Bay Municipal Utility District (EBMUD) main gravity feed zone (65,700 pipes). Simulation results show that the WDN damage degree varies tremendously even the earthquake fault line and magnitude are fixed (M7.05 Hayward Fault rupture). Due to the large size of the WDN, uncertainties of earthquake epicenter greatly impact the number of potential malfunctioning pipes in the system. Around 200–800 pipes were estimated to break during the simulated earthquake events. The other important finding is that the relationship between a pipe failure number and the resulting total water shortage ratio is nonlinear. Specifically, the total water supply shortage ratio increases faster when the number of pipe failures increases. As the damage state of a WDN increases, users in the less-damaged region may experience water shortages due to water path blockages from severely damaged regions. On average, 25% of demand nodes may experience insufficient water pressure level, which can rise to 78% for the worst-case scenario.

Chapter 5

Pipe Isolation Risks Under Valve Condition Uncertainties

5.1 Isolation Valves and Isolation Segment

A comprehensive analysis of the impacts of pipe failure events for a WDN requires the consideration of isolation valve conditions. In real-life situations, failed pipes need to be isolated from the main network by closing the corresponding valves to prevent the effects of individual events from spreading throughout the system [72]. Only when a pipe has operable valves at both ends, it can be isolated individually. However, many water utilities tend to use sparse valving strategies (e.g., not every pipe has two valves installed) to reduce cost [37]. Furthermore, the actual valves may not work properly at the time of usage [11]. Common causes of valve failure are a) broken valve stem (or other mechanical parts such as rounded operating nut) and b) inability to turn/reach the valves because of an obstruction [84]. Hence, it is necessary to consider the valve conditions of a pipeline system when analyzing the impacts of pipe failure events for a WDN.

Walski formulated the concept of isolation segment to facilitate the analysis of pipe failure risk considering valve conditions [138]. An isolation segment is defined as the smallest set of pipes that can be isolated after closing all the surrounding valves. Figure 5.1 illustrates the idea of isolation segment for a simple WDN. In this case, closing valves $V2, V3, V5$ forms an isolation segment that contains $P2, N2, N4$. Based on the concept of isolation segment, Walski developed the idea of Segment-Valve representation [138]. Unlike the conventional Junction-Pipe representation, whereas graph links are WDN pipes and graph vertices are WDN junctions, isolation segments and valves become the graph vertices and edges in the Segment-Valve representation, respectively. Figure 5.1 shows the equivalence of the Junction-Pipe and Segment-Valve representation for a WDN.

Using the Segment-Valve representation for a WDN allows one to analyze the consequences of failed valves and pipe isolation events through graph operations. Specifically, failure of a valve is equivalent to removing the corresponding valve edge from the Segment-

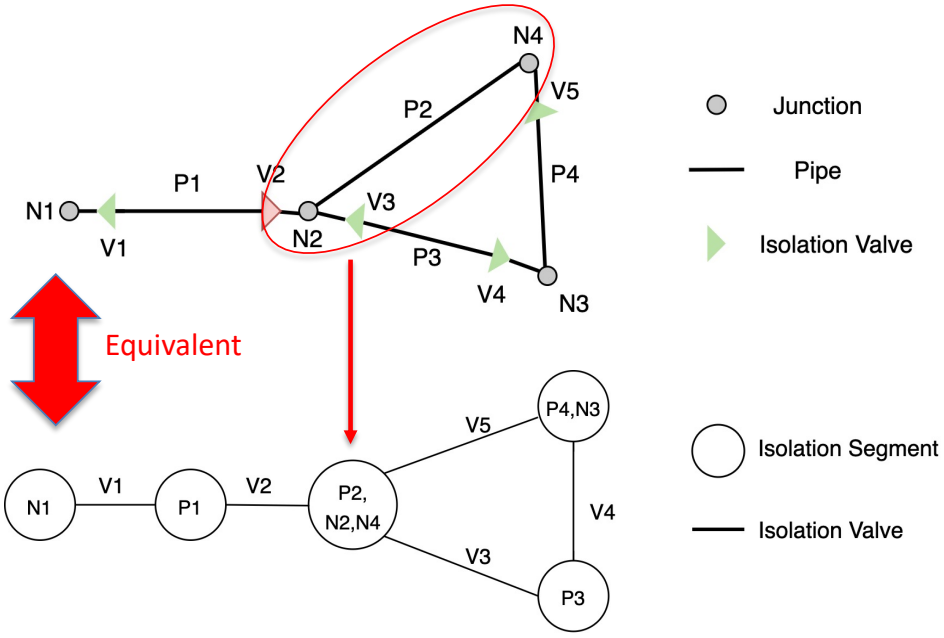
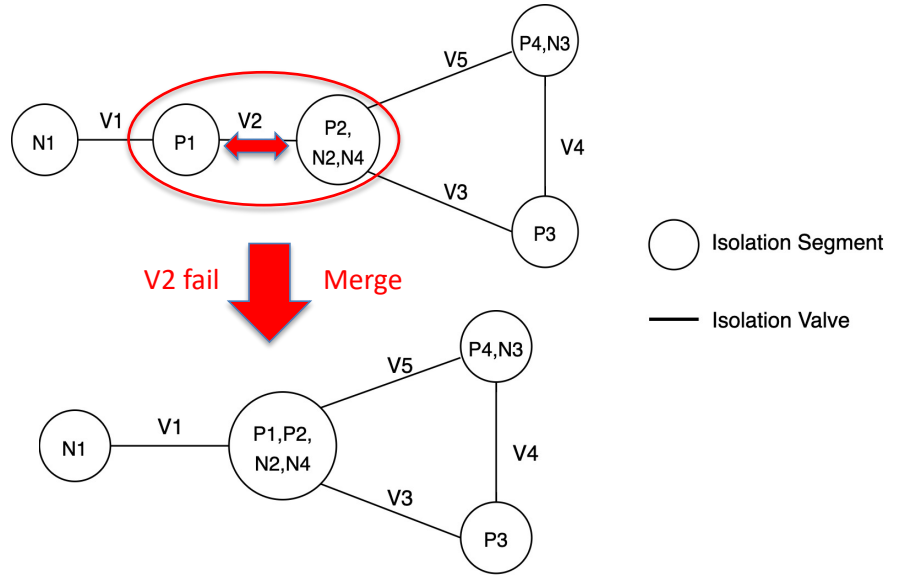


Figure 5.1: Junction-Pipe representation Vs. Segment-Valve representation. Isolation segments are found by closing all the available valves in the network

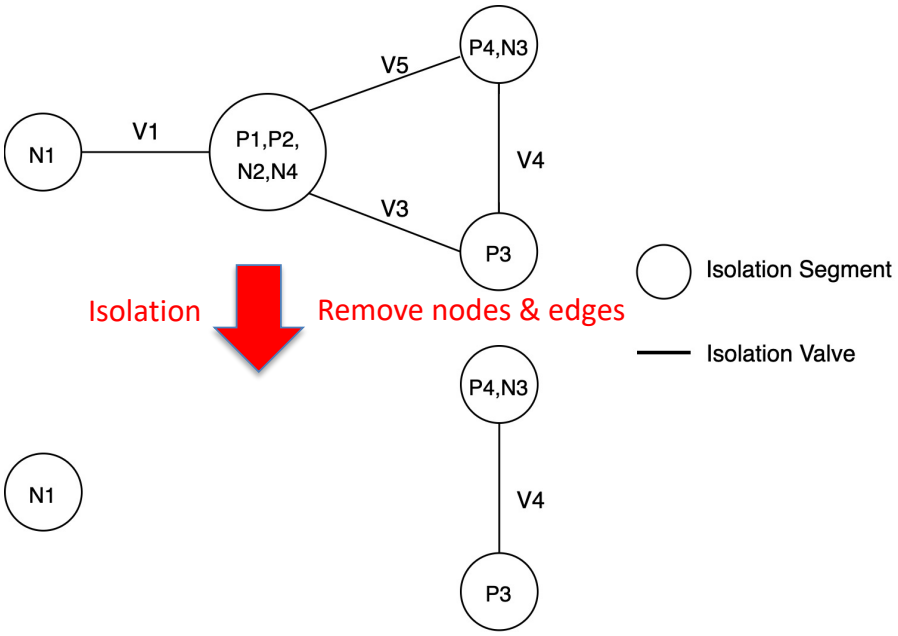
Valve graph. The removal operation is achieved through edge contraction, where the failed valve edge is deleted, and the pair of previously linked isolation segments is merged into a new isolation segment (node in the graph). To examine the consequence of isolating a failed pipe, one needs to identify the corresponding isolation segment first. The identified isolation segment is then removed from the graph by deleting both the segment node and related valve links. The state of the WDN after the pipe isolation is represented. Figure 5.2 shows an example of modeling valve failure and segment isolation of a WDN. In Figure 5.2a, when valve $V2$ fails, the $V2$ edge is removed from the graph, and the two linked segments $S1$ and $S2$ are merged. The newly formed segment contains all the components of the merging segments, and the connection relationships of other segments remain intact. Next, suppose a failed element inside the newly formed segment, such as the failed $P1$, needs to be isolated from the system. In that case, the post-isolation system is obtained by deleting the corresponding segment and related valve edges ($V1$,$V3$,$V5$), as shown in Figure 5.2b.

Since WDNs are conventionally represented in Junction-Pipe format (such as the EPANET .inp file [112]), a conversion tool is required to construct the Segment-Valve for a WDN. This study adopts the Junction-Pipe to Segment-Valve representation conversion algorithm proposed by Jun [72]. The workflow of the algorithm is stated as follows:

1. Given a pipe that needs to be isolated, the corresponding isolated nodes will be found



(a) Isolation valve failure event



(b) Segment isolation operation

Figure 5.2: Graph operations for the segment valve graph of a WDN. The Pipe-Junction representation of the network is shown in Figure 5.1

first, and they are pipe end nodes that do not have a valve installed on the searching pipe.

2. For a found isolated node i, identify its corresponding valve lacking pipes, which are pipes that do not have valves installed close to the node i.

3. Perform step 1 for the newly found valve lacking pipes.

4. Stop when there is no pipe and node to explore.

5. Add all the valves that are related to isolated pipes and nodes as the valves that need to be closed for this segment. Program ends.

This algorithm is illustrated by going through an example using the WDN shown in Figure 5.1. Assuming the failed pipe $P1$ needs to be isolated but valve $V2$ is also failed, the segment finding algorithm works as follows:

1. Searching for isolated nodes of pipe $P1$, which is node $N2$.

2. Searching for valve deficient pipes from node $N2$, which is pipe $P2$ ($P1$ has been explored).

3. Searching for isolated nodes of pipe $P2$, which is $N4$ ($N2$ has been explored).

4. Searching for valve deficient pipes from node $N4$, which is none ($P4$ has $V5$).

5. No new pipe or node to explore, add all related valves, which are $V1$, $V3$, $V5$. Isolating the pipe $P1$ will result the isolation of $P1$,$P2$ and $N2$,$N4$ by closing $V1$, $V3$, $V5$.

In this study, some modifications are made to the original algorithm to boost computation efficiency. Specifically, redundant computations are avoided by replacing the recursion algorithm in the original algorithm with lookup tables for column and row search operations. In addition, auxiliary functions such as locating the corresponding isolation segment for a given pipe, updating the graph status based on valve condition changes are modeled and implemented in the programming language Python.

5.2 Risk Formulation for Subsystem Isolation

Definition

This research focuses on quantifying pipe isolation risk by evaluating the immediate system-wise hydraulic response of a WDN after an isolation segment or a set of isolation segments is removed from the primary system due to pipe inspection and repair. The pipe isolation

risk is modeled as a combination of the likelihood and the consequences of a failure [53]. In this study, the risk is defined as the expected cost of a pipe failure event [32, 51]:

$$\bar{D} = \sum_{i=1}^{N} f_i D_i \tag{5.1}$$

where N is the number of pipes in the system, f_i is the failure probability of the pipe i within N pipes in a system (if equal probability $f_i = \frac{1}{N}$), and D_i is the cost related to the failure of pipe i. Commonly used costs are system-wise hydraulic properties, such as system-wise pressure drop and supply loss [51], and economic impacts, such as the total isolation length and repair cost [53].

As discussed in Section 5.1, valves need to be closed prior to the pipe isolation operation. Therefore, isolation segment of the failed pipes need to be removed from the pipe failure event. In this case, Equation 5.1 can be rewritten as the expected cost of isolating segments with failed pipes:

$$\bar{D} = \sum_{j=1}^{C} f'_j D'_j \tag{5.2}$$

where C is the number of segments in the system, f'_j is the failure probability of segment j, which is the sum of fail probabilities of pipes inside the segment:

$$\begin{aligned} f'_j &= \sum_{i=1}^{x_j} f_i \quad \text{in segment j} \\ &= x_j \bar{f}_j \end{aligned} \tag{5.3}$$

where x_j is the number of pipes inside the segment j, and $\bar{f}_j$ is the mean pipe fail probability for x_j pipes in segment j:

$$\bar{l_j} = \frac{1}{x_j} \sum_{i=1}^{x_j} f_i \quad \text{in segment j} \tag{5.4}$$

D'_j is the cost associated with the failure of segment j, which can be written as:

$$D'_j = x_j \bar{l_j} \tag{5.5}$$

where $\bar{l_j}$ is the mean cost for the x_j pipes in segment j. Note, D'_j may not equal to $\sum_{i=1}^{x_j} D_i$ since removing a sub-network from the system may induce additional cost (unintended cost).

x_j can be further decomposed as $\bar{x} + b_j$, where $\bar{x}$ is the average segment size (average number of pipes inside a segment), and b_j is the size deviation of segment j to the mean segment size ($b_j = x_j - \bar{x}$). Note that $\sum_{j=1}^{C} b_j = 0$. Using this notation, Equation 5.2

becomes:

$$\begin{aligned}\bar{D} &= \sum_{j=1}^{C}(\bar{x}+b_j)^2\bar{f}_j\bar{l}_j \\ &= \bar{x}^2\sum_{j=1}^{C}\bar{f}_j\bar{l}_j + \sum_{j=1}^{C}\bar{f}_j\bar{l}_j b_j^2 + 2\bar{x}\sum_{j=1}^{C}\bar{f}_j\bar{l}_j b_j\end{aligned} \tag{5.6}$$

Equation 5.6 provides a complete description of pipe isolation risk for a WDN with valve condition considerations. For the first term, note that $\sum_{j=1}^{C}\bar{f}_j\bar{l}_j$ is the expected per pipe cost pooled over all segments. This means that the first term $\bar{x}^2\sum_{j=1}^{C}\bar{f}_j\bar{l}_j$ can be interpreted as the system risk when the network is assumed to be configured with valves that produce segments that have the same number of pipes (i.e. 1 for the ideal case). The magnitude of this term depends on the square of segment size $\bar{x}$. When the number of installed valves is insufficient, or the valve failure rate is high, the average segment size $\bar{x}$ becomes large, driving the whole system risk up.

The other two terms act as correction terms to account for the actual segment size variations, which depend on the locations of functioning valves. The $\bar{f}_j\bar{l}_j$ term is positive because all pipes can fail ($\bar{f}_j > 0$), and every failure has a certain consequence ($\bar{l}_j > 0$). The second term is therefore always non-negative since $b_j^2 \geq 0$. This means that having segment size variations across the network always increase the risk.

The effect of segment size variance is a bit complex in the last term. A segment with a size larger than the average size ($b_j > 0$) will increase the system risk and vice versa. Such size-variance induced risk change is scaled by the mean size of the segments across the system ($\bar{x}$). That is, systems with a large average segment size ($\bar{x}$) are more sensitive to the segment size distribution than systems with small average segment sizes. The sign and magnitude of the last term also depends on the distribution of actual pipe failure probabilities, $\bar{f}_j$, and the spatial distribution of the isolation cost $\bar{l}_j$, which is related to both topological and hydraulic properties of the network, and the sign and magnitude of b_j (could be > 0 or < 0).

Further derivations

When the pipe fail probability f_i is unknown, which is common during the planning phase, one may assume that every pipes is equal likely to fail until further information is available.

Under the equal-fail assumption, $f_i = \frac{1}{N}$, and $f'_j = \frac{x_j}{N}$. Equation 5.5 becomes

$$\begin{aligned}\bar{D} &= \frac{\bar{x}}{N}\sum_{j=1}^{C} x_j \bar{l_j} + \frac{1}{N}\sum_{j=1}^{C} x_j \bar{l_j} b_j \\ &= \frac{1}{N}[\bar{x}N\sum_{j=1}^{C}((\bar{x}+b_j)\bar{l_j} + (\bar{x}+b_j)\bar{l_j}b_j)] \\ &= \frac{1}{N}[\bar{x}^2\sum_{j=1}^{C}\bar{l_j} + \sum_{j=1}^{C}\bar{l_j}(2\bar{x}b_j + b_j^2)]\end{aligned} \tag{5.7}$$

Similar to Equation 5.6, the first term of the above equation relates system risk to the quantity of functioning valves. Because the cost of isolation for every pipe is accounted in $\sum_{j=1}^{C}\bar{l_j}$, the variations of this term will not be significant over different valve configurations. Hence, quadratic growth with respect to the average segment size dominates the $\frac{1}{N}\bar{x}^2\sum_{j=1}^{C}\bar{l_j}$ term. In other words, systems with fewer valves (or large $\bar{x}$) suffer more (non-linearly) comparing to systems with more valves installed (small $\bar{x}$). The second term of Equation 5.7 is related to the variations of segment sizes. When $b_j > 0$, meaning the size of segment j is bigger than the average segment size, $2\bar{x}b_j + b_j^2$ is positive. On the other hand, when $b_j < 0$, $2\bar{x}b_j + b_j^2$ is always negative. To see this, we have

$$\begin{aligned}2\bar{x}b_j + b_j^2 &< 0 \\ b_j^2 &< -2\bar{x}b_j \\ b_j &> -2\bar{x}\end{aligned} \tag{5.8}$$

Because $b_j = x_j - \bar{x}$, Equation 5.8 holds when $x_j > -\bar{x}$. Since x_j (number of pipes inside an isolation segment) is a non-negative integer, Equation 5.8 always holds, proving that $2\bar{x}b_j + b_j^2$ is negative when $b_j < 0$. Therefore, large-size segments tend to contribute positively, and small segments contribute negatively to the overall system risk. Since the second term contains both positive and negative terms, it may either increase or decrease the overall system pipe isolation risk after the summation of all segment risk. That is, the sign of it depends on the distribution of averaged cost per pipe inside the segment, $\bar{l_j}$. Only when most large impact segments (large $\bar{l_j}$) have small size (negative $2\bar{x}b_j + b_j^2$) and small impact segments (small $\bar{l_j}$) have large size (positive $2\bar{x}b_j + b_j^2$), segment variations will decrease the overall system risk.

When all the pipes are assumed to have approximately equal isolation consequence, $\bar{l_j}$

can be replaced by a constant $\bar{L}$. In this case, Equation 5.7 can be further simplified:

$$\begin{aligned}\bar{D} &= \frac{1}{N}[\bar{x}^2 C\bar{L} + \bar{L}(2\bar{x}\sum_{j=1}^{C} b_j + \sum_{j=1}^{C} b_j^2)] \\ &= \frac{\bar{L}}{N}(\bar{x}^2 C + \sum_{j=1}^{C} b_j^2) \\ &= \bar{L}(\bar{x} + \frac{1}{\bar{x}} Var(\text{segments}))\end{aligned} \tag{5.9}$$

Note that $\sum_{j=1}^{C} b_j = 0$, $\bar{x} = \frac{N}{C}$ and $Var(\text{segments}) = \frac{\sum_{j=1}^{C} b_j^2}{N}$ are used in above derivation. Equation 5.9 means that when all the pipes are equally vulnerable and important, it is best to divide the system into homogeneous segments. Any segment size deviation will increase the overall system pipe isolation risk.

5.3 System Risk with Respect to the Segment Size Distribution

Equation 5.6 reveals that there are three major factors that govern the level of system risk upon a pipe isolation: the average segment size $\bar{x}$, segment size variations b_j, and segment-wise pipe risk $\bar{f_j}\bar{l_j}$. The existence of malfunctioning valves also changes the quantity of the three governing factors, influencing the overall system pipe isolation risk. The level of change for all three terms depends on both the quantity and position of the failed valves. The following sections explores the relationships between valve conditions and the three risk governing factors. We start with developing a model that relates isolation segment size properties (mean and variance) with the damaged degree of valves (valve failure rates) first.

Due to the vast amount, obtaining up-to-date field conditions of all valves in a WDN is impractical. In such a case, positions of the failed valves are often assumed as random [84]. Consequently, only the number of failed valves, determined by the overall valve failure rate, is essential. To explore the relationship between number of failed valves and isolation segment size properties, simulations need to be conducted due to the complexity of WDN topologies. Since only segment size properties are explored in this part of the work, only the isolation segment size distribution of a damaged network is simulated. The hydraulics behavior of the network will be considered separately in Section 5.4.

Isolation segment size properties on a grid network

The relationship between valve failure rate and isolation segment size properties (mean and variance) is illustrated by a toy lattice (grid) network, shown in Figure 5.3. The network contains 9 junctions, 12 pipes and 24 valves. At the initial state, every pipe of the network has

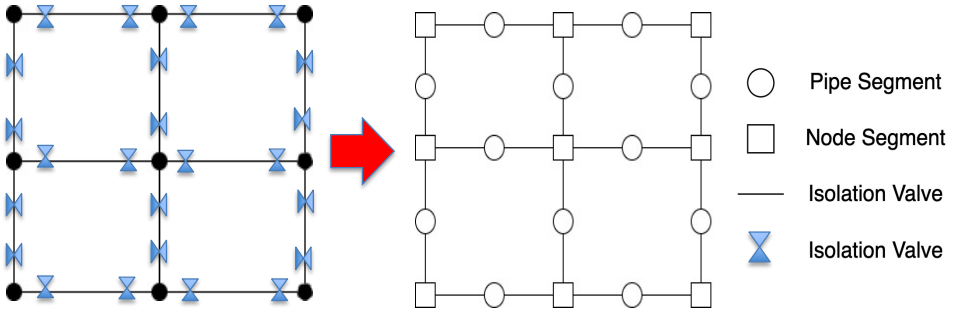


Figure 5.3: A simple grid network in Junction-Pipe (left) and Segment-Valve representation (right)

two functional valves installed at each end. Therefore, every node and pipe can be isolated individually by closing the corresponding valves. In the Segment-Valve representation, every segment will contain either a junction (node) or a pipe. The network composition will change as valves start to fail, as shown in Figure 5.4. If only a few valves failed (12.5% valve failure rate, 3 failed valves), the resulting Segment-Valve graph will contain three segments that have both pipes and nodes (Phase A). However, it is unlikely for a segment to have multiple pipes (large segment) due to valve redundancy. Hence the variance of segment sizes is close to 0 at this phase. As the valve failure rate increases to 25% (6 failed valves), the valve redundancy is depleted. Hence, some segments may contain multiple pipes (Phase B). The mean and variance of segment sizes will increase at this phase, but the increment level will not be significant as single-pipe segments dominates the network. As the valve condition continues to deteriorate (50% valve failure rate), the network condition reaches Phase C. At this stage, single-pipe segments no longer dominate the network. Hence, merging different types of segments is highly likely. The increasing trend for both mean and variance of segment sizes will speed up in this phase comparing to the previous phases. Also, the number of multi-pipe segments stays stable in this phase as it is equally likely to merge two multi-pipe segments (decrease the number of the multi-pipe segment by one) and merge two single-pipe segments (increase the number of the multi-pipe segment by one). When the valve failure rate reaches 62.5%, the system enters a critical state (Phase D). Multi-pipe segments dominate the network in this phase, so a valve failure event is very likely to trigger the merge of two segments with multiple pipes, which dramatically increases the mean and variance of segment size of the network. The number of multi-pipe segment decreases in this phase as well. If most valves fail in the network, only a few multi-pipe segments will remain (Phase E). When a remaining valve fails in this phase, the mean and std of segment sizes continues to increase. The statistics of the described process for the toy network is

summarized in Table 5.1.

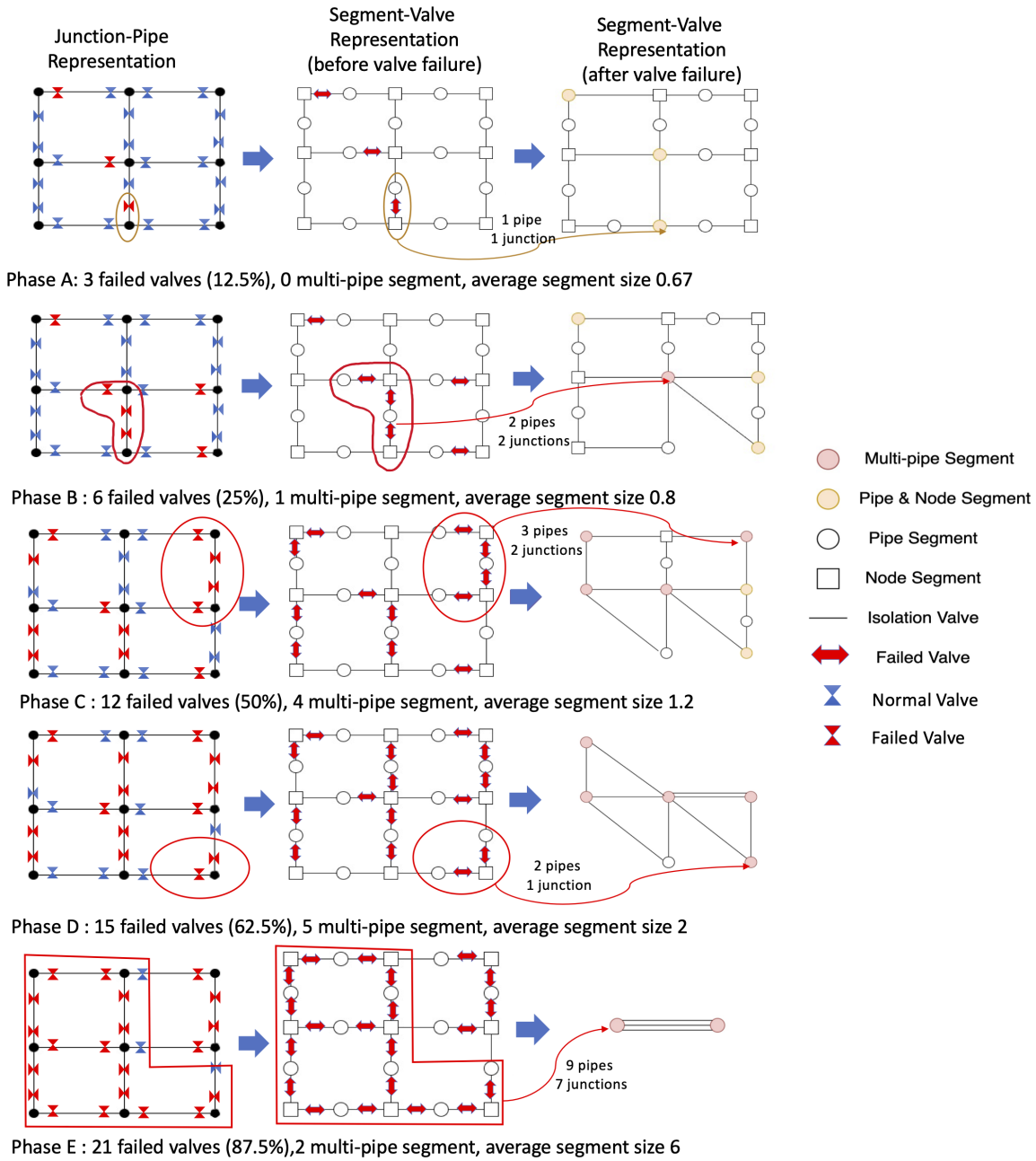
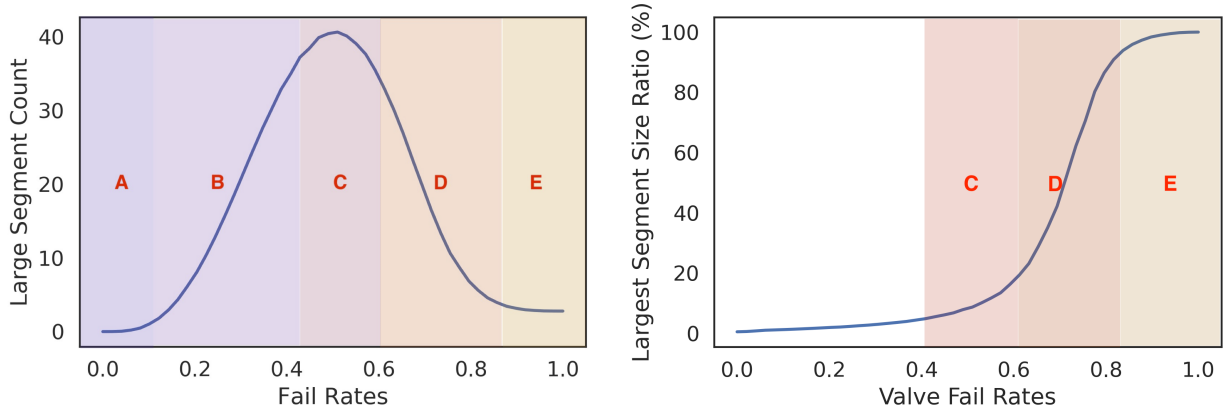


Figure 5.4: Illustration of the phase change phenomena of the toy network (Figure 5.3) as the number of failed valves increases

Table 5.1: Summary of important statistics for the toy network as valve failure rate increases

Phase	Valve failure rate	Number of multi-pipe segments	Average segment size	Segment size standard deviations (STD)
A	12.5%	0	0.67	0.47
B	25%	1	0.8	0.57
C	50%	4	1.2	0.78
D	62.5%	5	2	0.81
E	87.5%	2	6	3.5



(a) Multi-pipe segment count for the grid networks

(b) Size proportion of the largest segment in the grid network

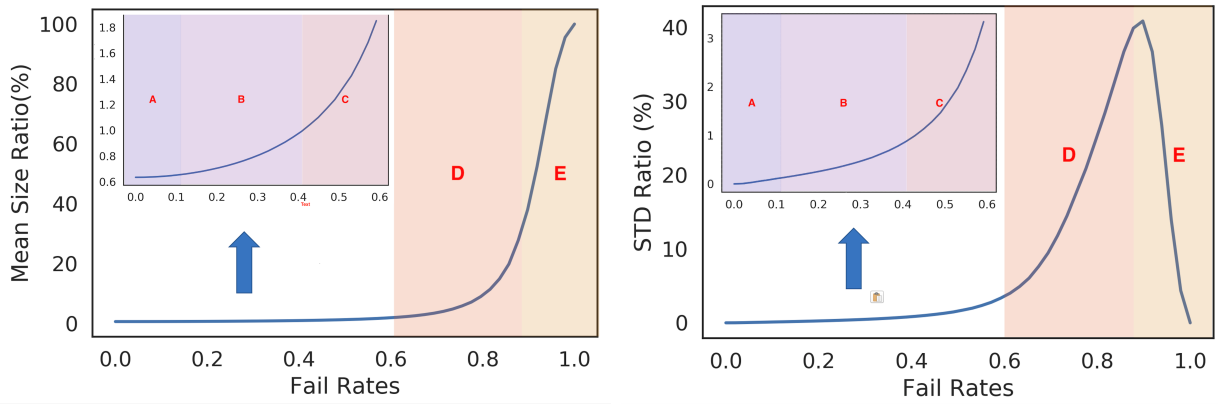
Figure 5.5: Changes of isolation segment properties with respect to different valve failure rates for a grid network

To validate the five-phase network dynamics shown by the simple network, simulation on a large size network (15 by 15 grid network) is conducted on various valve failure rates. Since valves at different locations can fail at each valve failure rate, Monte Carlo (MC) simulation is adopted in this study. Specifically, for each valve failure rate scenario, many realizations of the valve failure events are simulated. For each realization, we randomly fail the desired number of valves (calculated from the given valve failure rate) on the network and analyze the resulting network properties using the Segment-Valve graph. The number of segments that contain multiple pipes, the size of the biggest segment, and the mean and variance of segment sizes across the network are recorded for each realization. All network property measures (mean and variance of segment sizes etc.) under a given valve failure rate scenarios are estimated using the corresponding mean values across realizations. The simulation starts with a fixed size warm-up simulation round (100 realizations for each valve failure scenario). After the warm-up round, if the mean value of all the measures with additional simulations is within 2 percent difference compared with those without additional simulations, the Monte Carlo simulation is terminated.

Figure 5.5a shows the relationship between the valve failure rate to the number of isolation segments that contain multiple pipes inside the network. The previously discussed five phases patterns explain the observed bell shape curve. Initially, the number of multi-pipe segments stays small (Phase A) due to the valve redundancy. It starts to increase as redundancy is depleted (Phase B) and reaches the equilibrium state at around 40% valve failure rates (Phase C). The number decreases dramatically in Phase D when large-size segments dominate the network and stabilize in Phase E, where only a few segments are left.

Figure 5.5b further verifies the proposed five-phases mechanism, which shows the relationship between the size of the largest isolation segment (the segment that contains the most pipes) and the valve failure rate for the simulated network. The largest segment size is insignificant when the valve failure rate is below 40%, confirming that the network is mainly composed of small size segments in this stage (Phase A and B). It starts to rise considerably after passing the 40% valve failure rates, the threshold that the network begins to merge large segments with small segments after a valve failure event. The largest segment size increases dramatically after passing the 60% point, validating the hypothesis that large-segment merging is the most common scheme in this phase (Phase D). Lastly, a Giant Component [100], which contains most of the elements in the network (above 90% of all pipes in this case), appears in Phase E.

The proposed five phase mechanism also fits the observed patterns on both mean and standard deviation plots, shown in Figure 5.6. The mean value and standard deviation of the segment size are close to 0 due to the valve redundancy in phase A. Both measures start to increase slightly in phase B. The small increasing rate is because the network is dominated by a small (single pipe) segment in this phase. However, the growing speed gradually increases in phase C as more and more segments are turned into multi-pipe segments. Both measures increase dramatically in phase D, where large segments dominate the network. In this phase, any valve failure events will likely cause the merge of two large segments, dramatically increasing both the mean value and standard deviation of segment sizes. In phase E, the



(a) Mean value of segment sizes

(b) Standard deviation of segment sizes

Figure 5.6: Isolation segment size properties with respect to different valve failure rates on densely connected networks. The results are normalized to percentile (fraction of pipes over the total number of pipes)

network is dominated by a few substantial segments. The mean value increases tremendously as both merging segments have enormous sizes, but the standard deviation decreases since only a few large-size segments are left.

Heterogeneous network generation process

The previous section shows that the relationship between isolation segment size properties (multi-pipe segment count, mean and standard deviation of segment sizes) and the valve failure rate can be characterized by five patterns for a lattice (grid) network. However, real-life networks often tend to have network topologies other than grid. In this section, a method of generating networks with different topologies is developed to approximate real-life WDN topologies. Using the proposed network generation method, the relationship between valve failure rates and segment size properties for networks with different topologies is discussed in the following section.

In this study, network topology is characterized by the number of nodes (junctions) and node connectivity degree distribution (number of linked pipes for each node). Since hydraulic simulation is not required at the current stage, network properties such as pipe length, material, diameter are excluded from the network generation process. The goal of network generation is then to generate a network with a given node connectivity degree distribution (information of number of nodes is included inside the degree distribution). The proposed network generation process can be considered a spatial version of the configuration model in network theory [100]. In the configuration model, the exact degree of each vertex (node) in the network is fixed. The formation of the network is to randomly connect vertices with the constrain of the predefined vertex degrees. One more constraint is added to account

for the spatial nature of WDNs. Based on the fact that shortcuts (a long pipe that connects two distant junctions even though there are water paths between them) rarely exist in a WDN [14], it is assumed that connections can only occur between adjacent vertices.

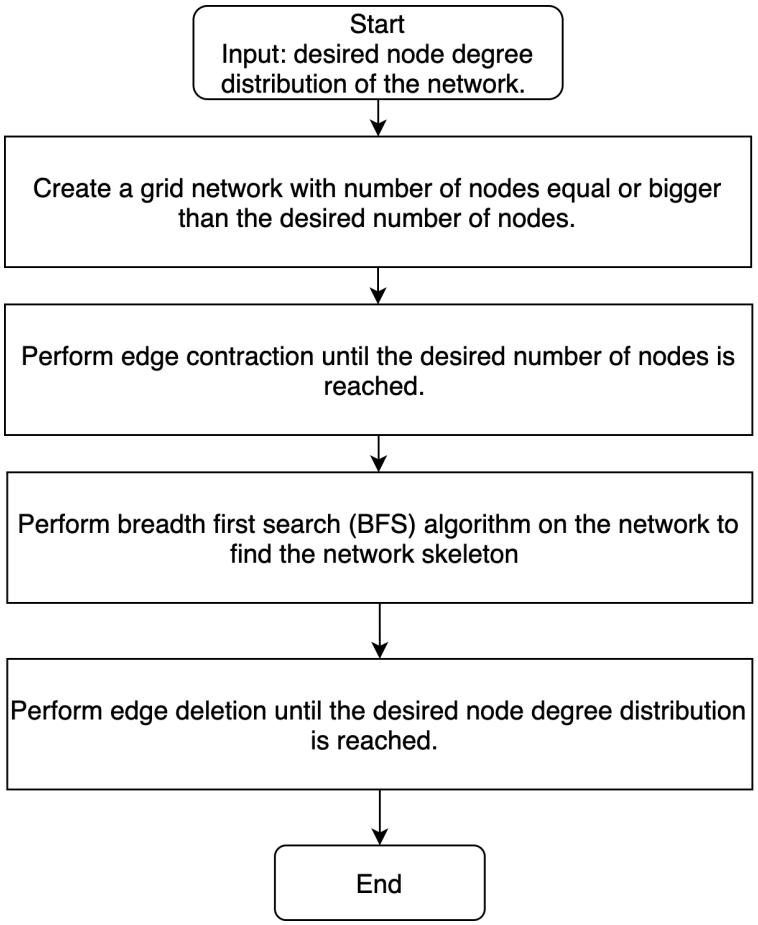


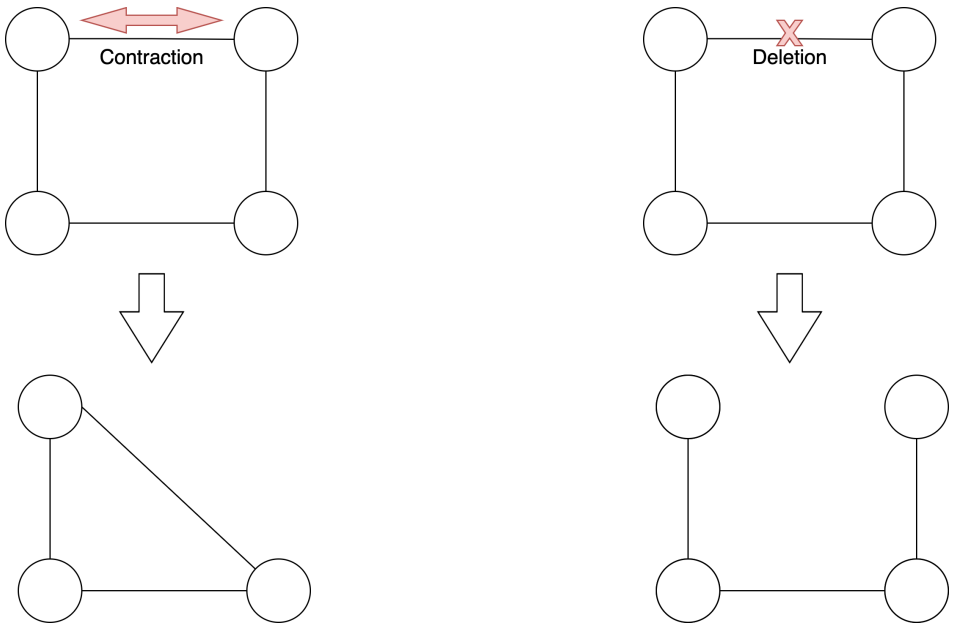
Figure 5.7: Flowchart of the network generation process

The details of the proposed network generation process is summarized in Figure 5.7, which has three major steps:

- Network initialization. A grid network with optimal valve configuration (two valves per pipe) is constructed as the base network. The base network is constructed as a grid because most real WDN networks are less connected than a grid network [14]. Therefore, different WDN topologies (variants) can be approximated by reducing the connectivity from the grid network. Note that the initial network sets the upper limit of the number of nodes and edges for all variants.
- Edge contraction. Performing edge contraction (remove edge and merge the corresponding two segments, see Figure 5.8a) reduces both the number of nodes and edges

from the initial network. It can also create new connection patterns. Edge contraction operations are randomly performed on the base network until the desired number of nodes is achieved.

- Edge deletion. Performing edge deletion (delete a certain edge from the network, see (Figure 5.8b)) on incident edges of a node reduces the node degree (number of incident edges to a node). Networks with a certain node degree distribution can be created by conducting edge deletion on different nodes. To make sure that all the generated networks are connected graphs, a spanning tree is generated for the initial network using the breadth-first search (BFS) algorithm [100]. Edges on the spanning tree will be excluded from the edge deletion operations. In this way, any pair of nodes in the generated graph will have at least one path between them due to the existence of a spanning tree.



(a) Edge contraction operation (b) Edge deletion operation

Figure 5.8: Graph operations used for the network generation process. Edge contraction operation removes the desired edge and merges the two ending nodes. Edge deletion operation removes the desired edge only.

Figure 5.9 shows how an actual WDN topology can be generated using the procedure described above. First, a two-by-four grid network is generated. Then two edge contraction operations are performed to reduce the network to the desired 6 nodes. Edge deletion operations are repeatedly conducted until the network has the same node degree distribution

as the desired one. The generated network is topological equivalent to the original one. Note that the network generation process is not unique.

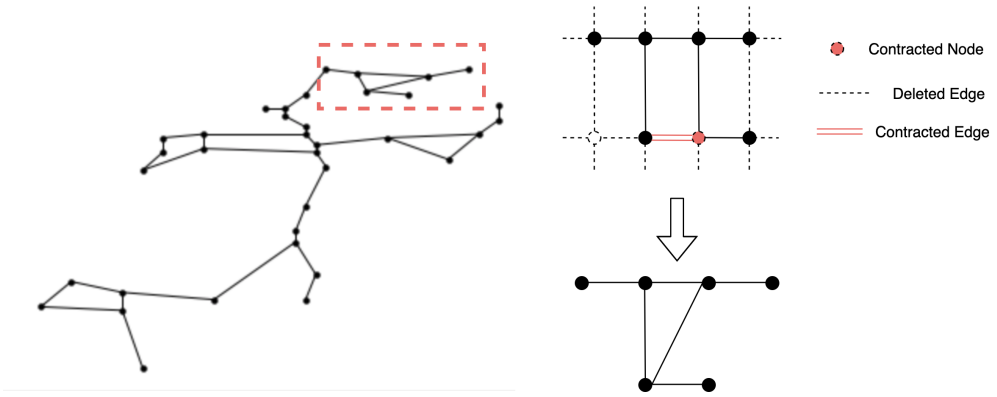


Figure 5.9: Illustration of the network generation process for a real-life WDN

Implementation of edge deletion and contraction on a network is nontrivial since both the network size and connectivity pattern may change after the operations. To perform edge deletion, two valves are set on the deletion pipe to be always closed. In this way, the desired pipe is isolated from the whole system, equivalent to an edge deletion. Closing valves to approximate edge deletion is always feasible because the network is initialized are set with two valves per pipe. To perform edge contraction, the valves on the contraction edge are set to be always open, and mark the contracted edge as the "fictitious" edge. "fictitious" means the edge is conceptual and does not convey any physical meaning. Comparing to directly change the network property (modifying graph adjacency matrix structure etc.), using valve operations to perform edge contraction and deletion is much easier to implement. Opening and closing valves can be directly performed on the Segment-Valve graph without changing the network structure. Furthermore, the modified Segment-Valve graph provides the distribution of isolation segment sizes (mean and variance) of the network, which greatly simplifies the simulation procedure.

Segment size properties on arbitrary heterogeneous networks

In this section, isolation segment properties of networks with different typologies regarding valve failure rate change is investigated. Networks are created using the network generation procedure discussed in the previous section. Specifically, a 15 by 15 grid network (225 nodes, 420 edges) is used as the initial base network. Networks with various node degree distributions are produced using graph operations on the base network. The generated networks are distinguished by the sparseness ratio, which is defined as:

$$S = 1 - \frac{E - E_{branch}}{E_{grid} - E_{branch}} \tag{5.10}$$

where E is the number of edges (pipes) in the network, E_{branch} is the number of edges of a spanning network with the same amount of nodes as the tested network, E_{grid} is the number of edges of a grid network with the same amount of nodes as the tested network. If the network is well-connected (grid-like), the sparseness ratio will be close to 0. On the other hand, if the network is sparsely connected (branch-like), the sparseness ratio will be close to 1.

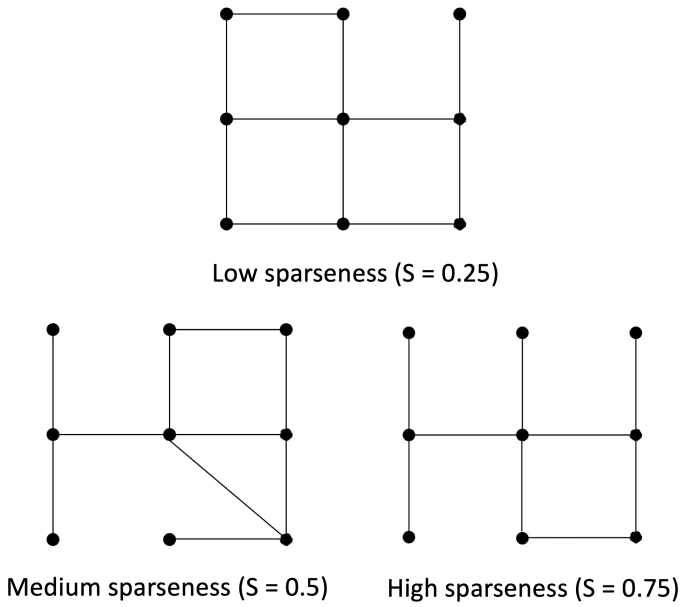


Figure 5.10: Example networks with different sparseness ratios

In this study, 10,000 networks with sparseness ratios range from 0 to 1 are generated. The generated networks are categorized into three groups according to their sparseness ratios to facilitate simulation results visualization and pattern identification. Specifically, networks with a sparseness ratio range from 0 to 0.3 are categorized as low sparseness networks. Networks with a sparseness ratio range from 0.3 to 0.6 are categorized as medium sparseness networks. The other networks are categorized as high sparseness networks ($S = 0.6 - 1.0$.) See Figure 5.10 for examples. The relationship of segment size properties to valve failure rates for different network categories is explored using the same simulation procedure as in the lattice (grid) network. Valves are initialized with the most redundant configuration (two valves per pipe) and randomly choose to fail according to the given valve failure rate. The isolation segment size properties of the degraded networks are estimated using Monte Carlo simulation for each valve failure rate scenario.

The number of large size segments inside the network with respect to valve failure rate is shown in Figure 5.11. Since tested networks have different size, the normalized large segment

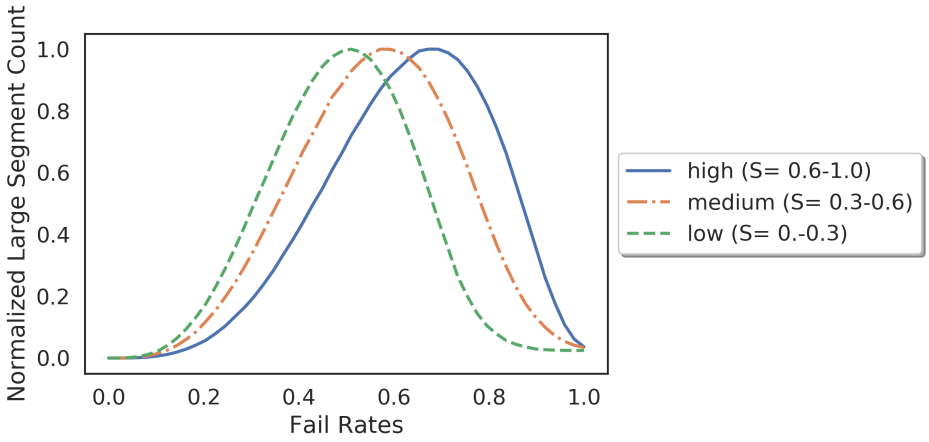
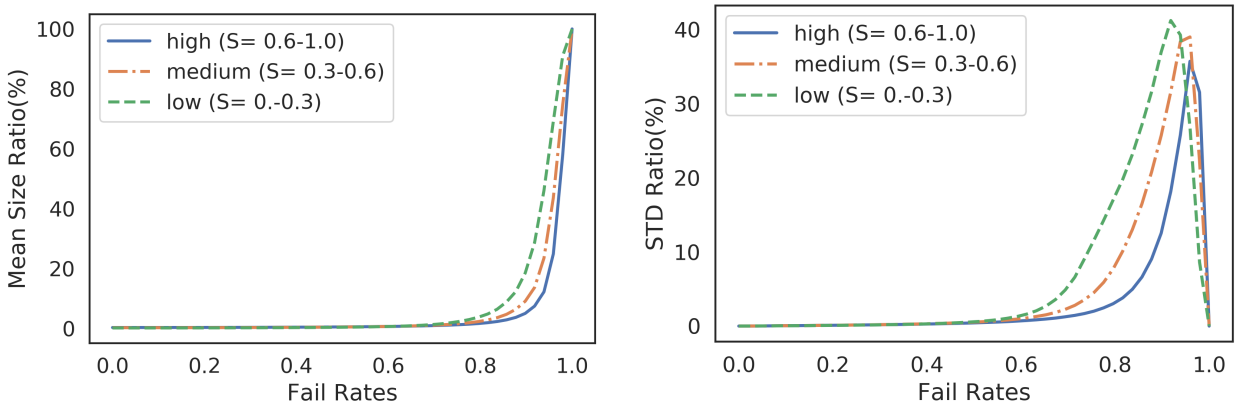


Figure 5.11: Multi-pipe segment count (normalized) for networks with different topology



(a) Mean value of segment sizes

(b) Standard deviation of segment sizes

Figure 5.12: Isolation segment size properties with respect to different valve failure rates. The results are normalized to percentile (fraction of pipes over the total number of pipes)

count is used for visualization purposes, which is defined as:

$$C_{norm} = C/C_{max} \quad \text{at valve failure rate } f \tag{5.11}$$

where C is the number of multi-pipe segments at valve failure rate f, C_{max} is maximum number of multi-pipe segments across all tested valve failure rate.

For all the tested network categories, the relationship between the number of multi-pipe segments and valve failure rates is characterized by the bell shape curve, same as in the lattice network case (Figure 5.6). Therefore, the five phases network behavior observed on the lattice network can be considered as the shared feature for all tested WDN topologies. Although the general shape for all the curves are same, networks with different topologies have different phase change thresholds. Sparse networks have a relatively high phase change threshold (requires a relatively large valve failure rate to change phase). Dense networks have a relatively low phase change threshold (requires a relatively low valve failure rate to change phase).

The differences between phase change thresholds on networks with different topologies can be understood as the following. By definition, networks with high sparsity are relatively less connected than networks with low sparsity. Therefore, segments are sparsely connected in the Segment-Valve graph for sparse networks comparing to dense networks. Because large-size segments are created by merging small segments, it is more challenging to form large size segments for a sparse network than a dense network due to the lack of segment connectivity. However, this does not mean that high sparsity networks are more robust than low sparsity networks. The isolation consequences of a high sparsity network are generally much more significant than those on a low sparsity network due to the high probability of network disconnection. The impact of segment risk distribution on the total system risk will be discussed in Section 5.4.

The mean and standard deviation of segment sizes regarding valve failure rate are shown in Figure 5.12. Because the total number of pipes may be different across networks, both measures are normalized as ratios using the total number of pipes in the network. The two measures show similar behaviors on networks with different topologies. The mean size stays relatively stable when the valve failure rate is low and increases dramatically after a certain threshold. The standard deviation of segment sizes also stays still for low valve failure rate cases but accelerates after passing a threshold. Unlike the mean size, the standard deviation of the segment size converges to 0 when valve failure rate reaches 1 because all pipes are included in one segment when all valves are failed. Note that the phase-changing threshold is also higher for sparse networks than dense networks for the two measures. The observed patterns are the same as those remarked in the lattice case, further validating that the phase-changing mechanism is fundamental for all networks.

5.4 System Risk with Respect to The Isolation Segment Risk Distribution

According to Equation 5.6, in addition to the system-wise segment size properties, the actual distribution of segment-wise isolation risk $\bar{f}_j\bar{l}_j$ is also critical to the quantification of system isolation risk. The segment-wise isolation risk depends on the isolation segment size and location. Isolating a segment that contains multiple pipes (large segment) is often risker than segments with fewer pipes because both pipe failure probability and isolation cost are great for large segments. The location of the isolation segment matters because the cost of isolation is spatially heterogeneous distributed. For instance, isolating a segment at the root of a branch sub-system may disconnect downstream pipes from sources, causing greater damage than isolating segments on densely connected locations. The actual impact scale of the isolation depends on the network topology and the spatial-temporal distribution of the demand, which varies from case to case. Therefore, quantification of segment-wise isolation risk requires hydraulic simulation with information of the isolation segment size and position.

However, both segment size and location are uncertain due to the uncertainties in valve condition (i.e. aging, corrosion, damage by earthquakes etc.). Consequently, the pipe connection pattern of a network is uncertain after segment isolation operations. Since hydraulic simulation requires pipe connection information as the input, it is infeasible to apply it directly to a network with unknown connectivity patterns for risk estimation.

In this study, a method is developed to map the segment-wise risks of the network considering valve condition uncertainties. Specifically, the isolation risk of every single pipe in the network is mapped with a set of predefined isolation segment sizes throughout simulation runs. Figure 5.13 illustrates the procedure for creating the risk map for a given pipe in the network. Before the simulation run, a set of segment sizes that need to be checked must be defined first (e.g. 5, 10, etc). A large size segment implies a large valve failure rate and vice versa. Then for a given segment size, the corresponding isolation segment for the selected pipe is approximated by incrementally adding the closest (minimum topological distance) pipes to the segment until the desired size is reached. The associated pipe component is randomly chosen when there are multiple choices for the closest pipe. Figure 5.14 shows an example of generating segments for a pipe with segment size 5,10,20 using this method. The generated hypothetical isolation segments are then isolated from the system to estimate the isolation risk through hydraulic simulations.

Note that the proposed method does not intend to enumerate all the possible isolation segment scenarios for a given pipe. This is because it is computationally infeasible to list all. For example, a 10 pipes segment may have up to 60,000 different configurations (3^10, assume each pipe connects to 3 other pipes). Instead, it serves as a heuristic for approximating the isolation consequence of different segment sizes for a pipe. The approximation is accurate when close-by pipes have similar risk levels. In such a case, the isolation risk for a segment mainly depends on its size; thus the impact of the exact choice of included pipes is insignificant. Since close-by pipes are often installed around the same period (similar pipe

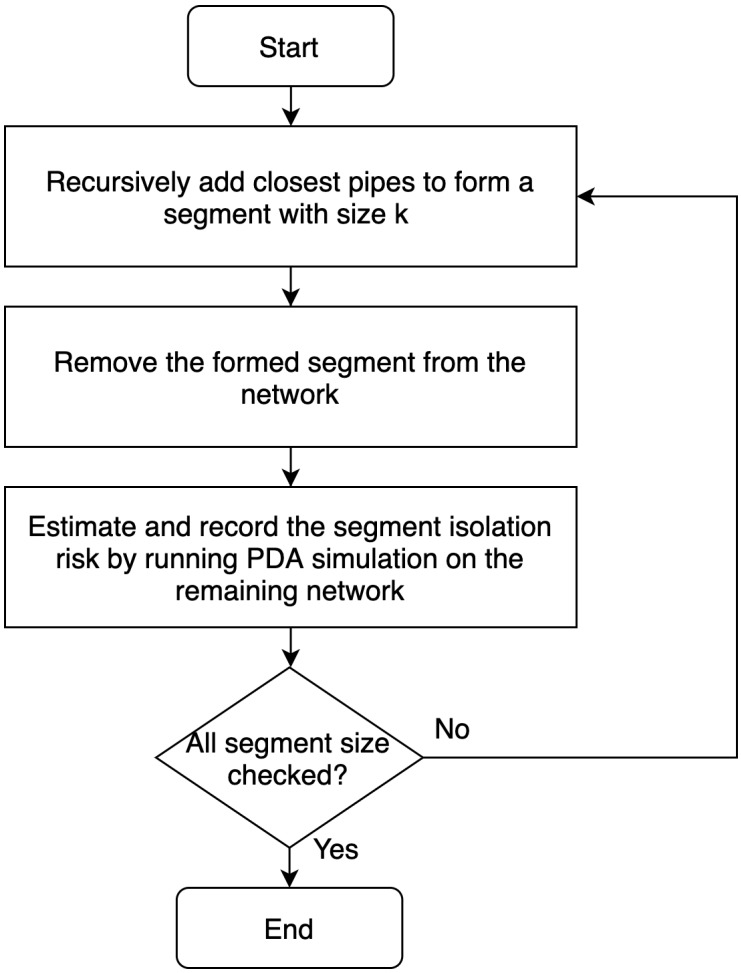


Figure 5.13: Flowchart of the segment isolation risk map generating procedure

material and age), serve the same community (similar demand pattern), and have similar underground conditions, such an assumption is valid in most cases.

A pipe risk map with different segment sizes for the whole system can be created by repeating the single pipe risk evaluation process for all the pipes in the system. The result is called as the "multi-scale pipe risk map". Using the multi-scale pipe risk map, the spatial properties of segment-wise risk distribution can be obtained, which is crucial in evaluating the potential system isolation risk as discussed previously. The importance of the multi-scale pipe risk map will be illustrated in later chapters (Chapter 6) and Chapter 7.

Figure 5.14: Example of generating isolation segments of various sizes for a given pipe

5.5 Isolation Cost

The risk of isolating a segment depends on two factors, the pipe failure probability and isolation cost (see Equation 5.5). In this study, the failure probability of a pipe is estimated as the ratio of the pipe length L_i on the total length of the network L_{tot} ($f_i = \frac{L_i}{L_tot}$). The system-wise demand drop is used as the cost of a pipe isolation operation following [32]. To calculate the system demand drop after the isolation operation, pressure-driven analysis (PDA) is used to simulate water distribution of damaged networks (after subsystem isolation) where some demands may not be met due to the lack of pressure [123]. In the PDA simulation, the

demand-pressure relation is formulated with the following equation [134]:

$$d = \begin{cases} 0 & p \leq P_0 \\ D_f(\frac{p-P_0}{P_f-P_0})^{0.5} & P_0 \leq p \leq P_f \\ D^f & p \geq P_f \end{cases} \tag{5.12}$$

where d is the actual demand, D_f is the desired demand; p is the pressure; P_f is the nominal pressure, which is the pressure threshold that the consumer will receive the requested demand if it is met; P_0 is the minimum pressure, which is the pressure threshold that the consumer will not receive any water if it can not be met.

PDA hydraulic simulation is performed on the initial network (no valve failure) with peak demand distribution to find the actual initial demand distribution using HydrauSim. If the network is configured correctly, all the demands are expected to be met for the undamaged case.

To find the demand cost of isolating a pipe with failed valves, one needs to locate the corresponding isolation segment first. Then all the components inside the isolation segment are removed from the system. Demand distribution after the isolation is reestimated by performing the PDA hydraulic simulation on the remaining network. The system demand cost of the isolation is the summation of the demand loss across all pipes in the network.

The cost of isolating a segment is often further decomposed to indicate different aspects of isolation consequences [32, 51]. In this study, the demand cost is decomposed into two sub-costs. The first sub-cost is the direct cost, which is the cost inside the isolated segment alone. It is the summation of the required demands of all pipes inside the isolated segment because isolated pipes lose all the demands. The other sub-cost is the indirect cost, which is the total cost on the rest of the system after the isolation. In other words, it is the total supply drop over all the pipes that are not part of the isolated segment. Indirect cost exists because isolating one segment may cause water path rerouting across the system, which may increase the water transportation energy loss, causing water pressure and supply drop. In extreme cases, isolation one part of the system may disconnect other parts of the system from the sources, increasing the total isolation cost significantly.

Revealing impact aspects of isolation through cost decomposition provides valuable information for water utilities. For instance, indirect cost implies the surprise level of an isolation operation as it measures the degree of having unintended (outside the isolation segment) demand loss. Moreover, effective risk mitigation procedures can be developed based on the two decomposed risk measures, which will be discussed in Chapter 7.

5.6 Summary

This chapter presents a framework that provides new understanding about the impact of having malfunctioning isolation valves for a WDN. Analyzing WDN using the Segment-Valve representation leads to the formulation of the system level pipe isolation risk (Equation 5.2),

which states that the risk depends on a) average size of segments; b) variance of segment sizes across the network; c) spatial distribution of the isolation cost, which is related to both topological and hydraulic properties of the network.

After identifying the controlling factors for the system risk, their behavior regarding valve failure rates is explored next. A method of generating networks with different topologies is developed to produce inputs for simulation. Simulation results reveal that the changes in the mean and standard deviation of segment sizes obey a distinctive five-phases phenomenon for all tested synthetic networks. Both measures stay small at the low valve failure rates due to the valve redundancy (Phase A). As the valve failure rate increases, redundancy starts to be depleted, driving up both the mean and variance (Phase B) until it reaches the equilibrium state (Phase C). Increasing failure rate further pushes the system out of the equilibrium, resulting in large-size segments dominating the network (Phase D, large mean and variance). Finally, Phase E is reached as almost all valves are failed, where only a few segments are left.

In addition to segment-size properties, this chapter develops a method to map the spatial distribution of segment-wise isolation cost considering valve condition uncertainties. Through simulation runs, the isolation risk of every single pipe in the network is mapped with a set of predefined isolation segment sizes. The proposed method serves as a heuristic for approximating the isolation consequence for a pipe as the valve failure rate (segment size) increases. Areas with risks that are sensitive to segment size changes can be pinpointed for maintenance purposes.

Chapter 6

Pipe Isolation Risks: Case Studies

Two real-life WDNs are tested in this study to demonstrate the validity of the hypothesis proposed in the previous chapter (Chapter 5). The first case study is performed on the EBMUD Alameda Island WDN, a densely connected mesh-like network. To illustrate the impact of network topology on the pipe isolation risk for a system, a sparsely connected network, the EBMUD Round Hill WDN, is used as the other case study. In the following sections, relationships between the valve failure rate and the pipe isolation risk are explored for the two cases, and the results are compared and discussed at the end.

6.1 Alameda Island WDN

Network description

A medium-sized WDN serving Alameda Island, an island in the San Francisco Bay Area, California, United States, is used as the case study for this research (Fig 6.1). The Alameda Island WDN is currently managed by the East Bay Municipal Utility District (EBMUD), a public utility district that provides water and sewage treatment services on the eastern side of San Francisco Bay. The network comprises 2508 internal nodes and 3028 pipes for a total length of around 194 km. 1999 valves are currently installed in the system. The network is fed by two transmission line pipes (source pipes) linked to other parts of the EBMUD network. Following suggestions from the EBMUD modeling team, two reservoirs with fixed heads 45.72 m (150 ft) are connected with the source pipes as water sources. No pumps are used for this service zone as the elevation for this area is low. The system has a total demand (average annualized day average in 2018) of around $15 m^3/min$ (3962.58 gals/min).

Figure 6.2b shows that pipes in the network have various lengths. About half of the pipes are less than 50 ft, and the number of pipes decreases dramatically as pipe length increases. Pipes in the study network are made with different materials, as shown in Figure 6.2b. About 38% of pipes are made with cast iron (CI), 28% made with steel, 25% with asbestos concrete (AC), and 8% with polyvinyl chloride (PVC).

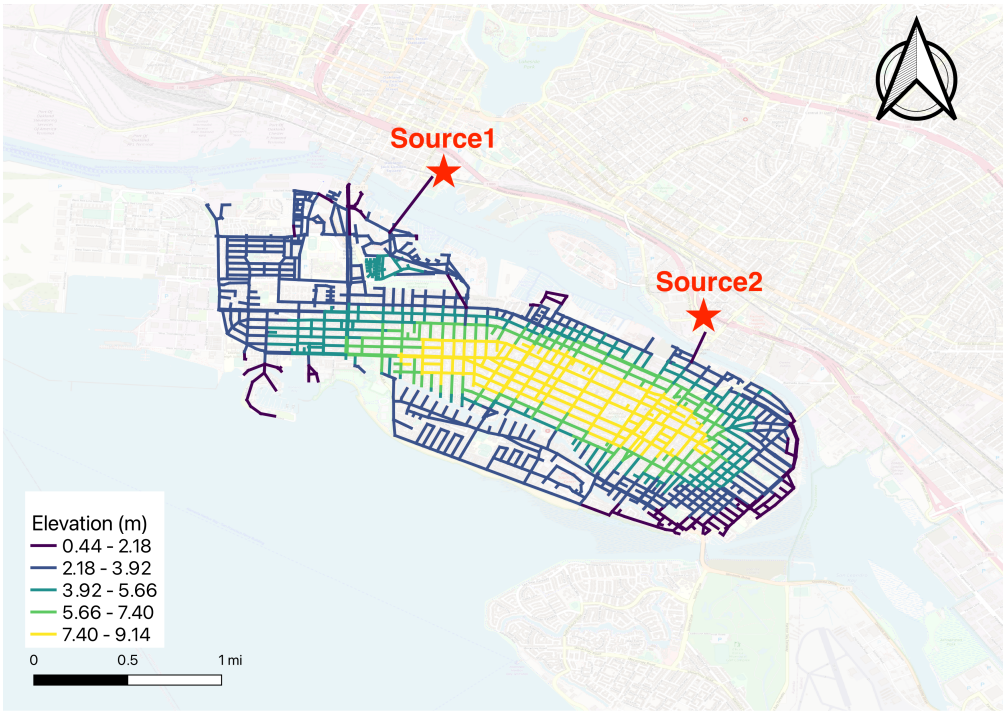
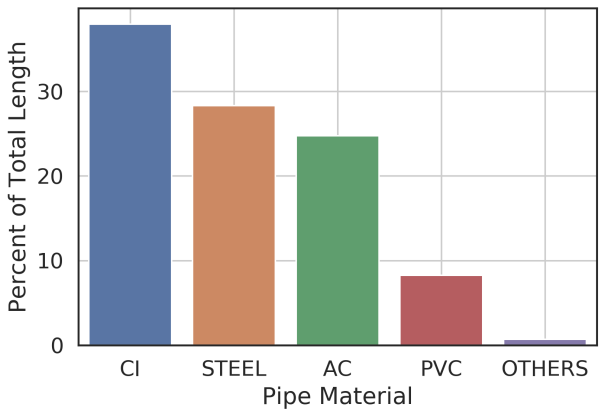
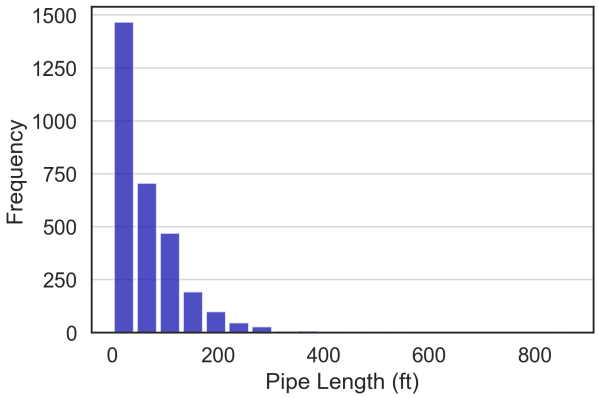


Figure 6.1: Network topology of the Alameda Island WDN



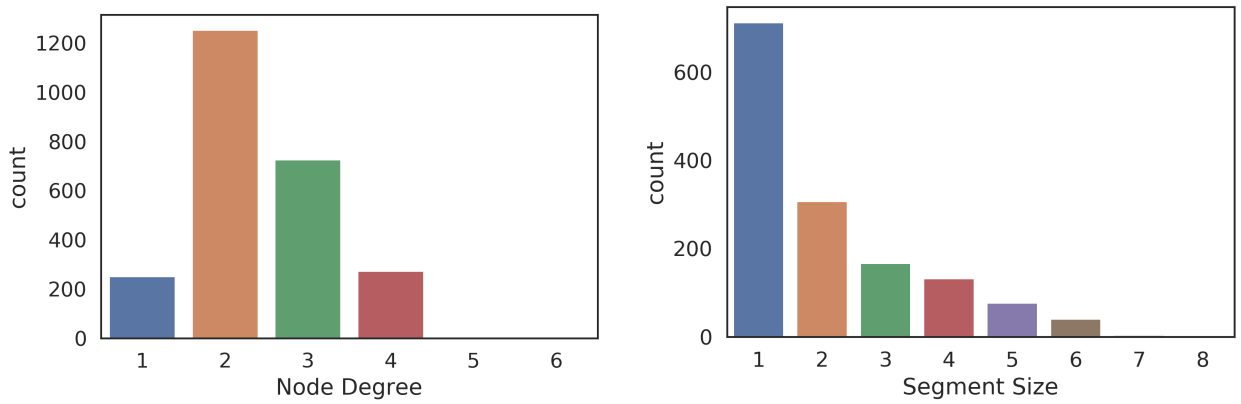
(a) Pipe material distribution



(b) Pipe length distribution

Figure 6.2: Pipe properties of the Alameda Island WDN

The network is mainly composed of well-connected grid-like structures, with a few branch-like sub-structures on the edges of the network. Figure 6.3a shows the node degree (number of connected pipes to a junction) distribution for the network. Most junctions are connected by multiple pipes (node degree > 1), and only a few junctions are linked by a single pipe (end nodes). Such node degree distribution is due to the fact that the network is well connected.



(a) Node degree distribution

(b) Isolation segment size distribution

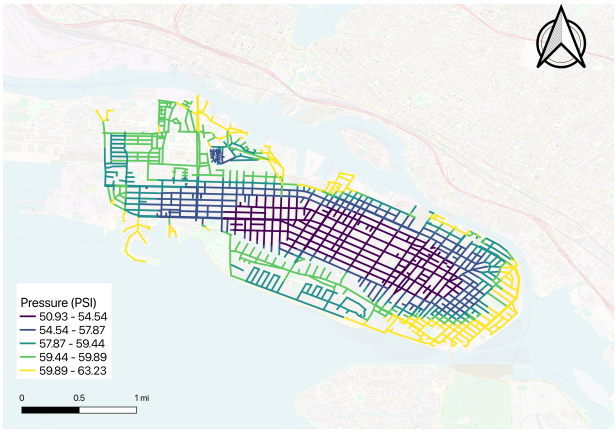
Figure 6.3: Network properties of the Alameda Island WDN

Figure 6.4 shows the pressure distribution across the study area under normal operating conditions. Note that pipe pressure instead of node pressure is used since the study focuses on pipe failure events. It is calculated as the average of the simulated pressure of the two end nodes. Areas located around the network center tend to experience relatively low pressure (around 50 PSI) due to the relatively high elevation. Regions on the edge of the network have slightly higher pressure (around 60 PSI) because of low elevation. Locally, the pressure variance across pipes is not significant (close-by pipes have similar pressure) due to small elevation variances.

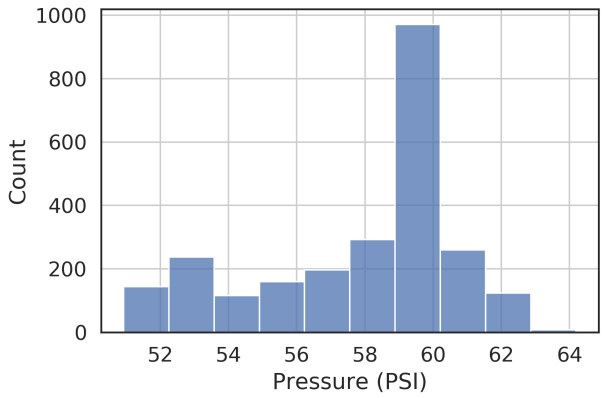
Figure 6.5 shows the demand distribution across the study area under normal operating conditions. Same to the pressure cost, pipe demand instead of node demand is used, which is the average of the simulated demand of the two end nodes. Under the normal operation condition, all the demands can be satisfied based on the hydraulic simulation result. Unlike the pressure distribution, the demand distribution is unevenly distributed. About half of pipes have trivial demand (< 1 GPM), and the other half lays on the range between 1-10 GPM. There are a few pipes that have exceptionally large demands (> 10GPM).

Isolation segment size properties

The effectiveness of the currently deployed valve configuration is explored first, assuming no malfunctioning valves. When all the available valves are closed, the network is divided into

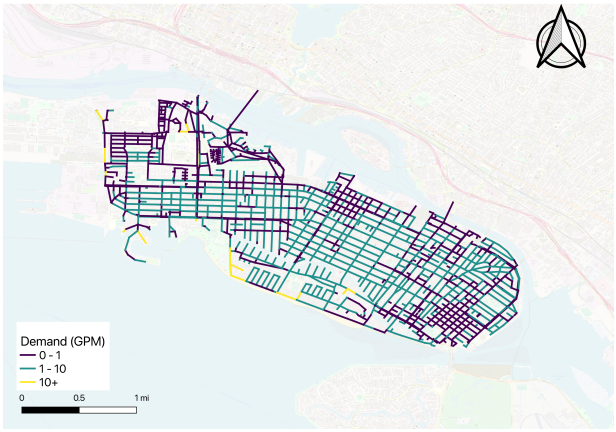


(a) Initial pressure map of the network

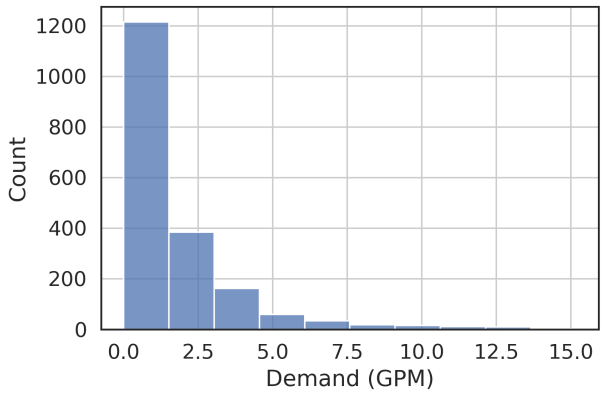


(b) Initial pressure distribution histogram of the network

Figure 6.4: Initial pressure distribution for the Alameda Island WDN under the normal operating condition with no subsystem isolation



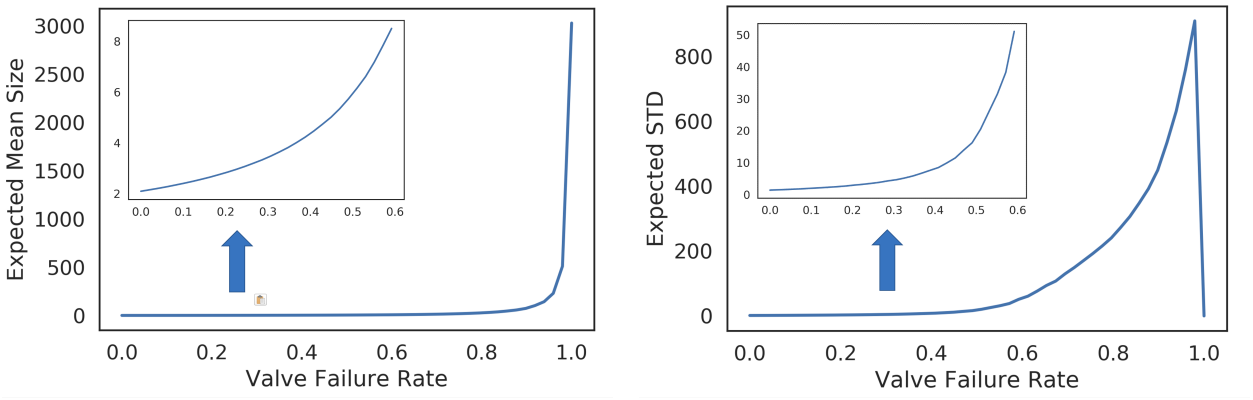
(a) Demand distribution map of the network



(b) Demand distribution histogram of the network

Figure 6.5: Demand distribution for the Alameda Island WDN under the normal operating condition with no subsystem isolation

1443 disjoint isolation segments. The distribution of segment sizes is shown in Fig 6.3b. The average segment size, defined as the number of pipes inside a segment, is 2.1. The variance of segment size is around 2. Almost half of the isolation segments are perfect (size equal to 1), and the number of segments decays as segment size increases. However, there are still non-negligible isolation segments containing multiple pipes (30 percent of the segments contain more than two pipes) since not all pipes have two valves installed.



(a) Average isolation segment size

(b) Isolation segment size variances

Figure 6.6: Isolation segment size properties with respect to different valve failure rates for the Alameda Island network. Results in range 0 to 0.6 are magnified to illustrate details

Isolation segment size properties change as valves start to fail. Figure 6.6 shows the simulation results of the mean and standard deviation of isolation segments in the system with respect to different valve failure rates. The rate of change for both mean and standard deviation (STD) of isolation segment size increases as valve failure rate increase from zero. After the valve failure rate passing 60%, both measures increase dramatically compared to the prior state. For example, raising the valve failure rate from 60% to 80% increasing the mean segment size from 3.7 to 39.1 (956% increase). On the other hand, raising valve failure from 10% to 30% only increases the mean segment size from 2.2 to 2.6 (18% increase). In other words, the network becomes exceptionally unreliable when the valve failure rate is above the 60% threshold. In the 0% to 60% range, it is clear that segment size STD increases dramatically after valve failure rate pass 30%. Similarly, the segment size mean value accelerates after the 30% threshold. Both contributes to increasing the risk as discussed earlier.

Figure 6.7 shows the simulation result of the number of the large segment in the network with respect to different valve failure rates. Comparing to Figure 5.5a, Phase A and B are missing in Figure 6.7 because the current deployed valving configuration is much sparser than the ideal configuration, where every pipe has two valves. Since segment size measures (mean and standard deviation) stays close to 0 in both phase A and B due to valve redundancy (see Figure 5.5a), missing the two phases implies that the case study network has no "buffer

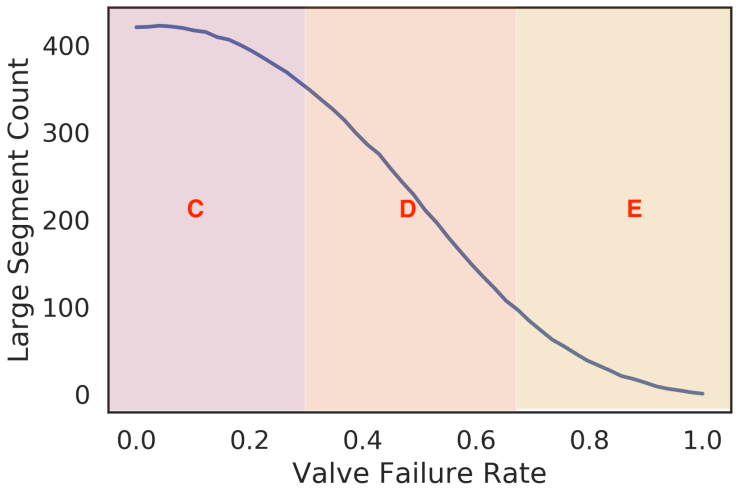


Figure 6.7: Number of multi-pipe segments with respect to different valve failure rates for the Alameda Island WDN

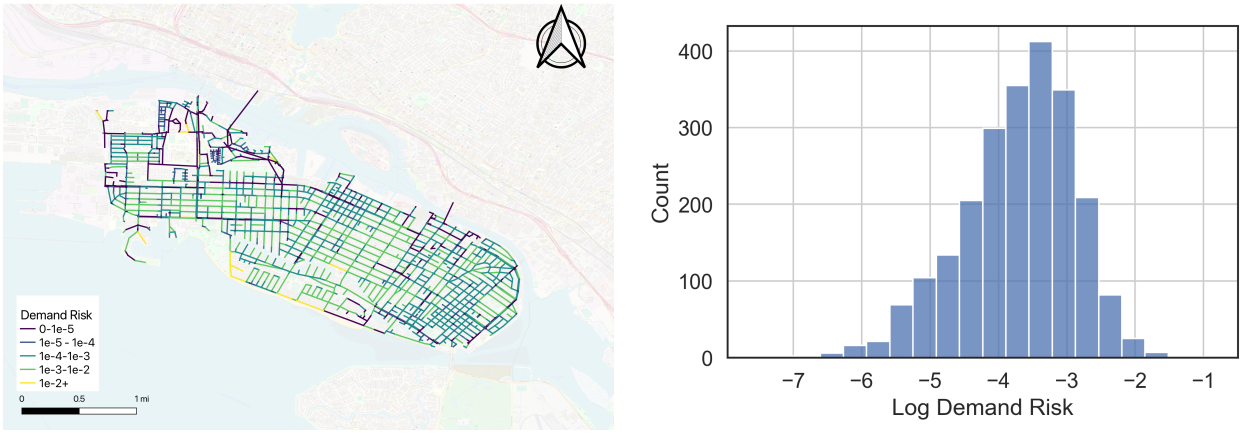
zone," where isolation risk stays small due to small segment size properties. The large segment counts remain relatively stable at first (Phase C) then transits to Phase D after around 30% of valves are failed in the network. In Phase D, large size segments dominate the network. Therefore, each valve failure event causes a merge of two large segments, decreasing the number of the large segment by one. When around 75% are failed, the network reaches Phase E. In this case, only a few huge segments remain, so the rate of change for large segment count decreases.

Segment-wise pipe isolation risk distribution

Figure 6.8 shows the isolation demand risk (direct risk) of each pipe, assuming each pipe can be isolated individually. Figure 6.8a illustrates the spatial pattern of pipe risk of the WDN. Pipes located at the central part of the system tend to have more risk comparing to peripheral pipes. Figure 6.8b shows the histogram of pipe demand risk. One key observation in Figure 6.8 is that demand risk varies over many magnitudes (from 10^{-6} to 10^{-2}) in the network. The significant variation is due to the fact that both demand (see Figure 6.5) and pipe length varies significantly (see Figure 4.6b) across the network pipes.

The multi-scale pipe isolation risk map described in Section 5.4 is generated for the Alameda Island WDN to illustrate the spatial distribution of segment-wise pipe isolation risk with different segment sizes. Specifically, segment sizes range from 1 to 30 are used to create the risk maps for the studied WDN.

Figure 6.9 shows the simulation result for both direct demand risk and indirect demand risk. Pipes with high demand risk (bigger than the total averaged per pipe risk) when



(a) The initial demand risk distribution

(b) Histogram of the initial demand risk distribution in log scale

Figure 6.8: Initial demand risk distribution for the Alameda Island WDN. Note that pipes with no demand (zero risk) are excluded for visualization purposes

isolating with small segment size (0-5 pipes) are marked in red (first-degree vulnerable pipes). Pipes with high risk only when the corresponding isolation segment size is large (5-30 pipes) are marked as yellow (second-degree vulnerable pipes). Pipes that have no risk regardless of tested segment sizes are colored in black.

Spatial patterns can be observed from the risk maps. For the direct demand risk map (Figure 6.9a), long pipes are more likely to be classified as first-degree vulnerable pipes due to relatively large pipe failure rates. There are very few second-degree vulnerable pipes in the direct risk map. This is because the segment-wise isolation risk is an aggregated measure ($\bar{f}_j\bar{l}_j$ in Equation 5.6), which remains stable if individual pipe risks inside the segment are similar. Hence, second-degree pipes can only appear in areas with high individual pipe risk differences, which are rare for the studied WDN.

Figure 6.9b shows the indirect demand risk map for the Alameda Island WDN. Only a few parts of the system have high indirect risk when segment size is small (first-degree vulnerable pipes). Even when the tested isolation segment size becomes large, the number of pipes with significant indirect risk (second-degree vulnerable pipes) is still small compared to the indirect risk insensitive pipes (marked as black). Although the total amount is small, indirect-risk sensitive pipes show spatial patterns. First, they tend to cluster around branch-like subsystems. Since pipes in the branch-like subsystem are connected to water sources through a few critical pipes (the root of the branch network etc.), cutting the water feed to such critical pipes will likely starve the whole branch. In other words, critical pipes of branch subsystems are first-degree vulnerable pipes. As the size of the isolation segment increases, these critical pipes might be included in the isolation segments of surrounding pipes, creating indirect demand losses. Thus, pipes that are closed to first-degree vulnerable

(a) Direct demand risk map

(b) Indirect demand risk map

Figure 6.9: Direct and indirect demand risk map of the Alameda Island WDN

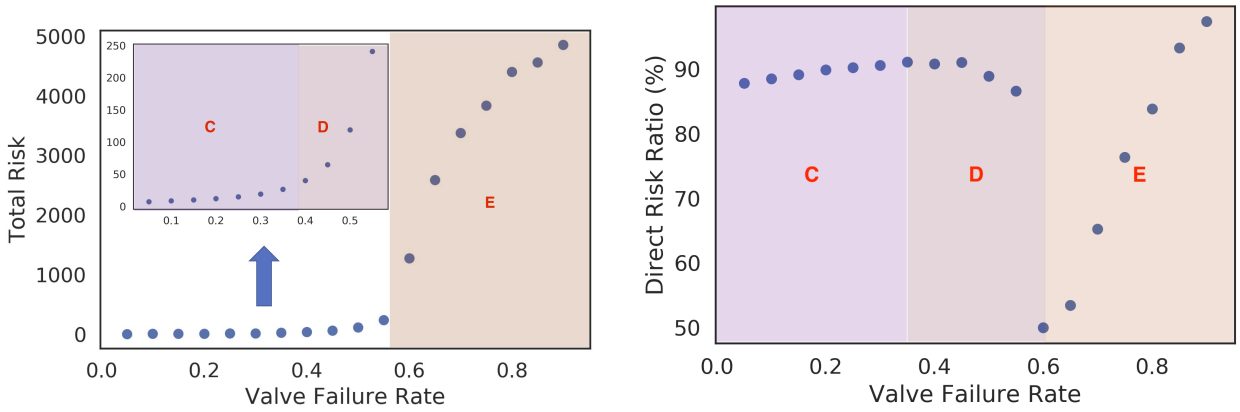
pipes are likely to become second-degree vulnerable pipes. Another vulnerable region for indirect risk is around water sources. As the isolation segment size increase, isolating pipes around the source pipes tend to limit the available number of water pathways to the rest of the network, causing a significant system-wise pressure and demand drop.

The final important observation from Figure 6.9b is that, although certain parts of the system do not have indirect demand risk when the isolation segment size is small, the risk emerges as segment size exceeds a specific threshold (see lower edge of the network). In other words, some areas of the WDN contain only second-degree vulnerable pipes. Such phenomena are due to the network connectivity pattern in these areas. Specifically, the network is moderated connected in such regions. They do not have critical root pipes like branched networks, but the number of water paths is limited due to the relative lack of connectivity compared to grid networks. When the isolation segment is small, there is no indirect pipe isolation risk due to the water path redundancy. As isolation segment size increases, it becomes more likely that some pipes will lose connection to water sources, creating indirect risk in the region.

System pipe isolation risk simulation

The hydraulic behavior of the case study network with different valve failure rates is simulated to verify the proposed hypothesis from previous discussions. Because any valve can fail in the system, Monte Carlo(MC) method is used to deal with uncertainties on which valve to fail. The MC pipe isolation risk estimation process is similar to the MC simulation used for isolation segment size quantification, discussed in Section 5.3. For a given valve failure rate scenario, risk calculation processes are repeatedly conducted with different selections on

failed valves. The system isolation risk under a given valve failure rate is estimated as the mean value from the MC results pool, with uncertainties captured by the standard deviation. The number of the MC simulation is determined during the simulation process from [123]. The simulation starts with a fixed size warm-up simulation round (100 simulations in this case). After the warm-up round, if the mean and variance of the estimated risk with additional simulations are within 2 percent difference compared with those without additional simulations, the Monte Carlo simulation is terminated.



(a) Demand risk of pipe isolation

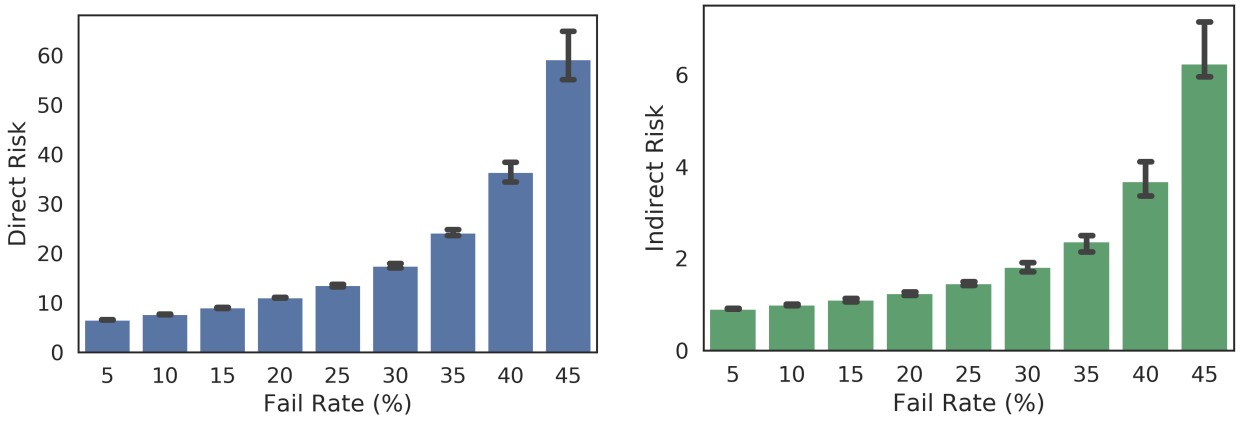
(b) Proportion of the direct risk to the total demand risk

Figure 6.10: Risk of pipe isolation for the Alameda Island network under various valve failure rate conditions

Figure 6.10a shows the simulation result for the total demand risk of pipe isolation with pipe failure rate. The pattern of the simulated demand risk with respect to valve failure rates matches the phase change behavior observed on segment size properties (Figure 6.7). The risk increases relatively slowly when the valve failure rate is lower than 35% (Phase C). After passing the 35% threshold, the system pipe isolation risk accelerates (Phase D). The increasing trend of the risk slows down when the most valves are failed in the system (Phase E).

The isolation risk composition, measured as the ratio of direct risk to the total risk, regarding various valve failure rates is shown in Figure 6.10b. The same phase-change mechanism discussed previously can also explain the observed patterns. The proportion of direct risk increases at the beginning stage when the valve failure rate is lower than 35% (Phase C). In Phase C, the Segment-Valve graph is composed of segments with various sizes (see Section 5.3). A valve failure event merges two segments with different sizes in this phase; thus, the average direct risk increases as the valve failure rate increase. Since there are no dominant segments (segment contains many pipes), it is unlikely to disconnect portions of the system with single segment isolation. Therefore, the indirect risk of the system remains

relatively stable. Because direct risk increases yet indirect risk remains almost the same, the ratio of direct risk increases as the valve failure rate increases in this phase. The direct risk ratio drops as the system enters Phase D. In this phase, large segments dominate the Segment-Valve graph. Consequently, it becomes more likely to disconnect the other parts of the system after the isolation than Phase C. In other words, the rate of change for indirect risk speeds up, resulting in a declining direct risk ratio as the system enters Phase E. A sharp discontinuity of the direct risk ratio is observed. An isolation segment that contains a large portion of system pipes, first appears at the beginning of Phase E. Isolating the pipes in the segments other than this large segment disconnects the large segment from the network, causing large demand loss. Therefore, the proportion of direct risk decreases sharply at the beginning of Phase E. As the valve failure rate increases, more and more small segments get included in the large segment, reducing the sources of indirect demand risk. Consequently, the proportion of direct risk recovers rapidly in Phase E.



(a) Direct demand risk

(b) Indirect demand risk

Figure 6.11: System demand risk under different valve failure rates

In real-life situations, it is unlikely to have an valve failure rate larger than 50%, which suggests at least 1 out of 2 valves is failed. Both direct and indirect risk are plotted in 5-50% valve failure rate range to illustrate the details of risk change, as shown in Figure 6.11. Both direct and indirect risk increase slowly when the valve failure rate is lower than the 35% threshold, and the increasing rate speeds up afterward. The indirect demand risk increases at a slower rate than the direct demand risk. For example, the direct demand risk increases from 6.5 GPM to 9.02 GPM (a 39% increase) when the valve failure rate increases from 5% to 15%, On the other hand, the indirect demand risk increases about 20% in the same range. The differences in the rate of change for the two risk measures can be explained using the multi-scale pipe risk maps described in the previous section. From Figure 6.9, it is clear that the amount of risk-sensitive pipes (first-degree vulnerable pipes) is larger for the direct risk comparing to the indirect risk. Therefore, it is more likely to fail valves on pipes with large

direct risk than large indirect risk, causing a relatively rapid change for the system direct risk.

The other observation is that the variance of the indirect demand risk is larger than the variance of the direct demand risk, but the differences are insignificant. For instance, the coefficient of variation for direct risk is about 14% when the valve failure rate is 45%, which is not far from the 16% estimation on the indirect risk at the same valve failure rate scenarios. The insignificant variance differences are because both direct and indirect demand risks are non-evenly distributed due to the variations of demand amount (cost), pipe length (pipe failure rate), and network topology.

6.2 Round Hill WDN

Network description

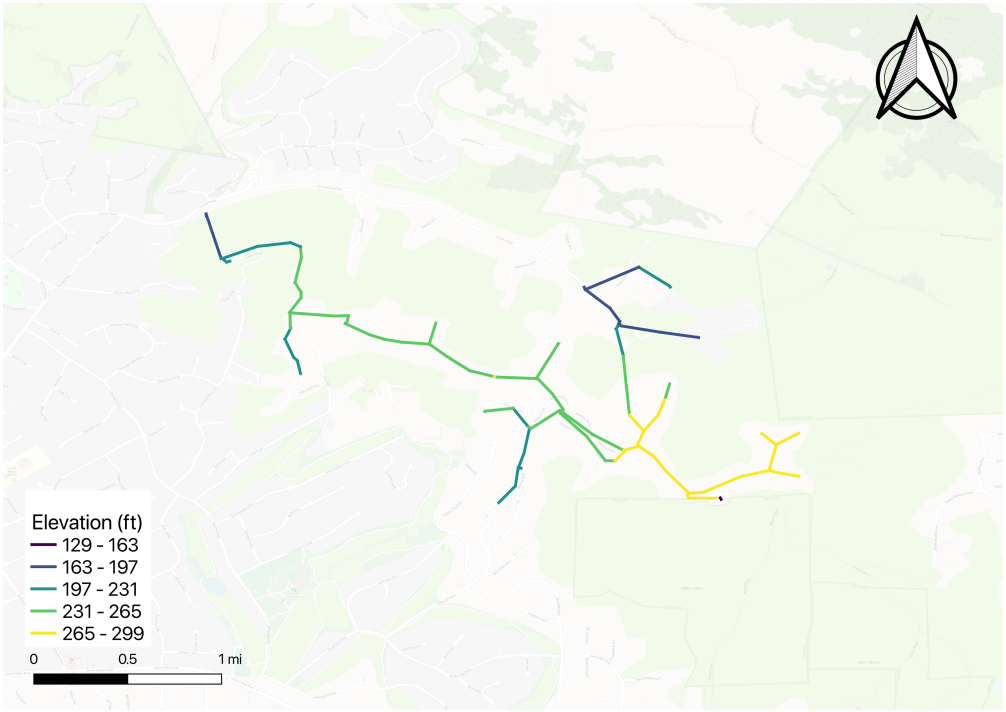
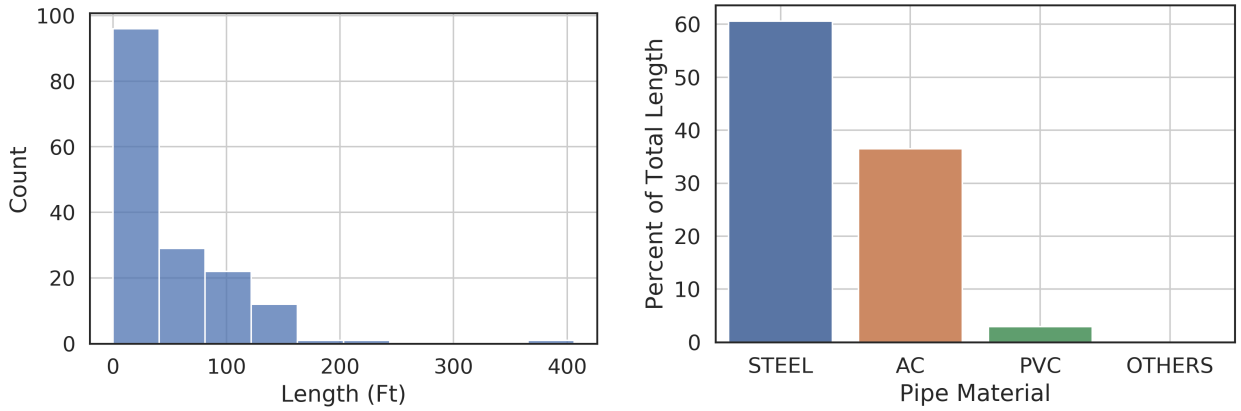


Figure 6.12: Network topology of the Round Hill WDN

A small size branch-like WDN, the Round Hill WDN, is used as a case study to illustrate the network topology impact on the system isolation risk. The Round Hill WDN, managed by the East Bay Municipal Utility District (EBMUD), serves a community in the East San Francisco Bay Area (City of Alamo). The topology of the WDN is shown in Figure 6.12. The network comprises 126 internal nodes and 126 pipes for a total length of around 2.1 km.

Currently, there are 48 valves installed in the system. The network is supplied by a reservoir with fixed head 997.9 ft. The system has a total demand (average annualized day average in 2018) of around 60 gals/min.



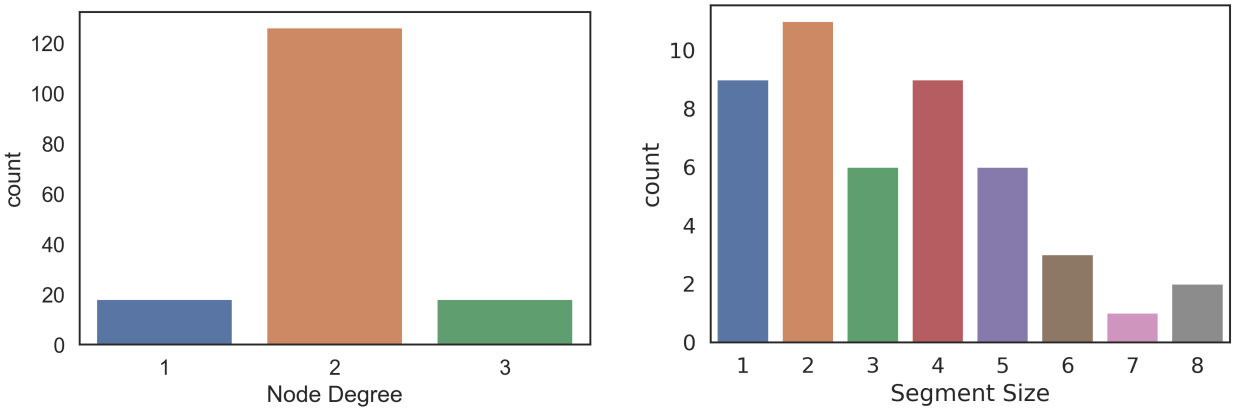
(a) Pipe length distribution

(b) Pipe material distribution

Figure 6.13: Pipe properties of the Round Hill WDN

Figure 6.13a shows the length distribution of system pipes. The majority of the pipes are less than 150 ft. There are a few exceptionally long pipes (longer than 200 ft) for water transmission purposes. Figure 6.13b shows the material distribution of the pipes. The majority of the network is made with steel (61%) and asbestos concrete (36%). Only a few pipes are made with PVC (3%). The node degree (number of connected pipes to a junction) distribution of the network is shown in Figure 6.14a. Since the overall network structure is branched (sparsely connected or tree like), it is not surprising to observe that most junctions are connected by two pipes (node degree equal to 1). Only a few three-degree junctions exist, and no junctions have more than three pipes connected to them.

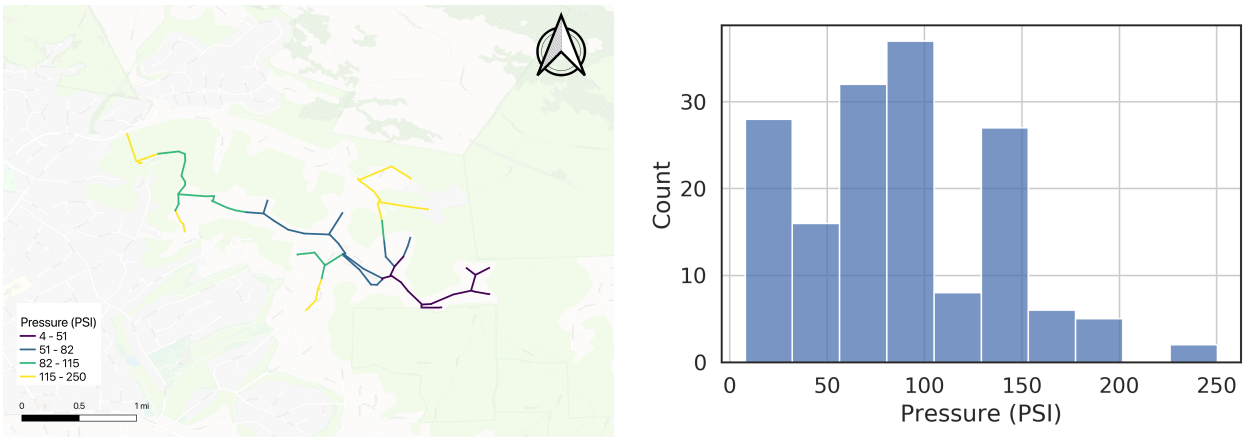
Figure 6.15 shows the pressure distribution of network pipes under normal operating conditions. All the demands can be satisfied based on the simulation result. Since the elevation variation for the study area is large (see Figure 6.12), the pipe pressure is distributed in a relatively wide range. The pipes that are close to the source, due to the relatively high elevation, tend to experience relatively low pressure. Conversely, the pipes located at the edge of the network experience high pressure because of low elevation. Figure 6.16 shows the demand distribution across the study area under the normal operating conditions. About 40% of pipes have very small demand (< 0.5 GPM). The demand for the rest of the pipes ranges from 0.5 to 2.0 GPM. Only a few pipes have demands that are greater than 2.0 GPM.



(a) Node degree distribution

(b) Isolation segment size distribution

Figure 6.14: Network properties of the Round Hill WDN



(a) Initial pressure distribution map

(b) Initial pressure distribution histogram

Figure 6.15: Initial pressure distribution for the Round Hill WDN under the normal operating condition (with no subsystem isolation)

(a) Demand distribution map

(b) Demand distribution histogram

Figure 6.16: Demand distribution for the Round Hill WDN under the normal operating condition (with no subsystem isolation)

Segment size properties

When all the available valves (assuming no valve malfunctioning) are closed, the network can be divided into 47 disjoint isolation segments. Figure 6.14b shows the sizes distribution of isolation segments in the network. The average segment size (number of pipes inside the segment) is 3.3, with a standard deviation of 1.89. Unlike the Alameda Island network, where the majority of isolation segments are small, there is no dominant isolation segment size for the Round Hill network. Only 20% of isolation segments are perfect (contains only one pipe), whereas 25% contain more than four pipes.

The number of the large segment in the network regarding different valve failure rates are simulated and plotted, shown in Figure 6.17. Similar to the case of Alameda Island network (Figure 6.7), Phase A and B (buffer zones) are missing in Figure 6.17 due to valve inadequacy. Therefore, it is expected that any valve failure event will likely increase the averaged segment size, driving up the total system isolation risk. After 70% of the valves are failed, the system enters phase D, where large size segments dominate the network. Phase E is not observed from the simulation because the resolution of the valve failure rate used in the simulation is too coarse. Since there are only 48 valves installed in the system, the smallest valve failure rate resolution is around 2%. On the other hand, the range of Phase E is expected to be small for a sparsely connected WDN like the current one (see Figure 5.11). Therefore, it is unsurprising that the small range phase E is overlooked from the simulation due to the relatively coarse valve failure rate resolution.

Isolation segment size properties (mean and standard deviation) regarding various valve failure rates are simulated and plotted in Figure 6.18. The rate of changes for both mean and standard deviation (STD) of isolation segment size increase as the valve failure rate

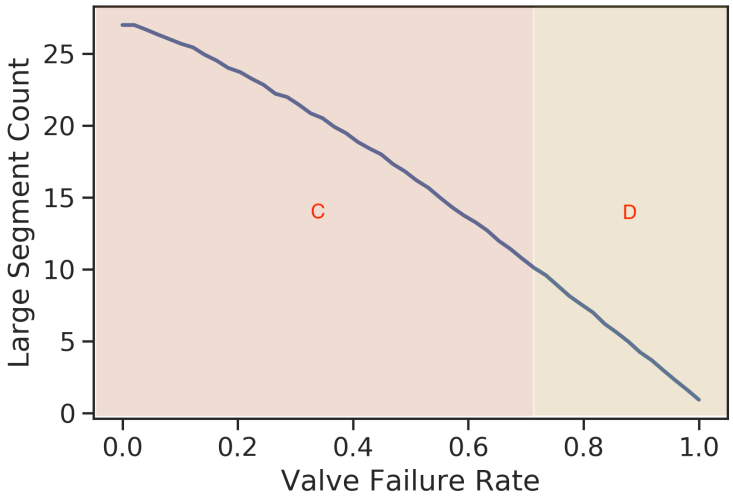
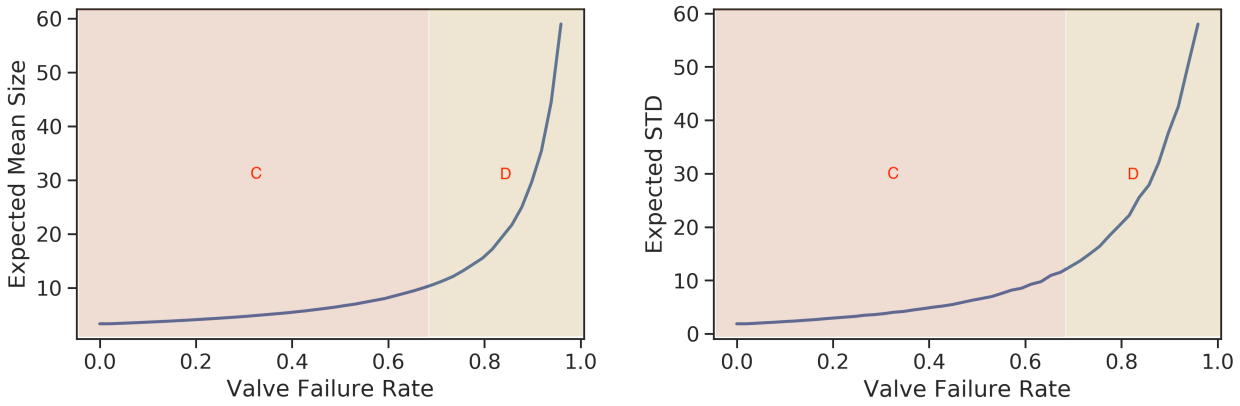


Figure 6.17: Number of multi-pipe segments with respect to different valve failure rates for the Round Hill WDN

increases. The overall increasing pattern matches the phases observed from Figure 6.17. Both measures increase steadily before the 70% threshold (Phase C) and rising dramatically after it (Phase D). For instance, raising the valve failure rate from 20% to 30% increasing the mean segment size from 3.76 to 3.94 (4.5% increase). On the other hand, raising valve failure from 80% to 90% increases the mean segment size from 28.93 to 44.13 (52.5% increase).



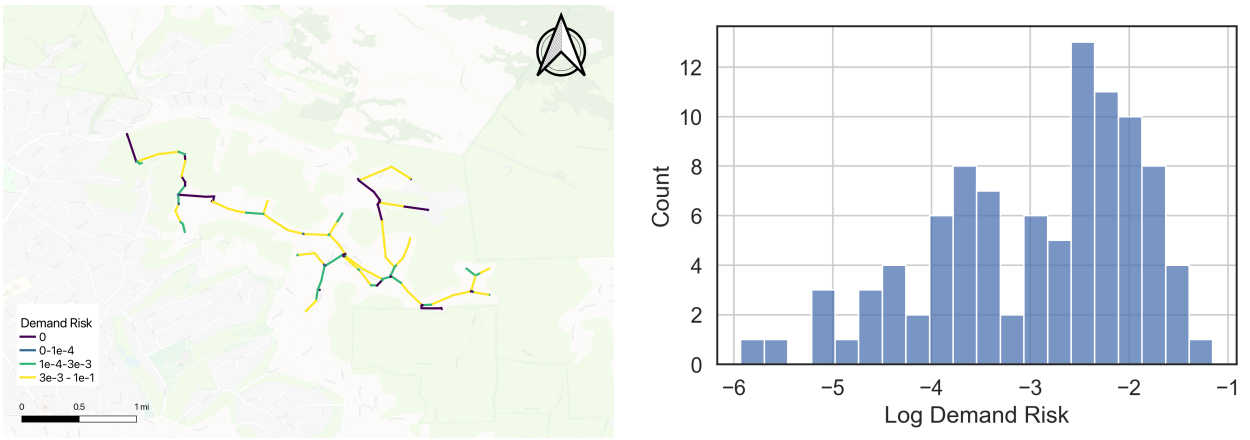
(a) Average isolation segment size

(b) Isolation segment size standard deviation (STD)

Figure 6.18: Isolation segment size properties with respect to different valve failure rates for the Round Hill WDN

Segment-wise pipe isolation risk distribution

Figure 6.19 shows the initial pipe isolation risk map (direct risk), assuming every pipe can be isolated individually. Figure 6.19a illustrates the spatial pattern of pipe risk of the WDN. Unlike the Alameda Island case, no significant patterns can be concluded. Figure 6.19b shows the histogram of pipe demand risk. Similar to the Alameda Island case, the isolation risk for each pipe varies significantly (from 10^{-6} to 10^{-2}) due to the uneven demand (Figure 6.16) and pipe length distribution (Figure 6.13a).



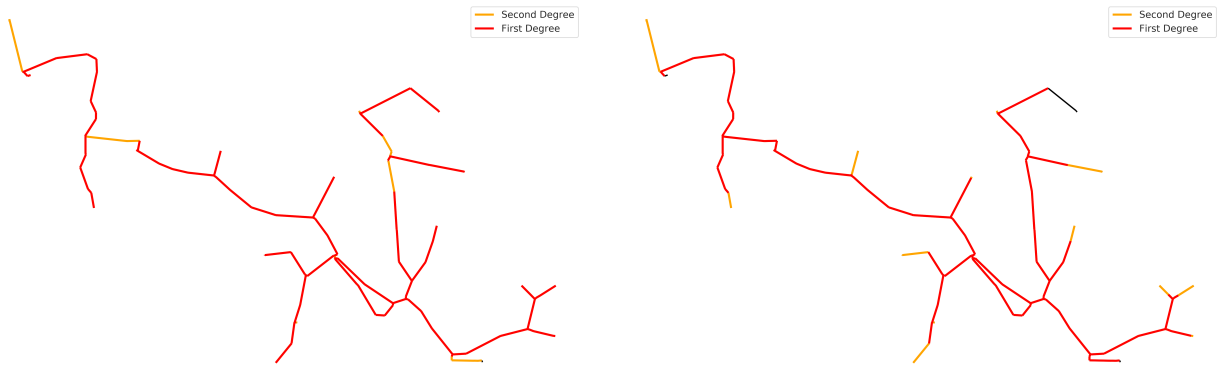
(a) The initial demand risk distribution

(b) Histogram of the the initial demand risk distribution

Figure 6.19: Initial demand risk distribution for the Round Hill WDN. Note that pipes with no demand (zero risk) are excluded for visualization purposes

The multi-scale pipe isolation risk map for both direct demand risk and indirect demand risk is generated, as shown in Figure 6.20. Since the network size is small, isolation segment size 1-15 is used instead of the 1-30 range used for the Alameda Island case. Pipes with high isolation risk, even when isolated with a small segment size (0-3 pipes), are marked in red (first-degree vulnerable pipes). Pipes with high isolation risk only when the corresponding segment size is large (3-15) are marked as yellow (second-degree vulnerable pipes). All other pipes are colored in black. Due to the lack of connectivity of the network, almost all the pipes are vulnerable to isolation for both direct and indirect demand risk measures. For the direct risk map shown in Figure 6.20a, only a few pipes with no demand associated with them are exempted from the first-degree vulnerable pipes. However, as the isolation segment size increases due to failed valves, the chance of including nearby pipes that have demand increases when isolating no demand pipes. Thus, all pipes are vulnerable to isolation when the valve condition is uncertain.

Similar patterns can be observed from the indirect demand risk map (Figure 6.20b. First, like the direct demand risk map, almost all the pipes are vulnerable to isolation due to the



(a) Direct demand risk map

(b) Indirect demand risk map

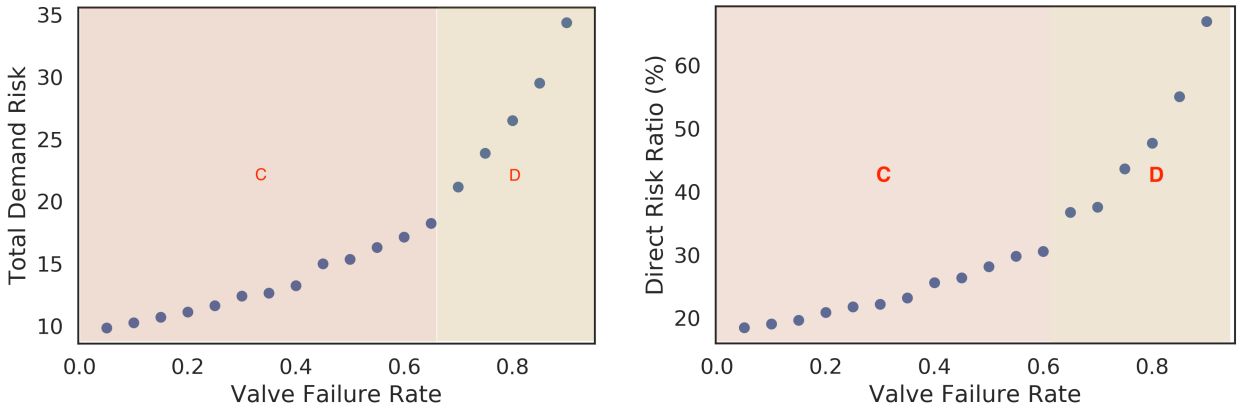
Figure 6.20: Direct and indirect demand risk map of the Round Hill WDN

lack of connectivity of the network under valve condition uncertainties. Isolating the pipes located at the central part of the network will disconnect a portion of the network from the only water source. Therefore, these central pipes are classified as first-degree vulnerable pipes. On the other hand, isolating the pipes around the edge of the network will not cause any indirect risk when the isolation segment size is small (no other pipes to disconnect). However, as valve conditions deteriorate, isolating the edge pipes may include some central pipes, creating indirect demand loss for the rest of the system.

System pipe isolation risk simulation

The hydraulic behavior of the Round Hill WDN under various valve failure rates is simulated to verify the proposed hypothesis from the previous discussion. The same MC simulation approach from the Alameda Island case study is used to deal with valve condition uncertainties. Figure 6.21a shows the simulation result for the total demand risk of pipe isolation with respect to various valve failure rates. Overall, the simulated risk behavior matches the phase change pattern observed on segment size properties (Figure 6.17). The risk increases relatively slowly when the valve failure rate is lower than 70% (Phase C). The risk change rate speeds up after passing the 70% threshold (Phase D).

Figure 6.21b shows the pipe isolation risk composition of the network regarding different valve failure rates. Indirect demand risk is the major contributor to the total demand risk when the valve failure rate is low. For instance, indirect risk accounts for over 80% of the total system isolation risk at the 10% valve failure rate level. The large proportion of indirect risk is due to the lack of pipe connectivity of the network. As shown in Figure 6.20, isolating a pipe even when the valve failure rate is low (small segment size) is likely to disconnect some parts of the network from the system, creating a large indirect isolation risk. As the valve failure rate increases, the average segment size increases as well. In other



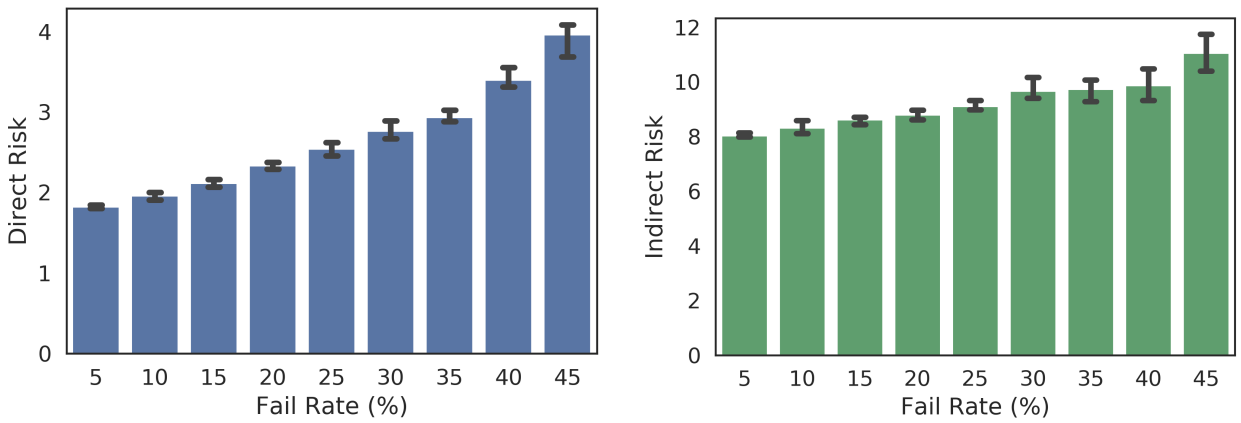
(a) Total pipe isolation demand risk

(b) Proportion of the direct demand risk to the total demand risk

Figure 6.21: System pipe isolation demand risks under different valve failure rates for the Round Hill WDN

words, more pipes are included inside a segment, driving up the direct isolation risk of the system. Therefore, the proportion of direct demand risk increases as the valve failure rate increases. The ratio of direct risk to the total isolation risk accelerates after passing the 70% valve failure rate threshold. In this phase (Phase D), the number of pipes inside each isolation segment increases rapidly, as shown in Figure 6.18. At the same time, the indirect risk does not increase as quickly as the direct risk, as the network is already disconnected even when the valve failure rate is low.

Figure 6.22 shows the demand risk change under the practical valve failure rate range (5%-50%). The direct demand risk increases steadily as the valve failure rate increases due to the rise in both mean and variance of the isolation segment size (see Figure 6.18). However, there is no dramatic acceleration of the direct risk since the studied valve failure range lies on the same phase (Phase C). On the other hand, Figure 6.22b shows that the indirect risk slightly increases as the valve failure rate increase. The lack of change is due to the sparsely connective nature of the Round Hill WDN. As shown in Figure 6.20, the majority of the pipes (first-degree vulnerable pipes) have indirect isolation risk even when the isolation segment size is small. Increasing their segment size will not increase their indirect risk because the other parts of the system are already disconnected. However, not all pipes have fixed indirect risk. Indirect risk of the second-degree vulnerable pipes, located at the network's edges, tend to increase as valve failure rates rise. Such pipes do not have indirect risk when isolating alone. But as the valve failure rate increases, their isolation segment might contain pipes that have indirect risk, driving up the total indirect risk. Because the number of second-degree vulnerable pipes is less than the number of first-degree vulnerable pipes, the overall system indirect risk increases slightly over the tested valve failure rate range.



(a) Direct demand risk

(b) Indirect demand risk

Figure 6.22: System demand risk variation under variuos valve failure rates for the Round Hill WDN

6.3 Discussion and Conclusion

This chapter conducted simulation on real-life WDNs to validate the novel five-phase network dynamics proposed in the previous chapter. In both case studies, the relationship between system pipe isolation demand risk and valve failure rates matches the proposed phase-change pattern. When the valve failure rate is relatively low, the system pipe isolation demand risk increases relatively slowly (Phase A, B and C). However, as the valve failure rate increases to the state that large-size segments dominate the Segment-Valve graph, the demand risk dramatically accelerates as additional valves fail (Phase D). The increasing trend of isolation risk slows down as the Segment-Valve graph is dominated by a major segment that contains most of the pipes (Phase E).

For networks with different typologies, the thresholds for phase changes are different. Specifically, compared to the densely connected Alameda Island WDN, the Round Hill WDN has a higher phase change threshold due to the sparsely connected topology. In other words, segment size properties (mean and variance) of the Round Hill WDN increase slower compared to the case in the Alameda Island WDN. Consequently, the total risk change level for the Round Hill WDN is smaller than the Alameda Island WDN for a same valve failure range. Specifically, the total demand risk increases from 10 to 66, a 560% increase for the Alameda Island WDN when valve failure rate increases from 5% to 45%. On the other hand, demand risk increases only by 60% for the Round Hill network on the same valve failure range.

Although the total demand risk of the sparsely connected network, Round Hill WDN, increases relatively slower than the densely connected Alameda Island WDN, it does not mean that Round Hill WDN is more robust. On the contrary, it is much more vulnerable.

In real-life situations, it is unlikely to have exceptional large valve failure rates (bigger than 50%). Hence, the average isolation segment size will be small for all networks regardless of topology features (assuming a reasonable amount of valves are deployed) under such conditions. Small segment size means that direct risk is locally constrained with a limited quantity. Therefore, the risk of having unintended isolation, reflected by the indirect risk, outweighs direct isolation risk under such conditions.

Due to the sparsely connected nature, each pipe in Round Hill WDN has limited water paths to the water source. Therefore, isolating a pipe in the Round Hill WDN may disconnect many other pipes from the source, causing high indirect costs. Hence, the indirect risk is the major contributor (around 80%) to the total risk when the valve failure rate is low (Figure 6.22b). On the other hand, water paths are abundant for most pipes in the Alameda Island WDN because of the densely connected topology. Consequently, the indirect risk is only a small part of the total risk (around 10%) when the valve failure rate is low (Figure 6.10b). Therefore, the Round Hill WDN is much more vulnerable to pipe isolation comparing to the Alameda Island WDN due to the large indirect isolation risk when valve failure rate is small.

However, as the valve failure rate becomes extremely large (e.g. over 50%), the densely connected network becomes more vulnerable to pipe isolation than the sparsely connected network. As discussed previously, isolation segment sizes increase faster for a dense network than for a sparse network. Consequently, both direct and indirect risk increase faster for a dense network than for a sparse network. Therefore, a dense network eventually becomes riskier than a sparse network as valve failure rate increase, especially at the stage that large-size segments dominate the Segment-Valve graph (Phase D and E)

Chapter 7

Valve Placement Strategy for Minimal System Pipe Isolation Risk

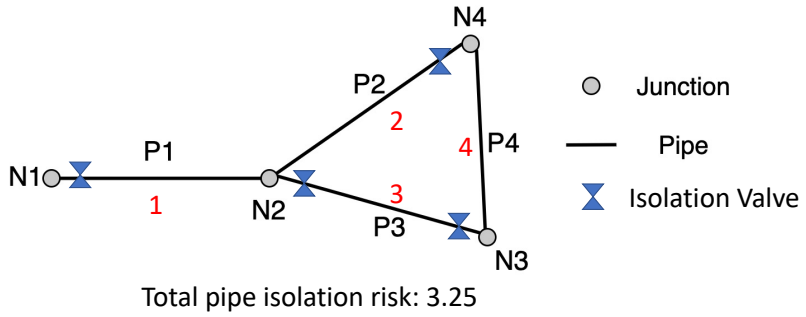
7.1 Overview

The previous chapters introduce a way of analyzing the pipe isolation risk under isolation valve uncertainties for a WDN. The proposed risk formulation gives insights about how the quantity (number of valves), quality (valve fail rate), and configuration (placement positions) of valves influence the system-wise pipe isolation risk. This chapter focuses on using such understandings to improve the resilience of a WDN regarding valve failure events. Specifically, an optimal valve placement algorithm is proposed using dynamic programming techniques. The proposed algorithm is intuitive and straightforward, yet the solution is proved to be both optimal and computationally feasible. Moreover, it provides theoretical insights on the effectiveness and robustness of some widely used rule of thumb isolation valve placement strategies. Code for the developed algorithm can be found in `https://github.com/rewu1993/wdn_valves`.

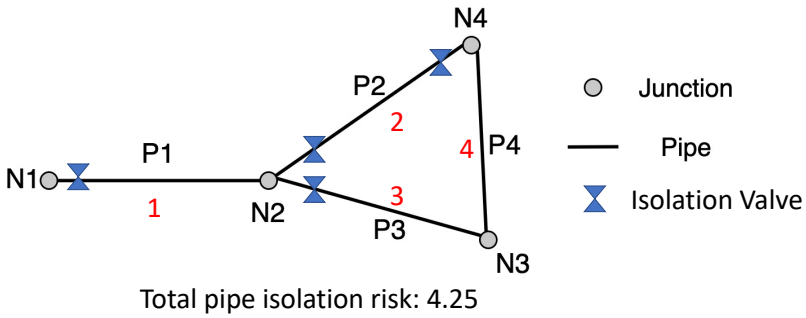
7.2 Optimal Valve Placement Strategy

As discussed in the previous chapter (Chapter 5), in addition to the quantity, the spatial locations of functioning valves also impact the pipe isolation risk of a WDN. In other words, the total isolation risk for a WDN might be different under different valve placement configurations, even though the number of functioning valves is the same. Considering Figure 7.1 for an example. Suppose there are four valves available to be installed for the network. Every pipe has an equal chance of failure (25%), and the cost (demand) for each pipe is marked as red in the figure. Under the valve configuration of Figure 7.1a, the total system pipe isolation risk is $(0.25 + 0.25) * (1 + 3) + 0.25 * (3 + 4) = 3.25$. However, if the valve that is installed close to junction N3 (on pipe P3) is moved to junction N2 (on pipe P2), the total system pipe isolation risk will increase to $(0.25 + 0.25) * (4 + 3) + 0.25 * (1 + 2) = 4.25$. The

pipe isolation risk for the system increased by around 30% from the replacement of one valve. Hence, finding a valve placement configuration that minimizes the system pipe isolation risk for a given number of valves is critical to maximizing the benefits of valve investment and improving the resilience of a WDN.



(a) One valve placement configuration with 4 isolation valves



(b) Another valve placement configuration with 4 isolation valves

Figure 7.1: Isolation valve placement strategies for a simple network. Pipes are assumed to have equal failure probability

In this study, an optimization process is proposed to find the least pipe isolation risk valve placement configuration of a WDN for a given number of valves. The key idea is to work backward: instead of incrementally adding valves to the system until reaching the desired number, valves are gradually removed from an ideal state. The process starts from a system with the ideal or safest valve placement configuration, which has the lowest pipe isolation risk among all possible valve placement configurations. The most redundant valve placement configuration, that is to install two valves on every pipes in the network, is used

as the initial configuration for this study. Then the consequence of valve removal, which is measured as the increment of system pipe isolation risk after the valve removal, is calculated for all the valves under the current configuration. The valve with the least removal impact is removed from the network. The valve removal process is repeated multiple times until the desired number of valves is reached. Mathematically, the optimization process can be expressed as:

$$\begin{aligned}\overline{D(n-1)} &= \overline{D(n)} + \{f_k\bar{D}_k - (f_i\bar{D}_i + f_j\bar{D}_j)\}_{min} \\ &= \overline{D(n)} + \{f_i(\bar{D}_k - \bar{D}_i) + f_j(\bar{D}_k - \bar{D}_j)\}_{min}\end{aligned} \tag{7.1}$$

where $\overline{D(n)}$ is the pipe isolation risk for the system with n functioning valves, f_i and f_j are failure probability for segment i and j, respectively; f_k is the failure probability of the new segment k, which is created by merging segment i and j (i.e., $f_k = f_i + f_j$). $\bar{D}_k$ is the pipe isolation risk of the new segment k, which may not equal the sum of risk for the before-merging components due to network effects (isolating a group of pipes may create additional isolation risk than isolating them individually).

Proof of the optimality

The optimality of the procedure described by Equation 7.1 can be proved by mathematical induction:

Theorem 1. *Pipe isolation risk for a WDN with N isolation valves is minimized when valves are placed following the process described by Equation 7.1.*

Proof. The conventional three-step induction procedure is used for the proof: base case, inductive hypothesis, and the inductive step.

Base case: The initial valve placement configuration is to have two valves installed at each end of a pipe. Thus every pipe can be isolated individually from the system at the initial state. Therefore, the system pipe isolation risk $\overline{D(N)}$ is minimum with $N = 2P$, where P are the total number of pipes in the system.

Inductive hypothesis: Suppose the theorem holds for all values of N down to some M, with $M \geq 1$.

Inductive step: Let $N = M-1$, then $\overline{D(N)} = \overline{D(M)} + \{f_k\bar{D}_k - (f_i\bar{D}_i + f_j\bar{D}_j)\}_{min}$. By the inductive hypothesis, $\overline{D(M)}$ is the minimum across all valve configurations with M valves. Since removing multiple valves together is equivalent to remove these valves in sequence, the desired N valves state can only be achieved by reducing one valve from the M valves state. Hence $\overline{D(N)}$ is the minimum. By the principle of mathematical induction, the theorem holds for all $0 \leq n \leq 2P$. □

Implementation details

Although the proposed valve placement method is theoretically optimal, it must be implemented with a reasonable computation complexity to be practically useful. This section

focuses on the implementation details of the proposed valve placement method. First, we lay out a straightforward implementation approach and discuss its limitations. An alternative approach with less computation complexity is presented next. The computation feasibility of the two methods on WDNs with various sizes is tested and discussed at the end.

Naive implementation

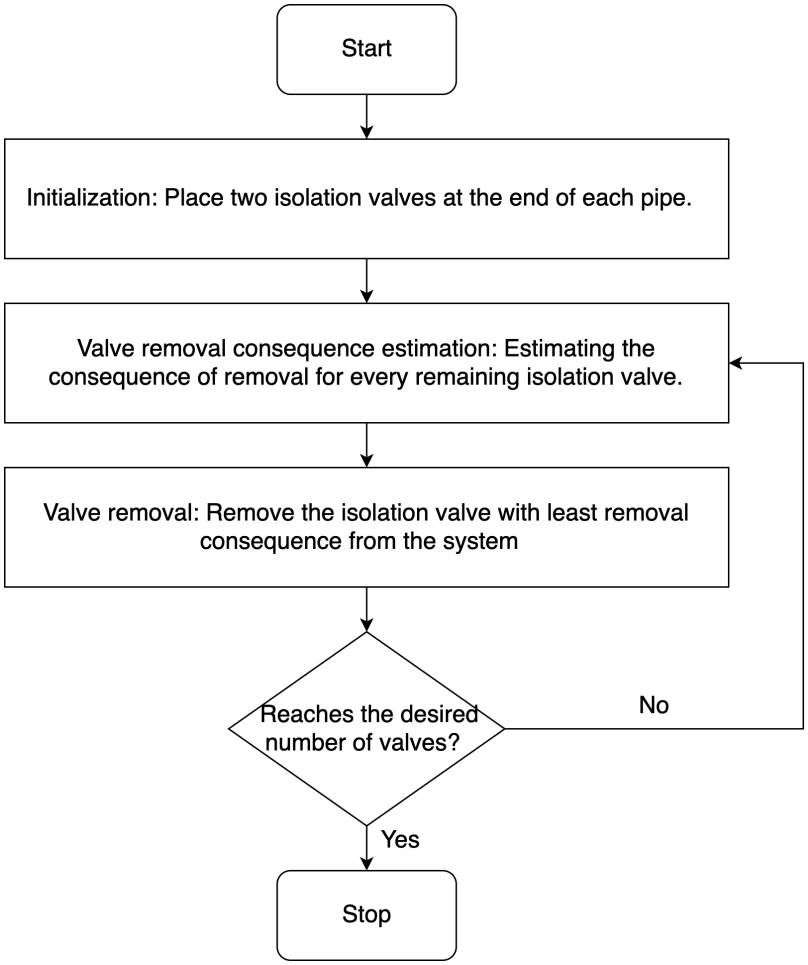


Figure 7.2: Flowchart of the naive implementation for the proposed valve placement strategy

A straightforward implementation of the proposed valve placement strategy is summarized in Figure 7.2. It has three major steps:

- Initialization. The network is initialized with the ideal valve placement strategy by applying two valves at the end of each pipe in the system. In this way, all the pipes can be isolated individually; thus, the system pipe isolation risk is minimal.

- Valve removal consequence estimation. The consequence of removal for every valve in the system is estimated in this step. The concept of Segment-Valve representation, introduced in Chap 5, is used to calculate the valve removal consequence. First, the Segment-Valve graph is constructed using the current valve placement configuration. Removing one valve means removing the corresponding valve edge from the Segment-Valve graph, resulting in the merge of two end-point segments. The removal consequence of the valve is then the pipe isolation risk difference between the edge-reduced Segment-Valve graph and the initial Segment-valve graph. Calculating the network pipe-isolation risk using the Segment-Valve graph is discussed in detail in the previous (Chapter (5), and the core ideas are restated as the following. For each isolation segment, calculate the segment failure probability, which is the sum of pipe failure probabilities for all the pipes in the segment. Remove all the pipes contained in the segment from the WDN and run the steady-state pressure-dependent hydraulic simulation (PDD) on the modified network. The isolation cost of the segment is the total demand loss across the network due to the segment isolation. The segment isolation risk is then the product of the segment failure probability and the segment isolation cost. The pipe isolation risk for the whole system is the summation of all the segment isolation risks.

- Valve removal. Remove the valve with the least removal consequence from the network. In this way, the remaining network has the minimum pipe isolation risk across all valve placement configurations with the same amount of valves. If the amount of remaining valves is still bigger than the desired amount, restart the valve removal consequence estimation step.

The most computationally intensive part of the naive implementation is the valve removal consequence estimation. For a network with P pipes, the total number of valves is $2P$ at the initial condition (two valves per pipe). Therefore, $2P - N$ "valve removal consequence estimation" steps need to be performed to obtain a minimum isolation risk network with N valves. For each such step, the removal consequence of every remaining valve needs to be estimated, which involves hydraulic simulation for the whole network. Consequently, the total number of hydraulic simulations required for obtaining the minimum risk valve placement configuration for a P pipes N valves network using the naive implementation is:

$$\begin{aligned} \text{SimulationRounds} &= (2P) + (2P-1) + ... + N \\ &= \frac{(2P)(2P+N)}{2} \end{aligned} \tag{7.2}$$

In this work, the big O notation is used as the measure to characterize the computation complexity of the implemented algorithm. Big O notation focuses on representing the upper bound on the growth rate of the function and is widely used in the field of computer science to describe how an algorithm behaves as its input size growth [95]. Formerly, suppose f(x) and g(x) are two functions defined on some subset of the real numbers.

$$f(x) = O(g(x)) \tag{7.3}$$

if and only if there exist constants N and C such that

$$|f(x)| \leq C|g(x)|, \text{for all } x > N \tag{7.4}$$

In summary, $g(x)$ is the asymptotic upper bounds of function $f(x)$, meaning f does not grow faster than g.

Using the big O notation, the computation complexity for the naive implementation is $O(P^2)$. In other words, for a network with thousands of pipes, which is common in real-life situations, over millions of full-scale hydraulic simulations need to be performed for the naive implementation of the proposed valve placement finding method. Suppose that each simulation takes 1 second to run, the total running time for the algorithm will be over 1 million seconds or 280 hours. Therefore, the naive implementation of the valve placement configuration algorithm can be ineffective in a practical sense. The following section presents an efficient implementation of the algorithm that overcomes the computation limitation of the one described in this section.

Efficient implementation

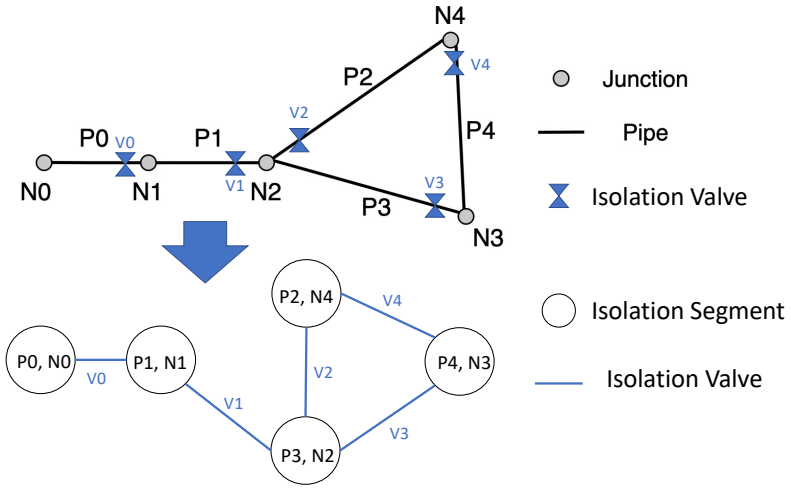


Figure 7.3: Example network for valve removal consequence estimation. Above: Junction-Pipe representation of the network. Below: Segment-Valve representation of the network

An efficient implementation of the algorithm can be obtained by eliminating the redundant computation steps from the naive implementation. As discussed previously, the most computationally intensive part of the naive implementation is the valve removal consequence estimation step, which uses system-wise hydraulic simulation to estimate the removal consequence of each valve in the system. Luckily, it turns out that we do not need to estimate

the removal impact of all valves in each state. Recall that a valve removal operation causes the merging of two segments. Therefore, the removal consequence of all the other segments remains the same after the removal operation. In other words, only the impacts of valves related to the newly merged segment need to be recalculated at each step.

Let's consider a simple example to illustrate this idea. Figure 7.3 shows a network with 5 pipes, 5 nodes, and 5 valves. The network can be converted to the Segment-Valve representation, which is consisted of 5 isolation segments and t valves. Suppose that V0 is removed from the previous step, the two end segments of valve V0 are merged (Figure 7.4). Comparing to the pre-removal Segment-Valve graph, the major structure of the after-removal graph stays almost intact. Only the removal consequence of valve V1 needs to be re-estimated since it is linked to the merged segments. The removal consequences of all other valves stay the same as in the previous step. Therefore, the required hydraulic simulation steps are reduced from 4 to 1 in this case.

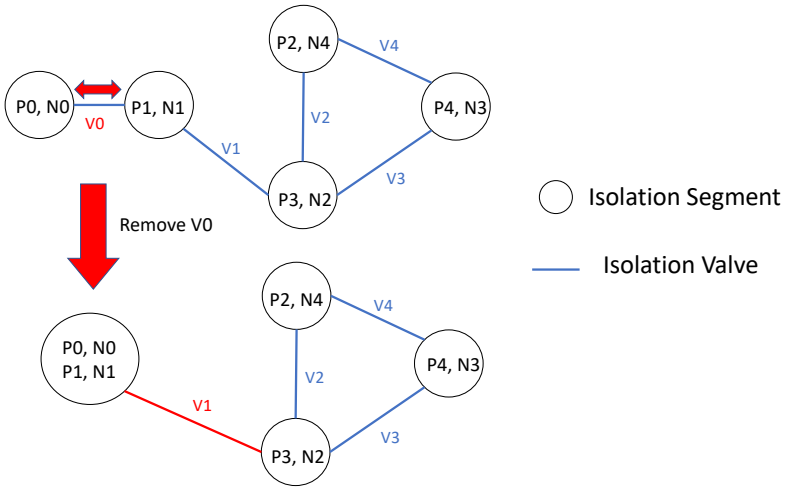


Figure 7.4: Network state before and after one isolation valve removal operation

Figure 7.5 shows the updated procedure to find the optimal valves placement configuration. After the initial valve placement, two operations are conducted at the initial stage. First, the Segment-Valve graph is constructed using the procedure described in the previous chapter (Chap 5). Next, the removal impact of each valve is estimated and memorized, which is the difference between total system isolation risk before and after the valve removal. After the initialization process, the valve with the most negligible removal impact is removed from the system, and the risk of the newly merged segment is computed and memorized to update the total system isolation risk. Then, the removal impacts of valves that are linked to the merged segment are updated. The removal consequence of all other valves stays the same. The removal/update procedure is repeated until the desired number of valves is reached.

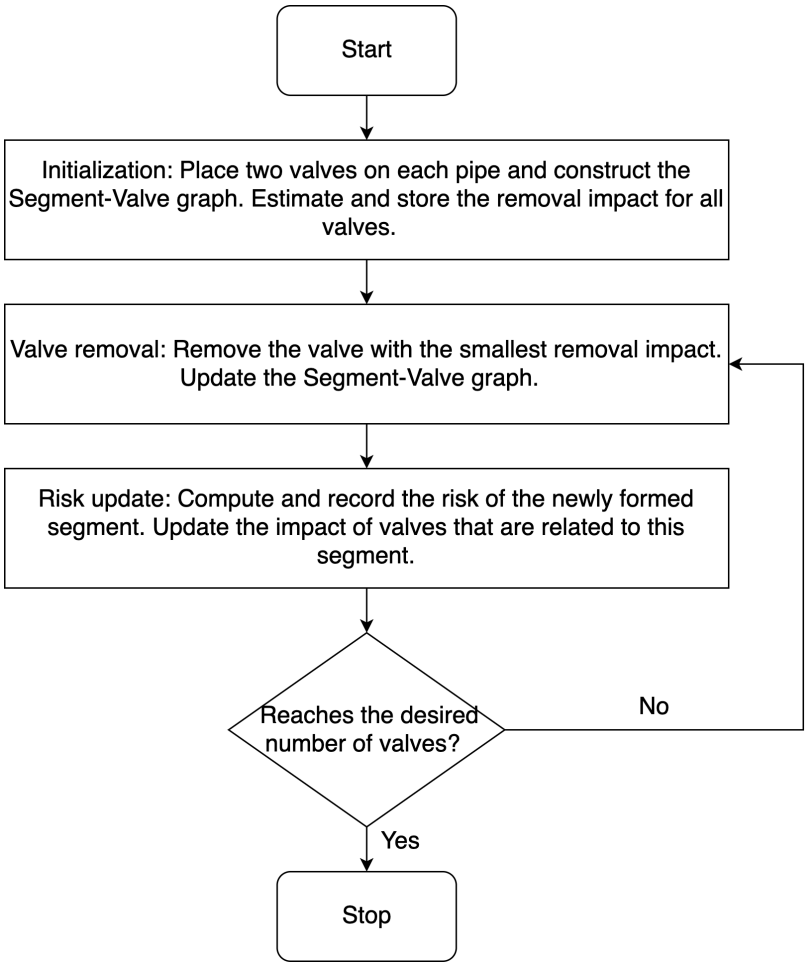


Figure 7.5: Flowchart of the efficient implementation for the proposed valve placement strategy

The updated optimal isolation placement algorithm is much more efficient than the naive implementation. In the initialization phase, system-wise hydraulic simulation is executed $2P$ times since each pipe has two valves installed. In the valve removal state, only k valves that are related to the merged isolation segment from the previous steps get updated, which takes k hydraulic simulation procedures. The removal/update procedure is repeated n times to get the final result. In summary, the required hydraulic simulation steps for the updated valve placement procedure is $2P + n\bar{k}$, where $\bar{k}$ is the averaged number of valves that are linked to the newly merged segment in each valve removal step. From Equation 5.6, large size segments tend to increase the overall system risk if the risk variations across segments are not excessively high. Hence, it is reasonable to assume $\bar{k} \lll P$. Therefore, the computation complexity of the implemented algorithm can be approximated as $O(P + n)$. Since n and P are in the same order, the computation complexity can be further simplified as $O(P)$,

which is much more efficient than the $O(P^2)$ complexity of the previous case. For instance, for a thousand pipes network, the updated algorithm only requires the order of thousand hydraulic simulations, which is much less than the million simulation requirements of the naive implementation.

Necessity of running backwards

The proposed optimization process is formulated backwardly. The least significant valves are gradually removed to reach the design with the desired number of valves starting from the safest valve placement configuration. Theoretically, one can also obtain the minimum isolation risk valve placement configuration through a reversed process. Starting with a network with no valves installed, one may add the most important valve progressively to the system until the desired number of valves is reached. Mathematically, Equation 7.1 can be rewritten as:

$$\begin{aligned} \overline{D(n)} &= \overline{D(n-1)} - \{f_k \bar{D}_k - (f_i \bar{D}_i + f_j \bar{D}_j)\}_{min} \\ &= \overline{D(n-1)} + \{f_i(\bar{D}_i - \bar{D}_k) + f_j(\bar{D}_j - \bar{D}_k)\}_{min} \end{aligned} \tag{7.5}$$

where $\overline{D(n)}$ is the pipe isolation risk for the system with n functioning valves, f_k, $\bar{D}_k$ is the failure probability and the pipe isolation risk of segment k, respectively; f_i and f_j are failure probability of the two split segments resulting from valve installation in segment k (i.e., $f_k = f_i + f_j$), and $\bar{D}_i$ and $\bar{D}_j$ are their corresponding segment isolation risk.

Since Equation 7.5 is equivalent to Equation 7.1, the valve placement strategy using Equation 7.5 is also optimal.

Figure 7.6 summarizes the workflow of the forward formulated optimal valve placement strategy. The process starts with a null network (i.e., no valves installed) and evaluates the potential benefit of adding an valve to all possible valve installation locations. Since there are no valves installed in the system, an valve can be installed on the end of each pipe. The benefit of valve installation is defined herein as the pipe isolation risk reduction for the whole system. After evaluating all potential valve installation locations, an valve is installed on the location that mostly reduces the overall system pipe isolation risk. Since a new valve is installed in the system, the Segment-Valve graph of the system needs to be updated. With the newly updated Segment-Valve graph, one can estimate the valve adding benefits for all other potential valve installations. The above process is repeated until the desired number of valves is installed in the system, and the resulting valve placement configuration is guaranteed to be optimal.

Similar to the backward formulated valve placement process, the most computationally intensive part of the forward configured process is evaluating the importance of an valve. For the backward case, the importance of an valve is the system isolation risk increase after the removal. In this case, it is the decrease of system isolation risk after the valve installation. Adding an valve into the system breaks the corresponding isolation segment into two small parts, reducing the overall system isolation risk. However, splitting a segment is much more

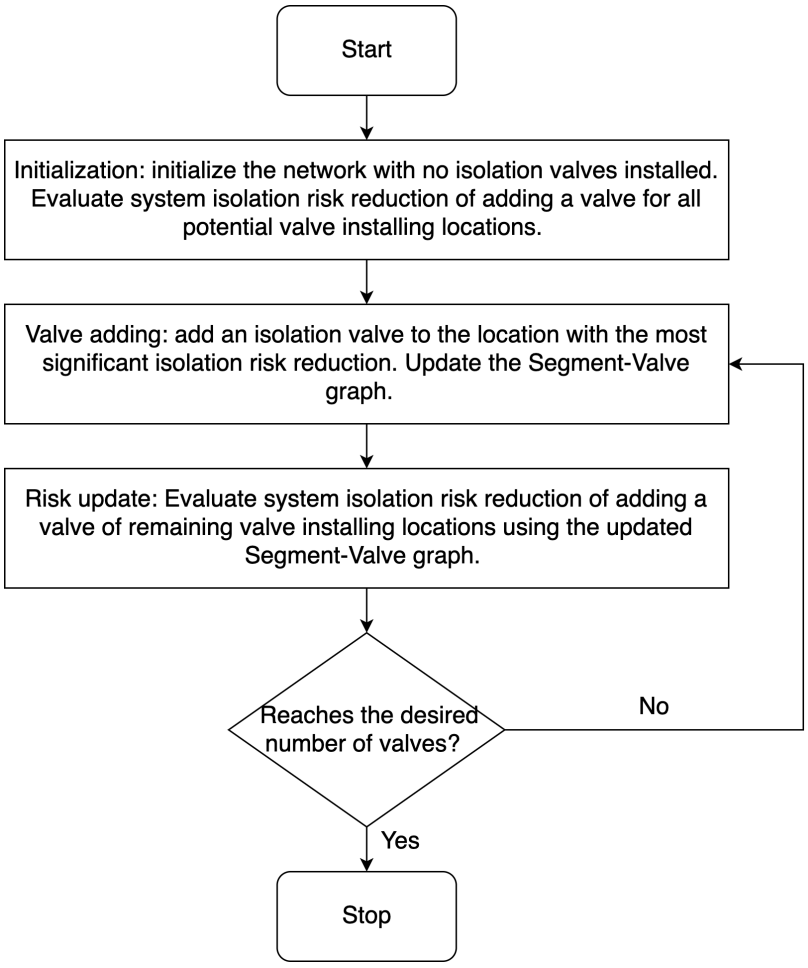


Figure 7.6: Flowchart of the forward formulated valve placement strategy

computationally intensive than merging two segments, the operation required to estimate valve removal consequences. Merging two segments only involves an edge contraction operation: delete the link and merge the segments of the two ends, which has $O(1)$ computation complexity. On the other hand, segment splitting is much more complicated. Specifically, for each added valve, the segment finding process from Chapter 5 Section 5.1 is required to split the original isolation segment. Since the segment finding process checks all the pipes in the segment, the computation complexity for segment splitting is $O(k)$, where k is the number of pipes inside the prior-splitting segment. Because the valve importance evaluation procedure is invoked multiple times during each step of the valve placement process, it is much more efficient to run the optimization process backward instead of forward.

Table 7.1: Common used rule-of-thumb isolation valve placement strategies

Valve Placement Strategy	Description
N rule	If there are n pipe legs at a node, n valves need to be installed at that node. The n rule represents a fully valved system with two valves associated with each pipe.
N-1 rule	if there are n pipe legs at a node, install n-1 valves at that node.
One valve per pipe rule	All pipes are installed with exact one valve on one end
One valve per pipe and one valve per node rule	The one valve per pipe rule may result in some nodes with no valves associated with them. One valve is installed on these nodes.

Relation to rule of thumb valve placement strategies

The proposed valve placement process provides theoretical explanations for the effectiveness of some commonly used rule-of-thumb valve placement strategies, listed in Table 7.1 [37]. Figure 7.7 shows examples of the listed rule-of-thumb valve placement strategies.

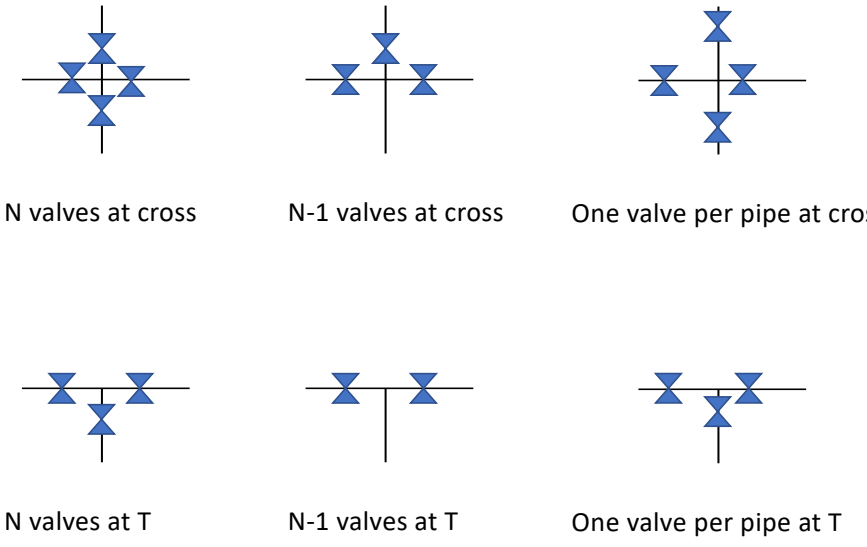


Figure 7.7: Examples of common rule-of-thumb valve placement strategies. Note that the one valve per pipe strategy can place the valve on either end of the pipe.

The most robust valve placement strategy is the N valving strategy, which requires the

number of valves installed at a junction equals the number of linked pipes in that junction. This is equivalent to having two valves installed at each pipe's ends, which is the initial state of the proposed valve placement process. In other words, a system with valves installed by the N valve placement strategy reaches the theoretical lower bound of pipe isolation risk.

The N-1 valve placement strategy uses one less valve for a junction comparing to the N configuration. Under the N-1 configuration, all pipes can be isolated individually. The N-1 placement strategy is the solution to the proposed valve placement process under certain constraints. Specifically, by conducting the optimization process with the requirement that the system pipe isolation risk remains at the lower bound (system risk at the initial condition), the N-1 configuration is the solution with the max number of removed valves. In other words, the N-1 valve placement strategy uses the least number of valves to guarantee that each pipe can be isolated individually, achieving the lower bound of system pipe isolation risk. Note that there is no valve redundancy in this situation, i.e., failure of one valve will increase isolation segment size, increasing the total system pipe isolation risk.

When budget is the concern, other heuristics such as the "one valve per pipe" rule, and the "one valve per pipe and one valve per node" rule are often adopted to reduce the number of valves even more. Assuming all pipes are equal-likely to fail, and the failure consequences are the same, the proposed process generates the "one valve per pipe" and the "one valve per pipe and one valve per node" valve placement configuration. According to Equation 5.9, segment size variances increase system pipe isolation risk under the equal-fail equal-consequence assumption. In other words, a WDN with minimum pipe isolation risk (generated by the proposed procedure) under such a condition has small variance regarding isolation segments sizes. The small segment size variance configuration can be achieved by ensuring each pipe (and node, if required) has at least one valve installed. Therefore, the one valve per pipe rule is optimal when pipe failure data is unavailable and hydraulic simulation can not be conducted to quantify the isolation cost.

In summary, both the N and the N-1 valve placement strategies are the solutions of the proposed valve placement process with theoretical lower bound of system-wise pipe isolation risk. However, N valving strategy is a more robust solution than the N-1 valving strategy due to the redundancy of valves. When information on the WDN is inadequate, one can adopt the "one valve per pipe" rule, or the "one valve per pipe and one valve per node" rule to achieve a low pipe isolation risk system with a limited amount of valves.

Additional valve placement

In most cases, adding additional valves into a WDN is more practical than redesigning the valve configuration to improve system resilience. The proposed valve placement procedure, with slight modification, can guide the installation of new valves for a WDN with installed valves.

Same as the original algorithm, a backward-configured optimization process is used to find the best installation locations for additional valves. The optimization process starts from the most redundant valve system (every pipe has two valves at the ends) and iteratively

removes the least impact valve. However, instead of considering all the installed valves at each removal step, the already deployed valves are excluded in the least-impact valve searching process. Mathematically, the modified optimization process can be expressed as:

$$\begin{aligned}\overline{D(n-1)} &= \overline{D(n)} + \{f_k\bar{D}_k - (f_i\bar{D}_i + f_j\bar{D}_j)\}_{min} \quad \text{for all available valves} \\ &= \overline{D(n)} + \{f_i(\bar{D}_k - \bar{D}_i) + f_j(\bar{D}_k - \bar{D}_j)\}_{min} \quad \text{for all available valves}\end{aligned} \tag{7.6}$$

where available valves are all valves other than the already deployed valves, $\overline{D(n)}$ is the pipe isolation risk for the system with n valves, f_i and f_j are failure probability for segment i and j, which are linked by an available valve; f_k is the failure probability of the new segment k, which is created by merging segment i and j (i.e., $f_k = f_i + f_j$). $\bar{D}_k$ is the pipe isolation risk of the new segment k.

Equation 7.6 is the constrained version of Equation 7.1. Specifically, the least-impact valve removal searching space of Equation 7.6 is a subset of the searching space of Equation 7.1. All the other aspects (e.g., valve removal impact assessment) are the same for the two algorithms. Therefore, the argument of optimality (Section 7.2) still holds for the modified algorithm. However, the interpretation of optimality is different in this case. Instead of obtaining the minimal risk valve configuration across all possible configurations, the modified algorithm finds the minimal risk configuration under the constrain that the already deployed valves must be included.

Starting from the most redundant valve configuration, the optimization process ends when all available isolation valves are removed (only the already deployed valves remain). It generates optimal valve placement configurations (constrained) from the current situation (no additional valve installed) to the most redundant case (each pipe has two valves installed at the ends). The difference between the obtained valve configuration and the existing one gives the installation locations for the additional valves.

7.3 Case Studies

This chapter uses the same case study networks introduced in Chapter 6 to test the effectiveness of the proposed valve placement method. The first case study considers the EBMUD Alameda Island WDN, a well-connected medium-size network. Next, the EBMUD Round Hill WDN, a small-scale sparsely connected network, is tested to illustrate the proposed method's effectiveness on networks with different scales and connectivity patterns.

Alameda island WDN

The EBMUD Alameda network comprises 2508 internal nodes and 3028 pipes for a total length of around 194 km. 1999 valves are currently installed in the system, see Figure 7.8. The network is fed by two transmission line pipes (source pipes) linked to other parts of the EBMUD network. Pipes in the network are made with different lengths and materials, and

the network is mainly composed of well-connected grid-like structures, with a few branch-like sub-structures on the edges of the network. See Chap 6 Section 6.1 for details of the network.

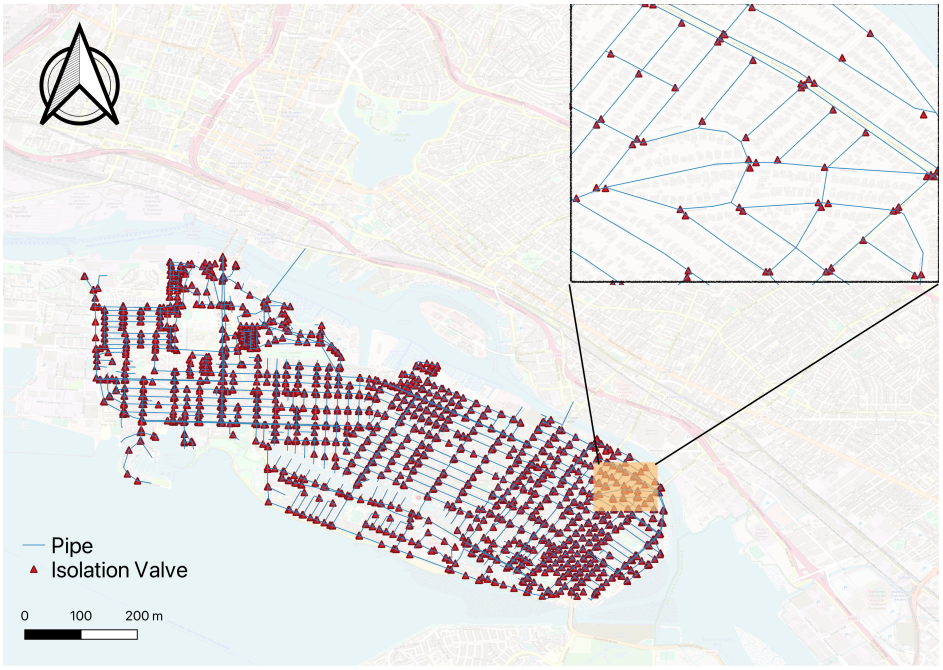


Figure 7.8: Location of deployed isolation valves for the Alameda Island WDN

Effectiveness of the proposed valve placement strategy

The first goal of the case study is to compare the pipe isolation risk between a system with the currently deployed valve placement configuration and a system with a valve placement configuration generated by the proposed method. A valve placement configuration with the same amount of valves (1999) as the current valving configuration is generated using the proposed valve placement method, shown in Figure 7.9. Compared to the deployed valve placement configuration, the proposed one distributes valves more evenly across the network. Considering region A in Figure 7.9 for an example. The original configuration placed valves on every terminal pipe (a pipe that has only one end connected to the WDN), which is inefficient in limiting pipe isolation risks. First, terminal pipes have a small failure probability due to short lengths. In addition, since terminal pipes are the least connected components in the system, their failures have small impacts on other parts of the system (small indirect cost). On the other hand, the proposed method only places valves on the transmission line, resulting in fewer valves than the original configuration. Since the water of terminal pipes comes from the transmission line, failures of transmission line pipes have

larger consequences than failures of terminal pipes. Placing valves on them reduces the probability of simultaneously isolating multiple transmission line pipes, lowering the system's total isolation risk. The proposed method places more valves than the original one in other parts of the system. For example, valves are sparsely placed under the original configuration in region B, which contains multiple loops. However, failure of a pipe in one loop requires the isolation of all pipes in it. By placing more valves in loops, the size of potential isolation segments is restrained in the region, resulting in small system risks.

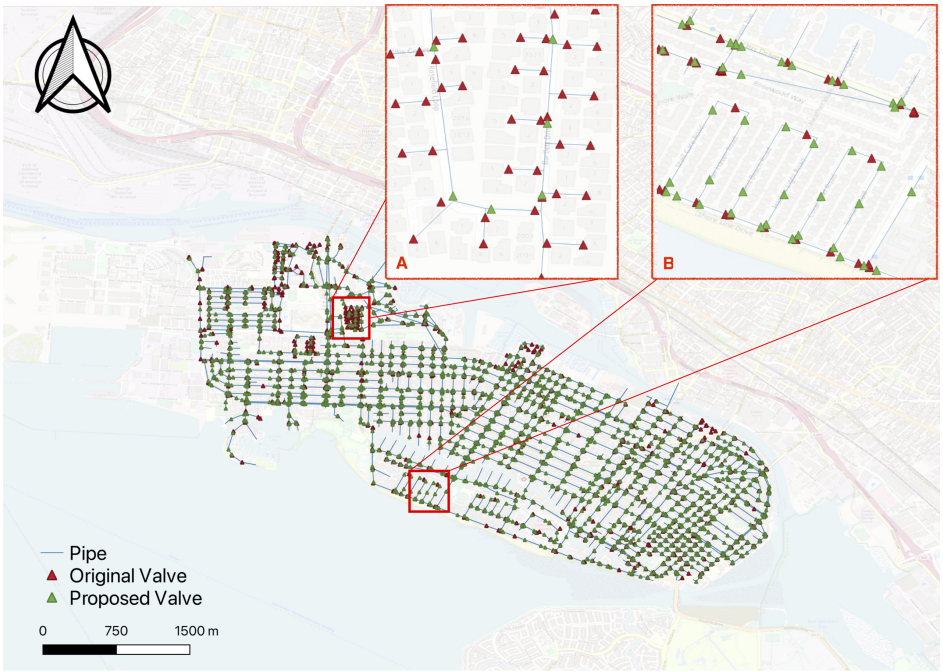


Figure 7.9: Location of proposed isolation valves for the Alameda Island WDN

Next, the system-wise pipe isolation risk of the proposed configuration and the current one is simulated using HydrauSim. The simulation results are summarized in Table 7.2. With the same number of valves, the proposed method's valve placement configuration reduced the overall system pipe isolation risk by 50%, from 6.59 GPM to 3.27 GPM, comparing to the current valving configuration. Properties of isolation segments for the system with the two valving configurations are different as well. Comparing to the original configuration, the valve placement configuration generated by the proposed method has a smaller mean and variance of isolation segment sizes. This matches the discussion from Chapter 5 Section 5.3 that a well-connected network with a small mean and variance of isolation segments implies a small pipe isolation risk.

To evaluate the consequence of having missing/malfunctioning valves, network risks under 5%, 15%, 25%, 30%, 35%, and 40% valve failure rates are simulated for both valve placement configurations. The valve failure rate range is chosen based on the estimated

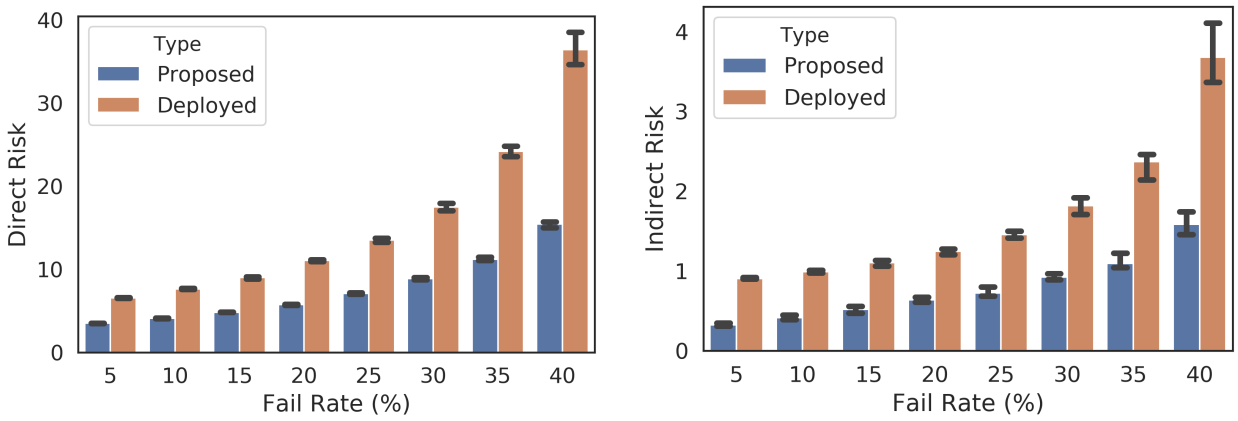
Table 7.2: Network performance under different valve placement strategies for Alameda Island WDN

Valve Placement Strategy	Number of valves	Average Segment Size	Segment Size Variation	Pipe Isolation Risk(GPM)
Current Strategy	1999	2.09	2.01	6.59
Proposed Strategy	1999	1.88	1.62	3.27

10-20% valve failure rate from EBMUD engineers. Because any valve can fail in the system, there are numerous choices on failed valve selections. The uncertainties of which valve to fail are simulated using the Monte Carlo(MC) method. Specifically, for each valve failure rate scenario, many realizations of the valve failure events are simulated. The desired number of valves (calculated from the given valve failure rate) are randomly failed for each realization on the network and the resulting network properties are analyzed using the Segment-Valve graph. Next, hydraulic simulation is performed on the generated realizations to calculate the pipe isolation risk of the system. Statistical measures, such as the mean and variance of risk for each valve failure rate scenario, can be obtained across realizations. The number of the MC simulation is determined during the simulation process [123]. The simulation starts with a fixed size warm-up simulation round (100 simulations in this case). After the warm-up round, if the mean and variance of the estimated risk with additional simulations are within 2 percent difference compared with those without additional simulations, the Monte Carlo simulation is terminated.

The simulation result for both valve placement configurations is shown by Fig 7.10, and the statistics are summarized in Table 7.3. Like the previous chapter, the total system pipe isolation risk decomposed into two sub-risk to indicate different aspects of isolation consequence. Specifically, the total risk is split into the direct risk, the risk inside the isolated segment alone, and the indirect risk, the total isolation risk on the rest of the system other than the isolated segment. Overall, the proposed valve placement strategy produces a safer network than the currently deployed valve placement strategy on both direct and indirect pipe isolation risk measures for every tested valve damage level. For instance, when the valve failure rate is 10%, the direct and indirect risks are 7.64 GPM and 0.34 GPM for the currently deployed valve placement configuration, respectively. On the other hand, the direct risk and indirect risk are 4.09 GPM and 0.15 GPM for the system using the proposed valve placement strategy. Therefore, the proposed valve placement strategy reduces system risk by 46% when the valve failure rate is 10%. The standard deviation (STD) of both risk measures for the proposed strategy is also smaller than the original strategy for all tested valve failure rate scenarios. On average, the STD of the total pipe isolation risk for the network with the proposed valve configuration is 70% less than the network with the original valve configuration. In conclusion, the network configured with the proposed

valving strategy has a more stable performance than the original one when valve conditions are uncertain.



(a) Direct pipe isolation risk

(b) Indirect pipe isolation risk

Figure 7.10: Comparison of the system pipe isolation risks under different valve placement strategies for Alameda Island WDN

As the valve failure rate increases, both the mean and standard deviation of the calculated risk increase for the tested two valving strategies. The increasing trend of both measures accelerates as the number of failed valves increases. In other words, failing a certain amount of valves on an already valve-damaged network will result in more considerable isolation risk and uncertainties compared to failing the same amount of valves on a relatively less-damaged network. Such a pattern of increase is due to the changes of isolation segment properties inside the network. See Chap 5 Section 5.3 for details. Note that the risk-increasing degree is much smaller for the network configured with the proposed valve placement strategy than the one configured with the original setting. Specifically, the mean value of direct risk increases by 12.39 GPM as the valve failure rate increases from 0% to 40% for the system with the proposed valve placement configuration. On the other hand, the mean value of direct risk increases by 30.69 GPM of the same range of valve failure rate for the system with the currently deployed setting. The differences in risk change are also apparent in the indirect pipe isolation risk. The indirect risk raises by 2.83 GPM for the system with the original valve configuration (valve failure rate 0% to 40%). This number decreases to 1.34 GPM for the system with the proposed valve configuration. Therefore, the proposed valve placement strategy produces a more robust network than the original one regarding valve failure scenarios.

Table 7.3: Statistics of system pipe isolation risks under different valve placement configurations with various valve failure rates for Alameda Island WDN

Valve failure rates	Valving placement strategy	Mean direct risk (GPM)	Direct risk STD	Mean indirect Risk (GPM)	Indirect risk STD
0%	Current	5.74	0	0.85	0
	Proposed	3.03	0	0.24	0
5%	Current	6.56	0.16	0.91	0.05
	Proposed	3.49	0.09	0.32	0.08
10%	Current	7.64	0.34	0.99	0.08
	Proposed	4.09	0.15	0.41	0.13
15%	Current	9.02	0.52	1.1	0.15
	Proposed	4.83	0.19	0.52	0.15
20%	Current	11.05	0.77	1.24	0.18
	Proposed	5.73	0.27	0.64	0.17
25%	Current	13.5	1.25	1.46	0.27
	Proposed	7.08	0.47	0.72	0.26
30%	Current	17.47	2.27	1.82	2.5
	Proposed	8.88	0.64	0.92	0.28
35%	Current	24.17	3.92	2.37	2.64
	Proposed	11.21	0.84	1.09	0.35
40%	Current	36.43	9.82	3.68	3.95
	Proposed	15.42	1.58	1.58	0.63

Lower bound of the system risk

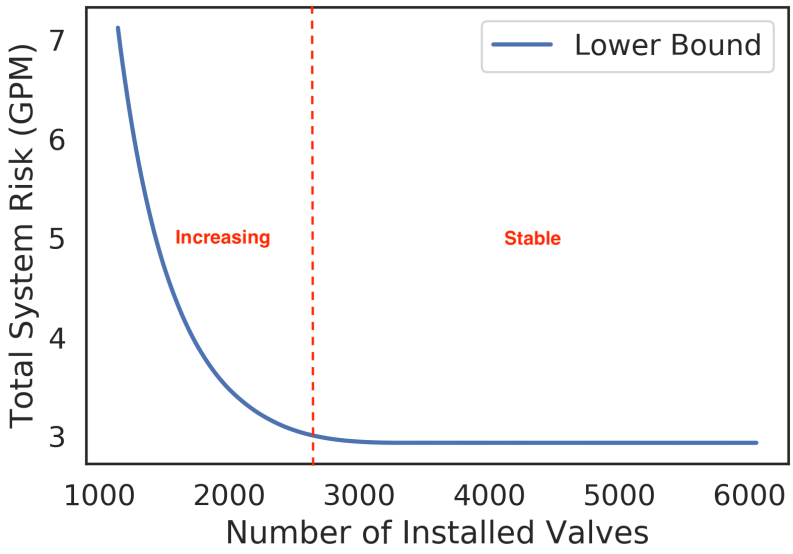


Figure 7.11: The lower bound of pipe isolation risk for the Alameda Island WDN with respect to the number of isolation valves

As discussed in the previous sections, the proposed valve placement strategy produces a system with minimal pipe isolation risk for a given number of valves. In other words, it provides the lower bounds of the pipe isolation risk for a WDN with a given number of valves. The lower bound of pipe isolation risk for the Alameda Island WDN with respect to the number of valves are computed and shown in Figure 7.11.

One critical observation from Figure 7.11 is that there are two distinct phases for system pipe isolation risk concerning the number of installed valves. As long as the number of valves is sufficient, which is above 2800 valves, the lower bound of the total pipe isolation risk remains the same regardless of the number of installed valves. However, as the number of installed valves drops from the 2800 threshold, the lower bound of pipe isolation risk of the system increase dramatically. Specifically, decreasing the number of installed valves from 2000 to 1000 increases the pipe isolation risk of the system from 3.57 GPM to 7.21 GPM, a 101% increase. The two phases phenomena have practical implications. It is essential to ensure that the system is configured with at least the number of valves as the threshold value, which is around 2800 valves in this case. Having additional valves installed provides redundancy for the system, making it robust again to potential valve failure events but does not decrease the system risk anymore. On the other hand, if the number of installed valves is smaller than the phase change threshold, the system will be vulnerable to valve failure events. Under such circumstances, the system's pipe isolation risk is high even when all

valves are functioning correctly, and a small number of valve failures will drive up the total pipe isolation risk to a greater level.

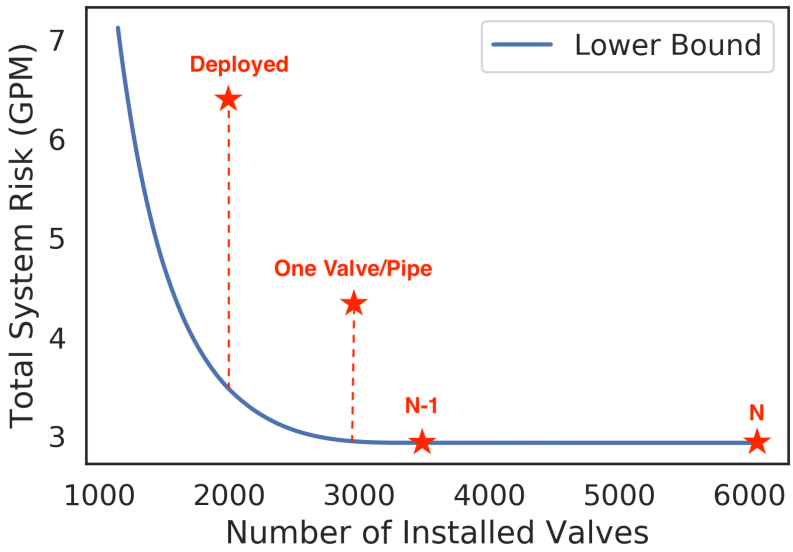


Figure 7.12: System pipe isolation risks of different isolation valve placement strategies for the Alameda Island WDN

Table 7.4: Pipe isolation risks for the Alameda WDN with different valve placement configuration

Valve Placement Strategy	Number of Valves	Direct risk (GPM)	Indirect risk (GPM)	Total risk (GPM)	Risk lower bound (GPM)
N	6052	2.58	0.35	2.93	2.93
N-1	3545	2.58	0.35	2.93	2.93
One valve per pipe	3026	4.67	0.73	5.4	3.01
Current Strategy (deployed)	1999	5.74	0.85	6.59	3.27

The effectiveness of the widely-used rule of thumb valve placement strategies is tested using the risk lower bound produced by the proposed method. Figure 7.12 shows the comparison result, and Table 7.4 lists the detailed numbers. Both the N and N-1 valve placement strategy achieves the lower bound of pipe isolation risk. Therefore, both strategies are optimal. The difference between the two is that the N valve placement strategy provides more

valve redundancy: the system configured with the N valve placement strategy is more robust against valve failure events than the system with the N-1 valve placement strategy. On the other hand, the rest placement strategies create systems with higher pipe isolation risk than the risk lower bound. Specifically, the total system risk is 4.67 GPM for the one valve per pipe strategy, which is 55% higher than the risk lower bound. Such difference is because the one pipe per valve strategy does not guarantee that each pipe will be isolated individually. See Figure 7.13 for an example. In this case, the four valves are installed at opposite junctions for the one valve per pipe strategy. Hence, there is only one isolation segment with four pipes in this case. Comparing to the N-1 valving strategy, which results in 4 isolation segments with individual pipes, the isolation risk is higher for the one valve per pipe strategy. Lastly, the currently deployed strategy is far from optimal, as discussed in the previous sections. The same pipe isolation risk level, a 6.59 GMP, can be achieved using only 1400 valves if the proposed valve placement strategy is adopted.

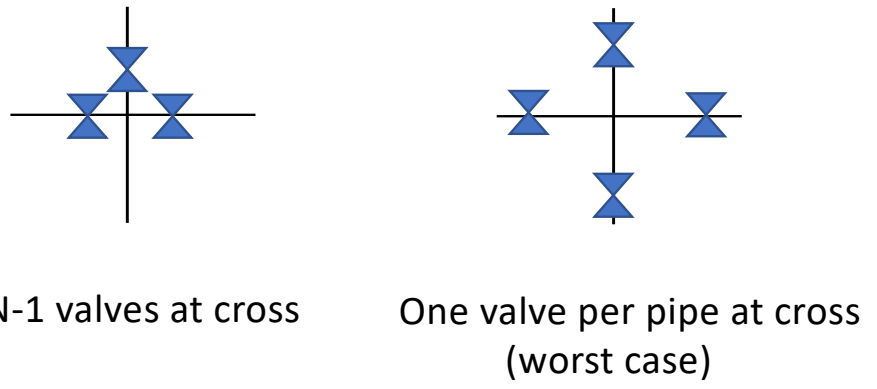


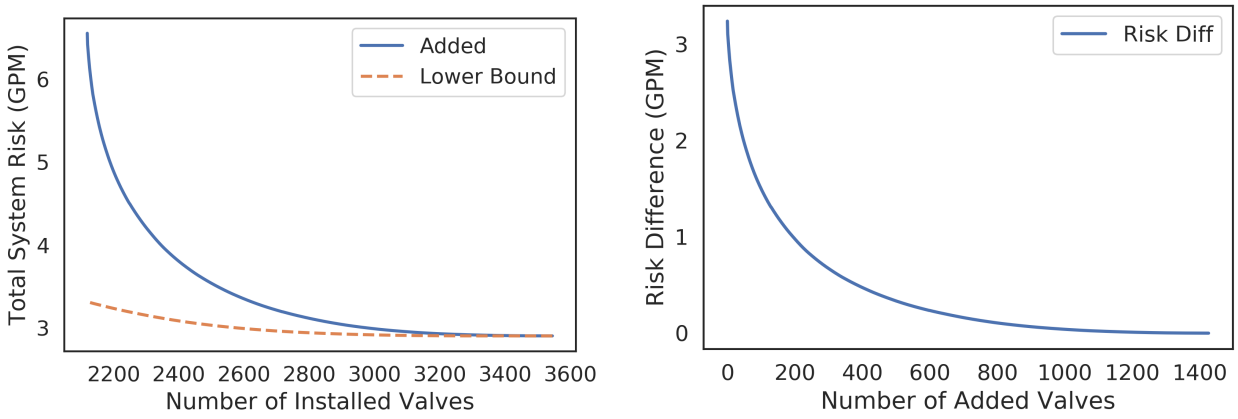
Figure 7.13: Examples of the N1 valving placement strategy Vs. One valve per pipe strategy (worst case). Note that all four pipes belongs to the same isolation segment for the One valve per pipe strategy

Additional valve installation

The proposed valve adding algorithm is tested for the Alameda Island WDN. Starting from a system with the currently deployed valve placement strategy (1999 valves), valves are added one by one until all pipes have two valves installed in both ends.

Figure 7.14a compares the pipe isolation risk of the resulting valve configuration with the risk lower bound at different levels of installed valves. There is a significant difference between the risks at the initial stage, signaling the inefficiency of the currently deployed

valve placement strategy. As more valves are installed according to the proposed method, the difference shrinks and eventually, the system risk reaches the lower bound. To analyze the convergence trend, the risk difference is plotted against the number of added valves, shown by Figure 7.14b. The risk difference decays and approaches 0 after about 800 valves are installed. Note that the rate of decay changes through the process. Specifically, the initial decay rate is high but decreases as the number of installed valves increases. For example, adding 200 valves into the system reduces the risk difference from 3.24 GPM to 0.95 GPM, a significant 70.7 % risk reduction rate. However, adding another 200 valves (400 added valves in total) reduces the risk difference from 0.95 GPM to 0.46 GPM, a less significant but still large 51.5 % reduction rate.



(a) System risk changes through the valve adding process

(b) Risk difference between the constrained and the optimal configurations

Figure 7.14: System risk changes of the valve adding procedure for the Alameda Island WDN

To understand and validate the proposed valve adding algorithm, the suggested locations of additional valves for the Alameda Island WDN are visualized, shown in Figure 7.15. Note that only the locations of the first 200 added valves are plotted herein. As discussed previously, the risk reduction rate decreases as the number of added valves increases. In other words, early added valves are more critical to decreasing system risk than valves added in a later stage. Therefore, analyzing the locations of early added valves may illuminate the proposed algorithm's decision-making process.

On a global scale, a general spatial pattern for added valves locations can be observed from Figure 7.15. Although valves are added all over the network, sparsely valved regions (before the valve adding process) are more augmented than other areas. The zoomed region in Figure 7.15 shows how added valves are configured on a local scale. The algorithm automatically identifies insufficiently valved intersections and suggests new valve installation around them. Since the number of available valves to add is limited, not all intersections have valves installed in every direction. Pipes with high failure probability or demand are more

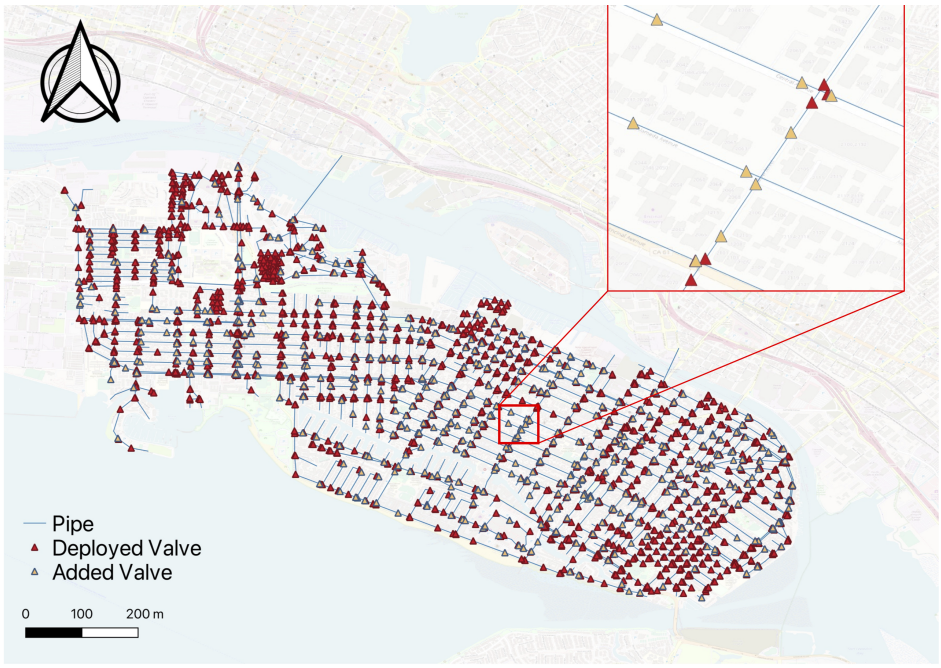


Figure 7.15: Location of the added valves for the Alameda Island WDN

heavily valved than other pipes. In summary, the proposed valve adding locations justifies the risk reduction result from the hydraulic simulation, which validates the effectiveness of the proposed method.

Round hill WDN

EMBUD Round Hill WDN is a small branch-like WDN serving a community in the East San Francisco Bay Area (City of Alamo). It has 126 internal nodes and 126 pipes for a total length of around 2.1 km. The majority of the pipes are less than 150 ft. There are a few exceptionally long pipes (longer than 200 ft) for water transmission purposes. The majority of the network is made with steel (61%) and asbestos concrete (36%). Currently, there are 48 valves installed in the system, as shown in Figure 7.16. The demand distribution for the network is non-uniform distributed. About 40% of pipes have very small demand (< 0.5 GPM). The demand for the rest of the pipes ranges from 0.5 to 2.0 GPM. Only a few pipes have demands that are greater than 2.0 GPM. See Chap 6 Section 6.2 for more details of the network.

Effectiveness of the proposed valve placement strategy

The first goal of the case study is to compare the pipe isolation risk between a system with the currently deployed valve placement configuration and a system with a valve placement

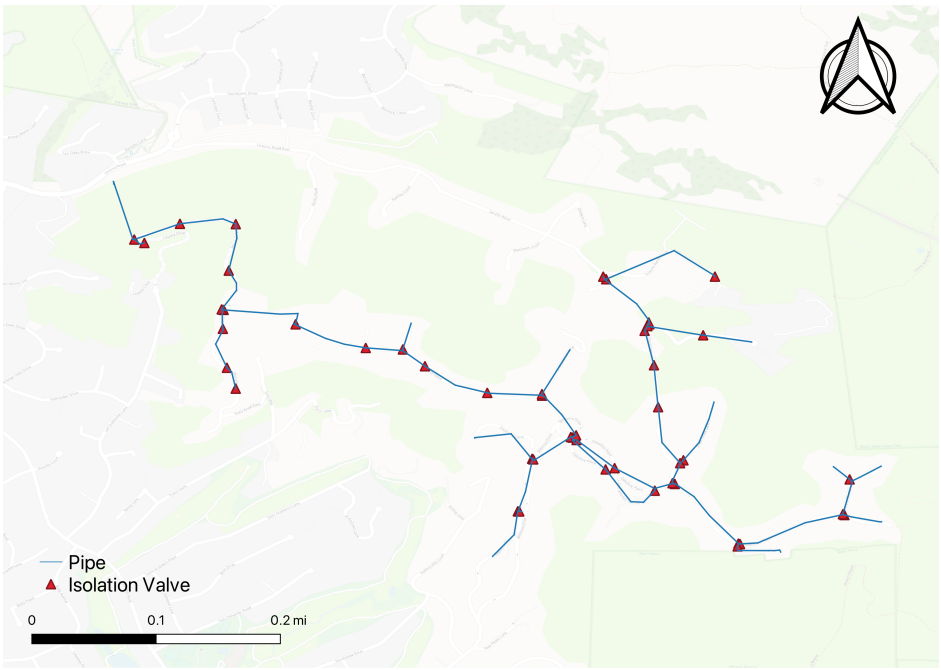


Figure 7.16: Location of deployed isolation valves for the Round Hill WDN

configuration generated by the proposed method. Figure 7.17 compares the location of valves from the proposed approach to the original one for the same number of valves. The proposed method considers the overall topology of the WDN when placing valves. For example, considering pipe $P0$ in the enlarged region of the network. Due to the sparse connective nature of the network, the removal of $P0$ disconnects many parts of the network from the water source. However, under the original valve placement configuration, isolation of $P1$ or $P2$ includes $P0$, creating high isolation cost for the system. On the other hand, the proposed method placed four valves around $P0$. Therefore, isolation of surrounding pipes will not include $P0$, which decreases the overall risk for the network.

Pipe isolation risks of the Round Hill WDN configured with both the currently deployed valve placement configuration, and the one generated by the proposed method are simulated and compared. Note that the number of installed valves is the same for all cases. Table 7.5 summarizes the results. Comparing to the current valving configuration, the proposed method's valve placement configuration reduced the overall system pipe isolation risk by 30%, from 9.61 GPM to 6.67 GPM. Different valve placement strategies result in distinct isolation segments properties. The mean value of isolation segment sizes of the system using the proposed valve configuration is similar to the value of the system with the original configuration. However, the variance of the segment sizes is different. The system with the proposed valve placement configuration has more size variances of isolation segments than the original setting. The variances difference of segment sizes is due to the fact that Round

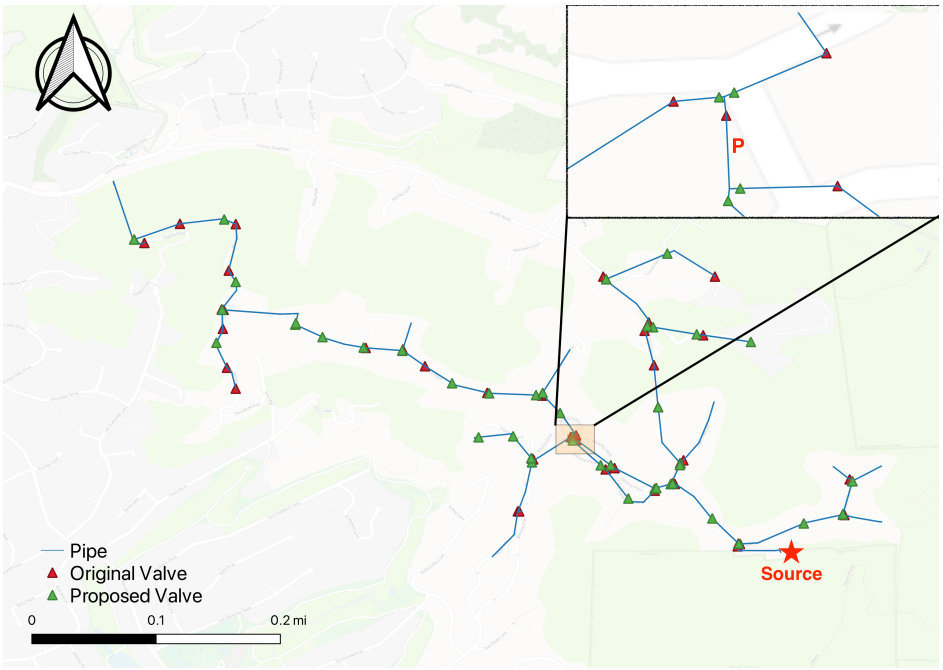


Figure 7.17: Location of proposed isolation valves for the Round Hill WDN

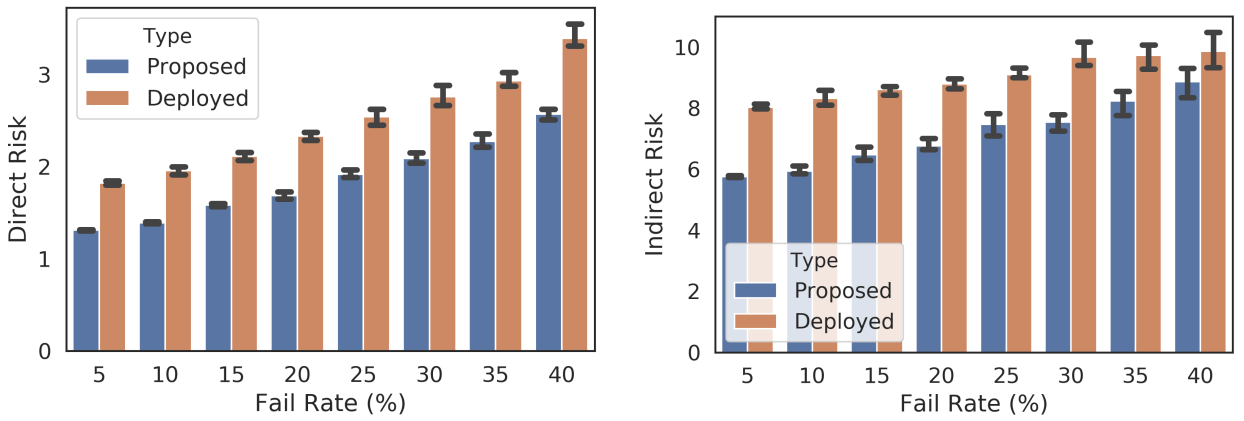
Hill WDN has non-uniform demand distribution. Placing valves around pipes with higher demands than other pipes reduces the overall system pipe isolation risk, thus increasing the variance of isolation segment sizes across the network.

Table 7.5: Network performance under different valve placement strategies for Round Hill WDN

Valve Placement Strategy	Number of valves	Average Segment Size	Segment Size Variation	Pipe Isolation Risk(GPM)
Current Strategy	48	3.31	3.58	9.61
Proposed Strategy	48	3.18	4.51	6.67

To evaluate the consequence of having missing/malfunctioning valves, network risks under 5%, 15%, 25%, 30%, 35%, and 40% valve failure rates are generated for both valve placement configurations using the same MC simulation method described in the previous case study. The result is shown in Figure 7.18, and the statistics are summarized in Table 7.6. Note that the total system pipe isolation risk is decomposed into two sub-risk to indicate different aspects of isolation consequence as the previous case study. Overall, the network configured with the proposed valve placement strategy has a smaller direct and indirect pipe isolation

risk than the one configured with the currently deployed valve placement strategy. For example, when the valve failure rate is 10%, the system using the currently deployed valve placement configuration has a direct risk of 1.96 GPM and an indirect risk of 8.32 GPM. On the other hand, the direct risk and indirect risk are 1.39 GPM and 5.94 GPM for the system configured by the proposed valve placement strategy. In other words, the proposed valve placement strategy reduces the total system risk by 28% when the valve failure rate is 10%. The standard deviation (STD) of both risk measures for the proposed strategy is also smaller than the original strategy for all tested valve failure rate scenarios. On average, the STD of the total pipe isolation risk for the network with the proposed valve configuration is 24% less than the network with the original valve configuration. In summary, the proposed valving placement configuration increases the system's ability against pipe isolation risk under valve condition uncertainties compared to the original configuration.



(a) Direct pipe isolation risk

(b) Indirect pipe isolation risk

Figure 7.18: Comparison of the system pipe isolation risks under different valve placement strategies for Round Hill WDN

As discussed previously, it is a general phenomenon that, as the valve failure rate increases, both the mean and standard deviation of the pipe isolation risk of the system increase with accelerating increasing rate due to network effects. In this case, the risk-increasing degrees of direct risks are similar for the two tested valve placement strategies. Specifically, the mean value of direct risk increases by 1.34 GPM as the valve failure rate increases from 0% to 40% for the system with the proposed valve placement configuration. On the other hand, the mean value of direct risk increases by 1.67 GPM of the same range of valve failure rate for the system with the currently deployed setting. The risk-increasing degree of indirect risk for the network configured with the proposed valve placement strategy is greater than the one configured with the original setting. The indirect risk raises by 1.99 GPM for the system with the original valve configuration (valve failure rate 0% to 40%). This number decreases to 3.16 GPM for the system with the proposed valve configuration. The fast-rising

of indirect risk for the proposed valve configuration is due to the sparse connected topological nature of the Round Hill WDN. Since the available water paths for every pipe to the source in this network are scarcer compared to densely connected networks such as the Alameda Island network, each valve failure event has a large chance of disconnecting pipes from the water source, increasing the overall indirect pipe isolation risk of the system. Therefore, the indirect risk between different isolation segments tends to converge as the valve failure rate increases. Considering that the total number of installed valves is small (only 48 valves), the differences of indirect isolation risk between the two valve configurations becomes small when the valve failure rate is large.

Table 7.6: Statistics of system pipe isolation risks under different valve placement configurations with various valve failure rates for Round Hill WDN

Valve failure rates	Valving placement strategy	Mean direct risk (GPM)	Direct risk STD	Mean indirect Risk (GPM)	Indirect risk STD
0%	Current	1.73	0	7.87	0
	Proposed	1.23	0	5.71	0
5%	Current	1.82	0.11	8.03	0.42
	Proposed	1.31	0.04	5.76	0.4
10%	Current	1.96	0.19	8.32	0.64
	Proposed	1.39	0.07	5.94	0.56
15%	Current	2.11	0.21	8.61	0.95
	Proposed	1.58	0.08	6.48	0.87
20%	Current	2.33	0.24	8.8	0.93
	Proposed	1.69	0.12	6.77	1.93
25%	Current	2.54	0.31	9.1	15.92
	Proposed	1.91	0.2	7.47	13.52
30%	Current	2.76	0.42	9.67	27.48
	Proposed	2.09	0.26	7.54	12.45
35%	Current	2.94	0.46	9.73	22.3
	Proposed	2.28	0.3	8.23	23.46
40%	Current	3.4	0.64	9.86	15.91
	Proposed	2.57	0.32	8.87	12.6

Lower bound of the system risk

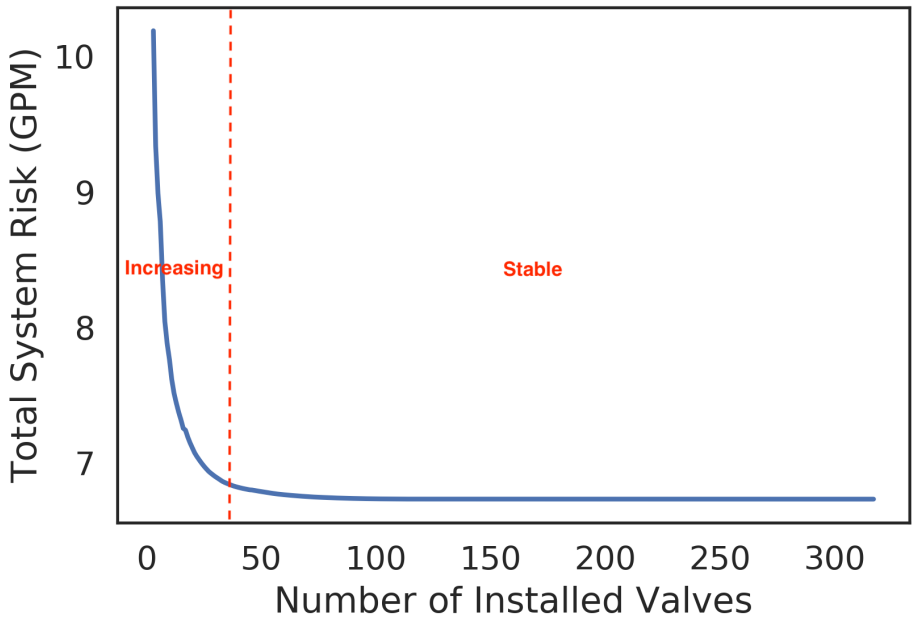


Figure 7.19: The lower bound of pipe isolation risk for the Round Hill WDN with respect to the number of isolation valves

Using the proposed valve placement method, the lower bounds of the pipe isolation risk for the Round Hill WDN considering various valves are computed and shown by Figure 7.19. Like the Alameda Island case study demonstrates, there are two distinct phases for system pipe isolation risk concerning the number of installed valves in the Round Hill WDN. The lower bounds stay almost the same regardless of the number installed after a certain number of valves are installed in the system. In this case, the threshold number is about 35 valves. On the other hand, if there are not enough valves installed (smaller than 35 valves), the total pipe isolation risk of the system increases dramatically. For instance, decreasing the number of installed valves from 30 to 15 increases the pipe isolation risk of the system from 6.67 GPM to 12.33 GPM, an 85% increase. Therefore, it is important to have at least 45 valves installed for the Round Hill WDN to achieve a small pipe isolation risk for the system. The more valves installed (in a proper configuration), the more robust the system is (far away from the phase change threshold).

Comparing Figure 7.19 with Figure 7.11, it is clear that the portion of the stable region for the lower bound of pipe isolation risk is larger for the Round Hill WDN than the Alameda Island WDN. Consequently, it requires relatively fewer valves (after adjusting for the size of the network) to achieve the optimal pipe isolation risk for the Round Hill network comparing to the Alameda Island WDN. Such difference is caused by the topological differences of the

two networks. Comparing to the Alameda Island WDN, Round Hill WDN is much sparsely connected. As discussed in Chap 5 Section 5.3, both the mean and variance of isolation segments of the system increase slower for sparsely connected network comparing to a densely connected one when valve failure rate increases. Therefore, as long as some "key" valves are presented in the network, which is typically around pipes with exceptionally high demand or prominent network positions (isolation such pipe disconnect a large portion of networks from the system), the pipe isolation risk of the Round Hill network stays low. On the other hand, since the average isolation segment size increases fast for densely connected networks, the Alameda Island WDN, as the number of available valves decreases, system isolation risk inevitably increases regardless of the valve configuration when a large number of valves are presented.

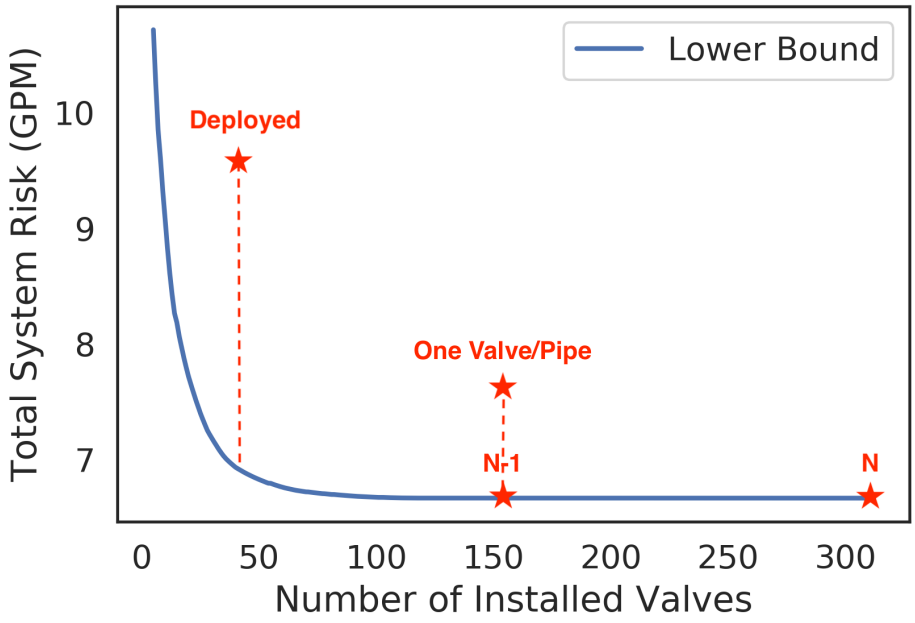


Figure 7.20: System pipe isolation risks of different valve placement strategies for the Round Hill WDN

Pipe isolation risks of the system equipped with different rule-of-thumb valve placement strategies are tested for the Round Hill WDN. Figure 7.20 shows the simulation result, and Table 7.7 lists the detailed numbers. Both the N and N-1 valve placement strategies achieve the lower bound of pipe isolation risk for the number of installed valves. Again, the difference between the two is that the N valve placement strategy provides more valve redundancy. Although the one valve per pipe strategy uses the same amount of valves as the N-1 valve placement strategy, it produces a system with a much large isolation risk. Specifically, the total system risk is 7.99 GPM, which is 16.5% higher than the risk lower bound (N-1 valve placement strategy). The deteriorated performance is because randomly

place valves on each pipe does not guarantee that each pipe will be isolated individually. In other words, the averaged isolation segment size is bigger for the one pipe per valve strategy than the N-1 strategy, causing the larger system pipe isolation risk. Lastly, the currently deployed strategy is far from optimal. Using the proposed method on placing valves for the system requires valves to achieve the same level of pipe isolation risk level, 9.61 GPM.

Table 7.7: Pipe isolation risks for the Round Hill WDN with different valve placement configuration

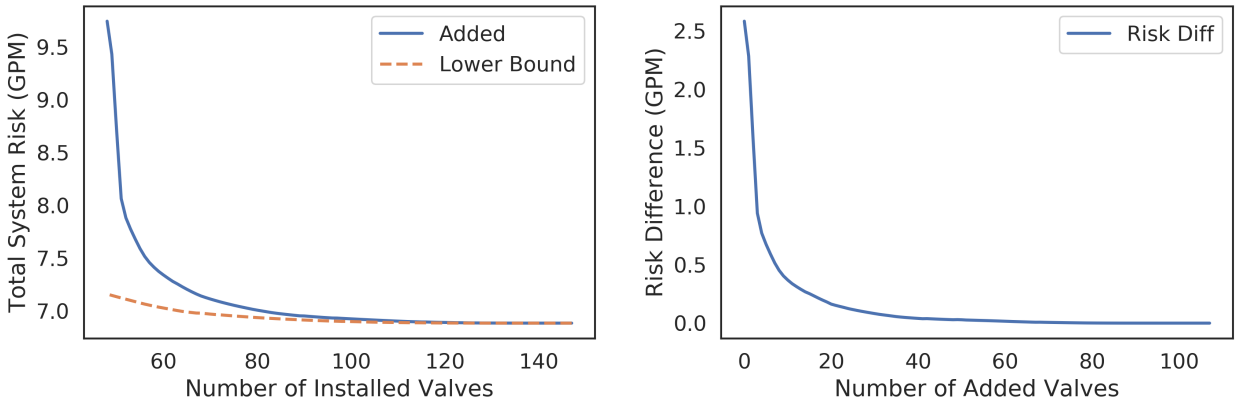
Valve Placement Strategy	Number of Valves	Direct risk (GPM)	Indirect risk (GPM)	Total risk (GPM)	Risk lower bound (GPM)
N	324	0.58	6.09	6.67	6.67
N-1	156	0.58	6.09	6.67	6.67
One valve per pipe	156	0.78	7.21	7.99	6.67
Current Strategy (deployed)	48	1.73	7.87	9.61	6.94

Additional valve installation

The proposed valve adding algorithm is tested for the Round Hill WDN. Starting from a system with the currently deployed valve placement strategy (48 valves), valves are added one by one until all pipes have two valves installed in both ends.

Figure 7.21a compares the pipe isolation risk of the resulting valve configuration with the risk lower bound at different levels of installed valves. Similar to the Alameda Island WDN case, the risk gap is significant at the initial stage, meaning the currently deployed valve placement strategy does not effectively minimize system pipe isolation risk. Adding more valves according to the proposed method shrinks the gap, and the two measures converge at the end. Figure 7.21b shows the relation between the risk difference and the number of added valves. The risk difference drops dramatically and quickly settles to 0 as more valves are added into the system. Specifically, adding only 10 additional valves into the system drops the risk gap from 2.58 GPM to 0.37 GPM, a remarkable 85.6% risk reduction. After installing 25 additional valves, the system's risk is insignificant to the optimal configuration's risk (lower bound). The risk gap is smaller than 0.1 GPM herein.

Figure 7.22 shows the locations of added valves for the round hill system. Since system risk approximates to the optimal level after 25 additional valves are installed into the system following the proposed method, only the first 25 added valves are shown for clear visualization. On a global scale, the algorithm successfully identifies pipes that previously lacked valves, and places new valves on them. Moreover, the algorithm considers the network's



(a) System risk changes through the valve adding process

(b) Risk difference between the constrained configuration and the optimal configuration

Figure 7.21: System risk changes of the valve adding procedure for the Round Hill WDN

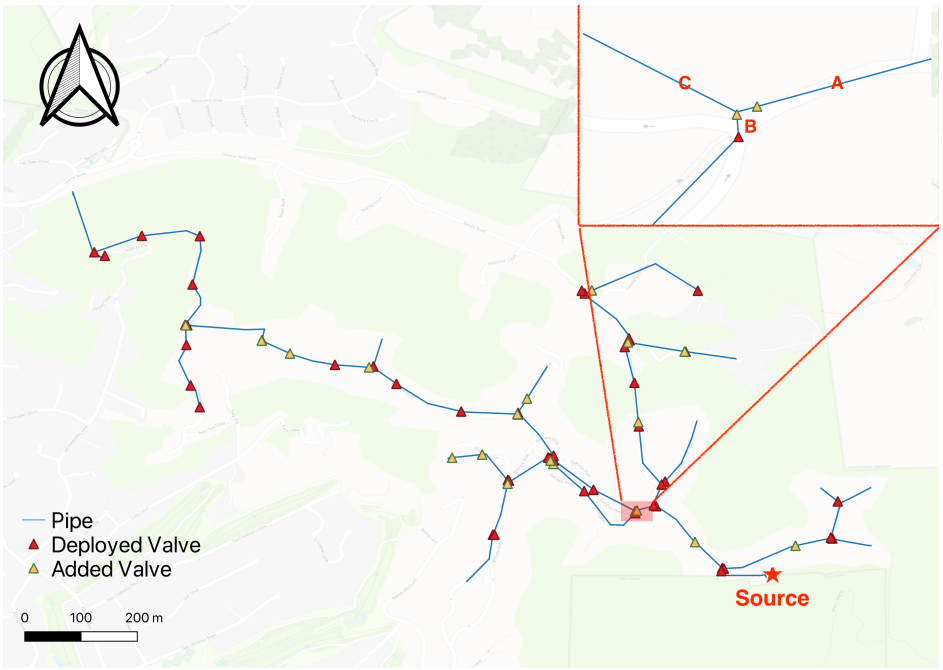


Figure 7.22: Location of the added valves for the Round Hill WDN

topology on choosing new valves locations. Considering the zoomed-in region in Figure 7.22 for an example. Pipe A is topological critical since it is the only pipe connecting a large part of the system (west to the pipe) to the water source. However, under the currently deployed valve configuration, Pipe B or Pipe C isolation includes Pipe A, causing significant impact to the system (west part of the system loses water source). Placing two more valves at the junction of the three pipes solves the problem, as suggested by the proposed algorithm. In this case, all three pipes can be isolated independently, which greatly reduces the system risks.

In summary, the proposed valve adding algorithm works well on the Round Hill WDN. It adds valves to the critical locations of the network to reduce the system pipe isolation risk. The effectiveness of the proposed valve configuration is quantified through hydraulic simulation, where system risk quickly settles to the lower bound as more valves are added to the system.

7.4 Summary

This chapter proposes an optimal valve placement method using dynamic programming. Compared to previous complicated optimization-based methods [52, 32, 93], the proposed algorithm produces optimal solutions while being intuitive and thus easy to follow. In addition, many widely used empirical valve placement strategies [92, 103, 136, 139] can be successfully interpreted and reproduced using the proposed method.

The general idea of the proposed method is to gradually remove the least impact valve from the most valve-redundant system until the desired number of valves is reached. The proposed method led to a system with minimal pipe isolation risk for all valve placement configurations with the same number of installed valves, which can be mathematically proved using induction. In addition to finding the best locations to install valves for a new system, the proposed method can be modified to find the best place to add new valves for an established system.

Although the idea of the proposed method is relatively simple, a naive implementation is proved to be computationally infeasible. An efficient implementation is developed to address this problem by eliminating redundant computation through graph operation and memorization.

The proposed valve placement method is tested on two real-life WDNs, the densely connected Alameda Island WDN and the sparsely connected Round Hill WDN. In both cases, the network with the valve placement configured by the proposed method produces much less pipe isolation risk than the currently deployed valve placement setting. Moreover, when valve conditions are uncertain, the system configured with the proposed strategy is more robust than the original one: both the expected mean value and variance of pipe isolation risk are low for every tested valve failure rate scenario.

Using the pipe isolation risk calculated from the proposed method, the lower bound of pipe isolation risk for a given number of valves is calculated for both cases. The behavior of

the risk lower bound can be categorized into two phases. When sufficient valves (phase change threshold) are installed in the system, the risk lower bound remains the same regardless number of installed valves. On the other hand, if the number of installed valves is smaller than the phase change threshold, the risk lower bound increases considerably. The exact value of the phase change threshold depends on both the topology and demand distribution of a WDN. The pipe isolation risk using commonly used rule-of-thumb valve placement strategies, the N, N-1, one pipe per valve strategies are simulated and compared with the risk lower bound. It is found that both N and N-1 valve placement strategies achieve the risk lower bounds for the number of valves used. The one pipe per valve strategy produces a higher system pipe isolation risk comparing to the N-1 valving strategy, even though the number of required valves are similar for the two strategies.

With slight modification, the proposed isolation placement method can be used to add additional valves into a system with deployed valves to reduce system risk. Instead of obtaining the minimal risk valve configuration across all possible configurations, the modified algorithm finds the minimal risk configuration under the constraint that the already deployed valves must be included. The effectiveness of the valve adding algorithm is validated for both cases. The system pipe isolation risk decreases and converges to the risk lower bounds as more valves are added to the system. The convergence speed is faster for the Round Hill WDN than the one in the Alameda Island WDN. As a sparsely connected network, Round Hill WDN has critical topological regions where removing a few pipes disconnects a significant part of the network from the water source. The proposed algorithm sharply reduces the overall system risk by identifying and adding more valves to such regions. On the other hand, there are few such critical regions in the Alameda Island WDN, a well-connected network. As a result, the system risk reduction speed is relatively slower than the Round Hill network as more valves are introduced according to the proposed method.

Chapter 8

Valve Improvement Strategy for System Pipe Isolation Risk Reduction

8.1 Overview

The previous chapter introduces an optimal valve placement algorithm to minimize the pipe isolation risk of a WDN for a given amount of valves. The valve placement method improves WDN resilience at the planning phase: it produces a valve configuration that reaches the lower bound of pipe isolation risk for the WDN. This chapter considers the improvement of WDN resilience at the operational phase, where valves have already been installed into the system.

As the valves in a WDN ages, their probability of failure rises due to various factors (e.g., soil corruption, material fatigue, etc.), increasing the pipe isolation risk of the WDN. Thus, it is necessary to perform regular maintenance (valve exercising, valve replacement, etc.) on the installed valves to reduce such a risk. Ideally, regular valves inspection should be performed to update the status of the valves on service frequently. Infrequently used valves should be exercised, and malfunctioning valves should be replaced on time. However, utilities often face many practical constraints to execute the ideal plan. Modern WDNs often contain an enormous amount of valves, making the cost of performing frequent valve maintenance activities to be significant. Considering their limited time and resources, it is often impractical for the utilities to carry out the ideal valve maintenance plan. Essential tasks such as water quality maintenance, aging pipe replacement, and customer services often require more attention from them comparing to valve maintenance. As a result, it is desirable to find a maintenance scheme that selects particular valves to maintain while decreasing the system pipe isolation risk as much as possible.

This chapter develops an efficient valve maintenance strategy that minimizes the system pipe isolation risk under valve condition uncertainties. Based on the insights about the relationship between valve failure rate and system pipe isolation risk (Chapter 5), maintenance priorities are assigned to valves in the system. Only the high-rank valves, which help reduce

the total system pipe isolation risk while valve conditions deteriorate, are recommended to be maintained under resource constraints. Code for the developed algorithm can be found in `https://github.com/rewu1993/wdn_valves`.

8.2 Factors that Influence System Pipe Isolation Risk

To devise an efficient valve maintenance strategy, one needs to understand the factors that influence the system pipe isolation risk under uncertainties of valve conditions. The entire Chapter 5 is dedicated to explaining the relationship between the valve failure rate and the total system pipe isolation risk. The key findings are summarized in the following equation:

$$\begin{aligned} \bar{D} &= \sum_{j=1}^{C} (\bar{x} + b_j)^2 \bar{f}_j \bar{l}_j \\ &= \bar{x}^2 \sum_{j=1}^{C} \bar{f}_j \bar{l}_j + \sum_{j=1}^{C} \bar{f}_j \bar{l}_j b_j^2 + 2\bar{x} \sum_{j=1}^{C} \bar{f}_j \bar{l}_j b_j \end{aligned} \tag{8.1}$$

where $\bar{D}$ is the total pipe isolation risk of the WDN, $\bar{x}$ is the average isolation segment size (number of pipes inside a segment) across the system, b_j is the size deviation of segment j to the mean size ($b_j = x_j - \bar{x}$), $\bar{f}_j$ is the mean pipe fail probability for M pipes in segment j, and $\bar{l}_j$ is the mean pipe cost for M pipes in segment j.

Equation 8.1 illustrates that two major factors determine the level of pipe isolation risk of the system. The first factor is the properties for the sizes of isolation segments in the system, the mean size $\bar{x}$ and size deviations b_j. The other one is the segment-wise pipe isolation risk $\bar{f}_j \bar{l}_j$, which depends on the network topology and distributions of both pipe failure rate and user demands. The meanings of the terms in Equation 8.1 are illustrated as follows. The first term of Equation 8.1 can be interpreted as the system risk when the network is assumed to be divided into equal size segments. Note that the magnitude of this term depends on the square of segment size $\bar{x}$, implying the significance of limiting the isolation segment size on reducing system risk. The real system may not be configured by equal-size isolation segments, and the last two terms of 8.1 account for this fact. Under typical conditions where all pipes are allowed to fail ($\bar{f}_j > 0$), and every failure has a certain consequence ($\bar{l}_j > 0$), the $\bar{f}_j \bar{l}_j$ term is positive. In this case, the second term is always non-negative, meaning that having segment size variations across the network will always increase the magnitude of this term. The last term captures the effect of variance of isolation segment size at a more detailed level. Under every pipe can fail and each failure has consequence condition, a segment with a size larger than the average size ($b_j > 0$) will increase the system risk and vice versa.

Based on the previous discussion, an effective valve maintenance strategy should aim at two purposes. First, reduce mean ($\bar{x}$) and variances (b_j) of the size of isolation segments, which can be achieved by prioritizing the maintenance of valves that can lead to a large

increment of the size of the isolation segment when fail. Second, limit the segment-wise risk $f'_j\bar{l}_j$ across the system. In this case, the maintenance of the valves around the pipes with large isolation risks (high failure rate, large demand) should be prioritized. The following sections discuss the way of achieving these two goals in detail.

Method to control isolation segment size properties

The Segment-Valve graph of a WDN fully captures the relationship between valves and isolation segments (see Chapter 5 for details). Since edges in the Segment-Valve graph represent valves in the system, failure of an valve results in the removal of the corresponding edge, which changes the topological characteristics of the whole graph. In other words, the Segment-Valve graph evolves as valves fail. As discussed previously, Segment-Valve graphs with small mean and variance on isolation segment sizes have less system pipe isolation risk comparing to others. Since failures of different valves have different consequences, one may reduce the chance of the system being at an undesirable state (large mean and variance for segment sizes) by performing valve maintenance on certain valves. In this way, the maintained WDN becomes less vulnerable regarding pipe isolation risk than the system before valve maintenance.

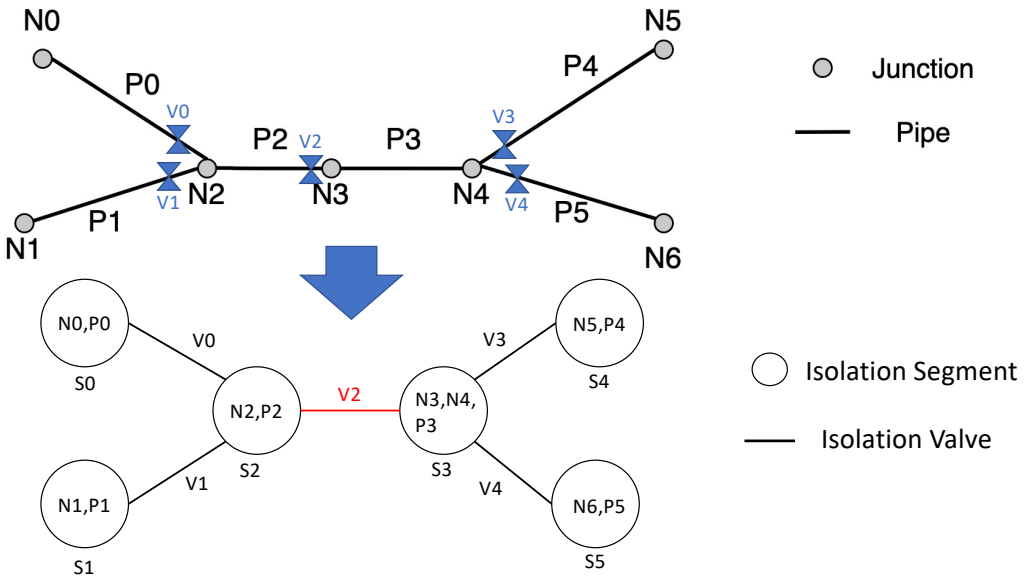


Figure 8.1: An example network for isolation valve priority illustration

Figure 8.1 shows an example that maintaining certain valves with priority can reduce the total system pipe isolation risk when valve conditions deteriorate. The example network contains 6 pipes, 7 junctions, and 5 valves. Each pipe is assumed to have the same failure

probability ($\frac{1}{6}$) and failure cost (C). The valves are placed according to the common-used rule of thumb N-1 valve placement strategy, where every junction with n pipe legs has n-1 valves installed around it. Using the technique introduced in Chapter 5 Section 5.1, the original Pipe-Junction representation of the network is converted into the Valve-Segment representation. The Valve-Segment graph contains 6 isolation segments connected by 5 valves.

When every valve functions properly, the total pipe isolation risk for the network is C. In this case, every isolation segment contains a single pipe. Consequently, the pipe isolation risk for each isolation segment is the same since each pipe has the same failure probability and consequence. However, even under such an equal isolation risk (segment-wise) condition, failure impacts of valves in the system are different.

In this example, valve $V2$ is more important than the other four valves in controlling the total system pipe isolation risk, which can be illustrated by a thought experiment. Suppose that only one of the five valves functions properly in the system (equivalent to four failed valves scenario). The Valve-Segment graphs for all cases are illustrated by Figure 8.2. Table 8.1 provides the statistics of the corresponding isolation segment size distribution and system pipe isolation risks for each scenario.

Among all five scenarios, the operation of valve $V2$ produces the least variance of isolation segment sizes. As a result, the system with $V2$ has the minimum pipe isolation risk comparing to others. The importance of the valve $V2$ lies in the structural significance of the corresponding edge in the Valve-Segment graph. As shown in Figure 8.1, edge $V2$ acts as the only "bridge" that links segments between the two sides. In other words, edge $V4$ must be included in the path to connect isolation segment $S0$ (or $S1$) with $S4$ (or $S5$). Therefore, by making sure the "bridge" is always closed (valve $V2$ won't fail), $S0$ (or $S1$) can never be merged with $S4$ (or $S5$). In summary, maintaining valve $V2$ provides the minimum pipe isolation risk compared to the other cases under the four failed valves condition.

Table 8.1: System pipe isolation risks under different choices of maintained valve

Valve to maintain	Mean segment size	Segment Size Variation	Pipe Isolation Risk
V0	3	4	4.33C
V1	3	4	4.33C
V3	3	4	4.33C
V4	3	4	4.33C
V2	3	0	3C

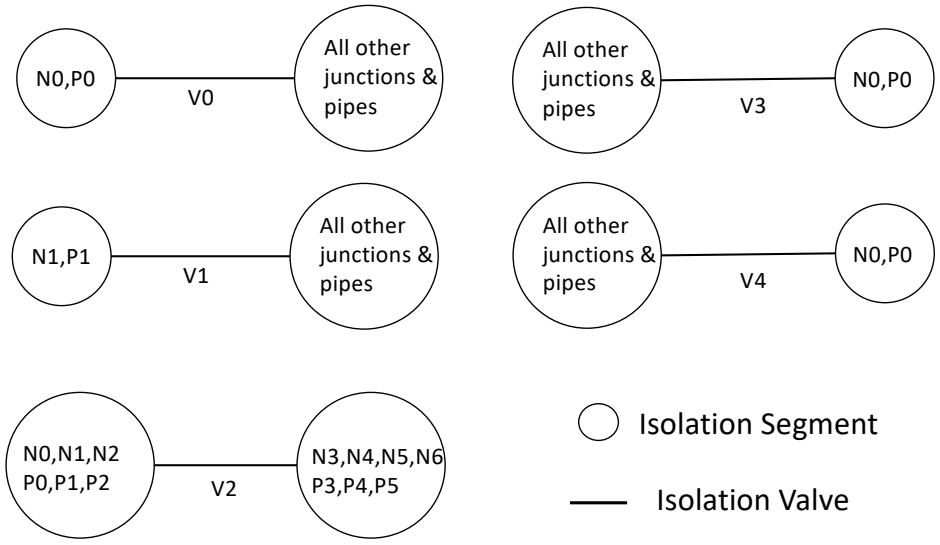


Figure 8.2: Valve-Segment graph of the network after the valve failure events under different choices of maintained valve

Centrality measures

To find valves to maintain in priority, one can enumerate all the possible scenarios and rank the valves based on their failure consequence. However, such a technique is only feasible for simple networks such as Figure 8.1. Real-life WDNs are complex. For example, pipes may have different properties (length, material, age, etc.), thus different failure probability. It is also common for a WDN to have unevenly distributed demand (isolation cost). Moreover, pipes' connectivity patterns vary significantly across different parts of the network, making risk calculation computational intensive. As a result, an efficient, automatic valve ranking method is proposed herein.

The key idea for the proposed valve ranking method is based on network centrality analysis. Centrality analysis provides measures of structural importance for network components regarding their network position [20]. Centrality analysis can be viewed as a function that maps each component of a network to a real value number, providing a ranking that identifies the significance of each component [18]. Network centrality analysis is widely used in many fields. For instance, social scientists used it to identify the most influential person(s) in a social network [21].

One of the centrality measures for network analysis, the edge betweenness centrality, is used to rank the importance of valves for a WDN regarding pipe isolation risk control. Edge betweenness centrality of an edge e is defined as the sum of the fraction of all-pairs shortest paths that pass through edge e [20]. Mathematically, the edge centrality for an edge e in a

network is defined as

$$c_B(e) = \sum_{s,t \text{ in V}} \frac{\sigma(s,t|e)}{\sigma(s,t)} \tag{8.2}$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest (s,t) paths, and $\sigma(s,t|e)$ is the number of those paths passing through the edge e.

In this study, the edge betweenness centrality is calculated for each edge in the Segment-Valve graph. Valves with large edge betweenness centrality values are ranked higher (to be maintained with high priority) than low-value ones. The logic for the ranking is based on the property of the Segment-Valve graph. Recall that each node represents an isolation segment, and each edge represents an valve in the Segment-Valve graph. Therefore, a path (sequence of edges) between two nodes in the Segment-Valve graph represents the failed valves required to merge the two segments. Following this interpretation, the shortest path between two nodes in the Segment-Valve graph represents the smallest set of failed valves that can merge the two segments. Since edge betweenness centrality measures the frequency of the edge being included in the shortest path between a pair of nodes, valves with high edge betweenness centrality are critical to the segments merging process in the network. Therefore, decreasing the failure probability of such valves through maintenance may prevent the spread of successive merges of segments in the network, leading to a network with smaller isolation segments compared to the non-maintained ones.

To illustrate the valve ranking method proposed above, considering the simple network in the previous section (Figure 8.1). Calculation of the edge betweenness centrality requires the knowledge of the total number of paths in the graph. Since the Segment-Valve is undirected (path s to t is the same as path t to s), the total number of paths can be calculated as the binomial coefficient C_r:

$$C_r = \frac{n!}{r!(n-r)!} \tag{8.3}$$

where n is the number of nodes (isolation segment) in the system, and r is the number of nodes that need to be drawn. Since only source and destination are need to be selected, r is equal to 2 in this case.

For the example network, the total number of isolation segments is 6. Using the equation above, it is found that the network has 15 paths in total. The edge betweenness centrality value for each edge can be found by counting its frequency in all paths. The result is summarized in Table 8.2. Valve $V2$ has a significantly larger edge centrality measure comparing to the other four valves. Therefore, valve $V2$ is ranked with the highest priority. All the other valves have the same priority following the valve $V2$. The ranking of valves calculated by edge betweenness measure matches the result obtained by enumerating all the possibilities of valve maintenance (discussed in the previous section), validating the effectiveness of the proposed method.

Table 8.2: Edge centrality for the example network

Isolation valve	Number of passing shortest path	Edge centrality
V0	5	$\frac{1}{3}$
V1	5	$\frac{1}{3}$
V3	5	$\frac{1}{3}$
V4	5	$\frac{1}{3}$
V2	11	$\frac{11}{15}$

Method to control segment-wise pipe isolation risk

Equation 8.1 suggests that restricting the mean ($\bar{x}$) and variance (b_j) of isolation segment sizes across the network is not the only way of minimizing the system pipe isolation risk when valve conditions deteriorate. Segment-wise pipe isolation risks ($\bar{f}_j\bar{l}_j$) also plays an essential role in risk calculation. Since pipes vary in failure probability (due to material, length, age differences, etc.) and failure consequence (due to uneven demand distribution and network topology), the isolation risk of each pipe might be different. In other words, segment-wise pipe isolation risks are spatially uneven distributed across the network. In this case, reducing the isolation chance of certain segments may be more effective in lowering system risk than other segments in a network, which can be achieved through a selective valve maintenance strategy.

Valves in two types of isolation segments need to be maintained with priority to improve system resilience. The first type is the isolation segment with exceptionally high isolation risk. There are several causes for an isolation segment to have high isolation risk. First, a large size isolation segment often has a large isolation risk due to the number of contained pipes inside it. The size of the isolation segment is determined by the conditions of valves around it. For instance, if a particular region of the network has a scarcer distribution of valves than the rest of the system, the sizes of isolation segments in the area tend to be larger than the average segment size of the system. However, two same-size isolation segments may have significantly different pipe isolation risks. The risk difference may be due to a) differences in pipe failure probability and b) differences in pipe failure consequence. For example, an old, long, cast-iron pipe crossing a major geological fault line (prone to earthquake damage) typically has a much greater chance of failure than a newly deployed PVC pipe on a region with no natural hazard risk. Consequence-wise, failures of pipes serving water-demanding industrial properties or critical public service buildings (hospitals, etc.) have severe consequences than pipes' failure in a low-density community. Maintaining valves in such isolation segments reduces the risk of accidentally isolating critical properties while isolating pipes in other parts of the system.

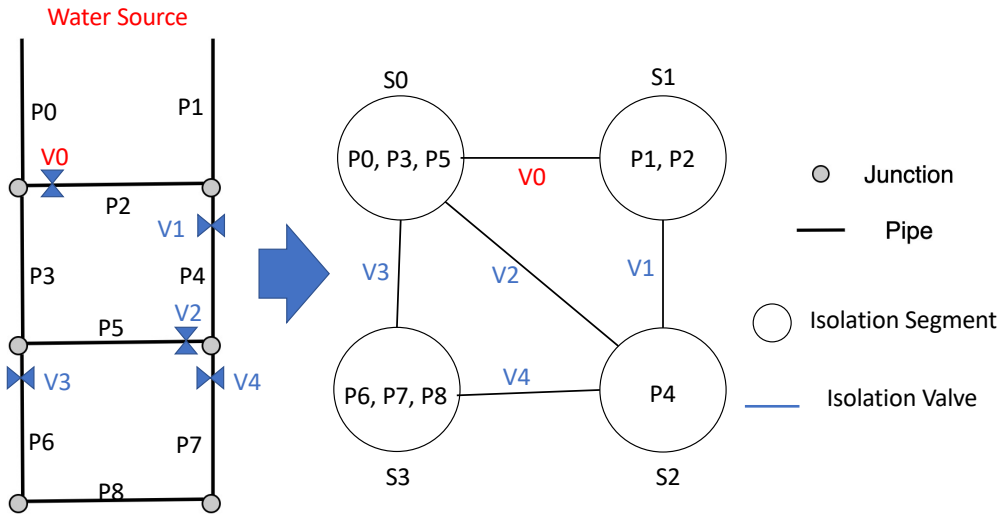


Figure 8.3: Example of unintended isolation risk due to valve failure events

Isolation segments with relatively small risk may also cause significant risk to a network as the valve failure rate increases. As discussed in Chapter 5 Section 5.4, certain network regions may have insignificant risk when the valve failure rate is low, but the risk increases dramatically as certain valves in the region fail. The pipe connectivity patterns of a region determine its vulnerable degree. Specifically, regions of a network with sparse pipe connections are more vulnerable than parts with dense pipe connections. In sparsely connected regions, isolating multiple pipes simultaneously (due to failed valves) may disconnect other parts of the system from the water source, leading to a significant pipe isolation risk increase for the system.

Considering Figure 8.3 for an example. The system consists of nine pipes, six junctions, and five valves. The system is connected to a water source through both pipe $P0$ and pipe $P1$. By closing all five valves, the system is divided into five isolation segments. Segments $S0$ and $S3$ contain three pipes, $S1$ has two pipes, and $S2$ has only one pipe. Assuming each pipe has the same probability of failure ($\frac{1}{9}$), and each failure has the same consequence C, the isolation risks for all the segments can be calculated, and the result is summarized in Table 8.3. Note that the total isolation risk of the segment is decomposed into direct isolation risk and indirect isolation risk. Since each isolation segment has more than one path to the water source, there is no indirect isolation risk for the system at this state. The total isolation risk for the system before any valve fails is $\frac{23C}{9}$. The total isolation risk of the system increases dramatically as particular valves fail, shown by Table 8.4. After

valve $V0$ fails, isolation segment $S0$ merges with isolation segment $S1$. Note that the direct risk of the merged segment ($\frac{25C}{9}$) is greater than the summation of the two components ($\frac{13C}{9}$) because both the failure probability and failure consequence increase for the merged segment. In addition to the increased direct risk, the failure of valve $V0$ introduces indirect isolation risk to the system. Isolating the newly merged segment disconnects the rest of the system from the water source, increasing the total system pipe isolation risk. Combing the two effects, the total pipe isolation risk for the system increases from $\frac{23C}{9}$ to $\frac{55C}{9}$, a 96% increase. On the other hand, the failure of the other valves does not produce any indirect pipe isolation risk for the system due to water path redundancy. In summary, the failure of valve $V0$ introduces more isolation risk for the system comparing to the failure of the other valves. Thus, prioritizing the maintenance of valve $V0$ is crucial to improve the robustness of the WDN against valve failure events in this case, even though the corresponding isolation segment has a relatively small isolation risk.

Table 8.3: Segment-wise pipe isolation risk for the example network prior to any valve failure event

Isolation segment	Direct isolation risk	Indirect isolation risk	Total isolation risk
S0	C	0	C
S1	$\frac{4C}{9}$	0	$\frac{4C}{9}$
S2	$\frac{C}{9}$	0	$\frac{C}{9}$
S3	C	0	C

Table 8.4: Segment-wise pipe isolation risk for the example network after valve $V0$ fails

Isolation segment	Direct isolation risk	Indirect isolation risk	Total isolation risk
S0 and S1	$\frac{25C}{9}$	$\frac{20C}{9}$	$\frac{45C}{9}$
S2	$\frac{C}{9}$	0	$\frac{C}{9}$
S3	C	0	C

Pipe risk map and clusters

Isolation segments with high pipe isolation risk can be identified using the multi-size pipe isolation risk map described in Chapter 5 Section 5.4. Recall that the multi-size pipe isolation risk map includes the risk of every pipe in the network with a set of predefined isolation

segment sizes. Therefore, pipes with high isolation risk, even when the corresponding isolation segment size is small, can be directly chosen from the risk map. The multi-size pipe isolation risk map can also be used to locate "vulnerable regions" of a WDN, which are subnetworks that experience rapid pipe isolation risk increase as the valve failure rate increases. Specifically, the changes of pipe isolation risk regarding segment size increase can be calculated using the multi-size pipe isolation risk map for every pipe in the system. A vulnerable region is then a spatial cluster of pipes that experience considerable risk increases as their corresponding isolation segment sizes rise (due to failures of valves). The task of identifying "vulnerable regions" then boils down to finding such pipe clusters automatically with precision.

In this study, a data clustering algorithm, the Density-Based Spatial Clustering (DBSCAN) [118], is applied to identify vulnerable regions for a WDN automatically. DBSCAN is a density-based non-parametric clustering algorithm. It automatically groups data closely packed together (high density) using user-specified distance measures (Euclidean distance, etc.). Comparing to other popular clustering algorithms such as the K-Means clustering algorithm [60], one advantage of DBSCAN is that it does not require the number of clusters as the input. Instead, the number of clusters inside a dataset is automatically generated through the clustering process. Since the number of vulnerable regions inside a WDN is unknown before the clustering process, DBSCAN becomes the perfect choice for the task.

The basic concept of the DBSCAN algorithm is summarized below. To find the clusters inside a dataset, the DBSCAN algorithm takes two input parameters from the user, the minimum distance (ϵ) and the minimum number of points ($minPts$) inside a cluster. ϵ specifics how close points should be to each other to count them belonging to the same cluster. If the distance between two points is smaller than ϵ, they are grouped into the same cluster. $minPts$ specifies the lower bound of cluster size. For instance, if the value for the parameter is 3, then only clusters with at least 3 points are valid clusters. The general steps for the algorithm are listed below [118].

1. For every data point, find the points that have a distance smaller than the minimum distance parameter, ϵ. Mark all the points that have more than $minPts$ neighbors as core points.

2. Find the connected components of core points on the neighbor graph.

3. For each non-core point, assign them to a nearby cluster if the distance is smaller than ϵ.

As previously discussed, the DBSCAN algorithm requires a clear definition of the "points" that need to be clustered and the distance measure between them.

In this study, "points" are pipes that experience large isolation risk (indirect) increase as their corresponding isolation segment sizes increase, which can be obtained from the multi-size pipe isolation risk map. The distance between the two pipes is defined as the minimum euclidean distance between them. See Figure 8.4 for examples of distance calculation between

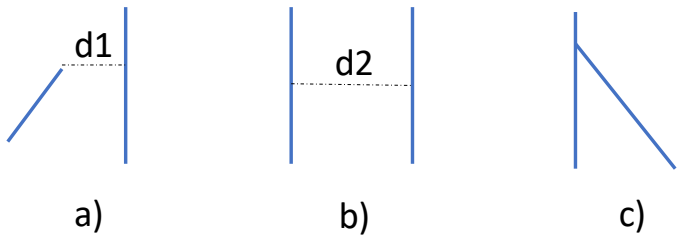


Figure 8.4: Examples for the calculation of distance between two pipes. a) non-parallel pipes that does not intersect (distance $d1$); b) Parallel pipes (distance $d2$); c) non-parallel pipes that intersects (distance 0)

a pair of pipes. Regarding input parameters for the DBSCAN, the minimum distance ϵ is set as the minimum pipe length of the network. The minimum number of points for a valid cluster ($minPts$) depends on the user's purpose. Using a small $minPts$ produces more clusters than the case with a large $minPts$. Therefore, if the user aims to choose a small set of valves to be maintained in priority due to budget constraints (time/money), a small $minPts$ value can be used, and vice versa.

Valve maintenance strategy

Based on the previous discussions, a valve maintaining strategy is proposed to improve WDN reliability regarding pipe isolation risks under isolation valve condition uncertainties. Figure 8.5 summarizes the proposed procedure of choosing valves to be maintained in priority. The workflow has two parts. The first part aims to control the distribution of isolation segment sizes across the network as valve conditions deteriorate. To achieve this goal, valves with high edge betweenness centrality measures on the Segment-Valve graph are selected to be maintained in priority. Reducing the failure probability of such valves (through maintenance) decreases the chance of having uneven distributed isolation segment sizes in the network, leading to smaller system pipe isolation risk as valves fail. In this study, a python program is developed to convert a WDN to the corresponding Segment-Valve graph automatically, and the edge betweenness centrality of every edge (valve) in the graph is calculated using Python package NetworkX [58]. Valves with high centrality measures are selected to be maintained in higher priority than low-value valves.

The second part of the maintenance procedure aims to control the segment-wise pipe isolation risk for the system. Two types of valves are chosen to be maintained in this stage. Valves that are related to pipes with high pipe isolation risk are selected first. Next, valves included in the "vulnerable regions", which are the subnetworks that experience rapid pipe

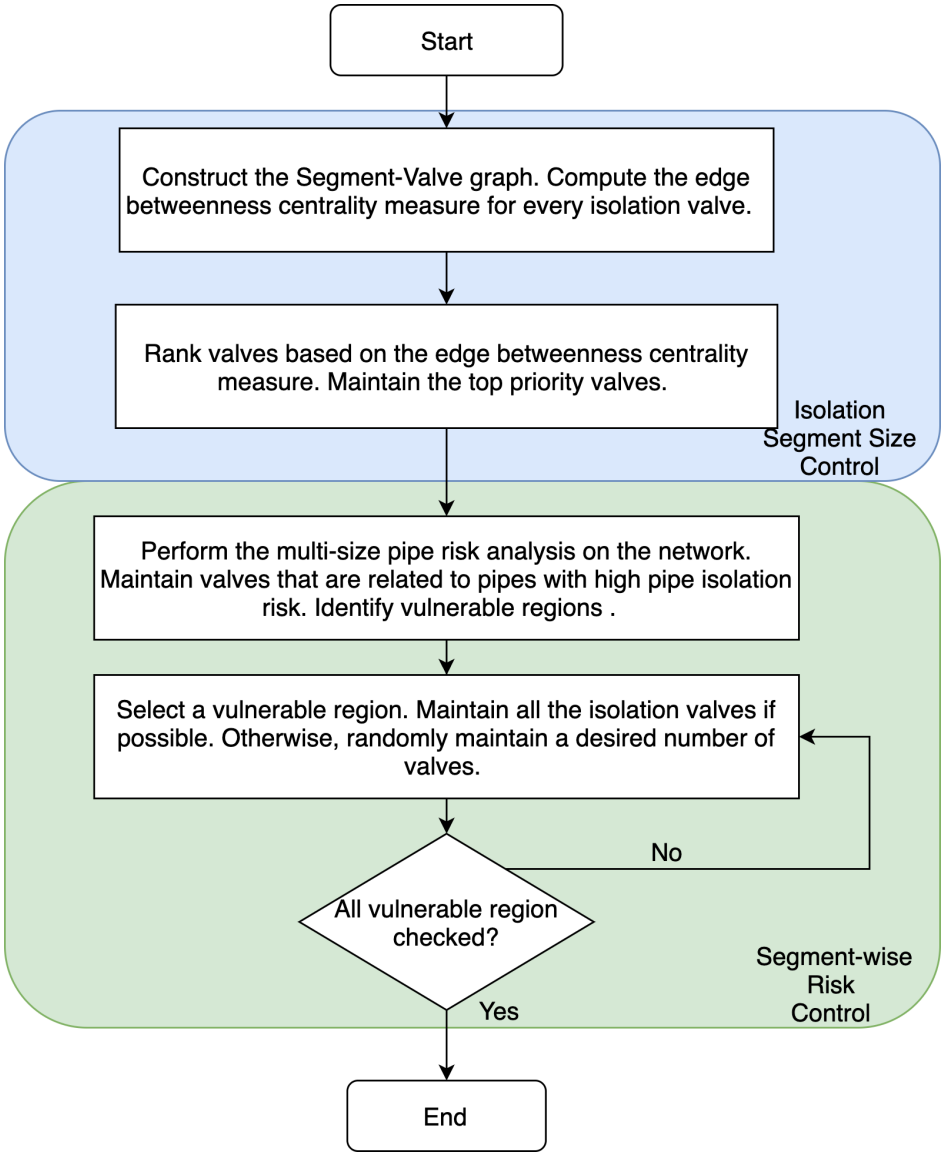


Figure 8.5: Flowchart of the proposed isolation valve maintenance strategy

isolation risk increase as the valve failure rate increases, are also chosen to be maintained. To make the selection, the multi-size pipe isolation risk map of the WDN is created using the procedure described in Chapter 5 Section 5.4 through hydraulic simulations. Since the multi-size pipe isolation risk map contains the isolation risk information for every pipe in the system regarding various isolation segment sizes, one can directly identify pipes with high isolation risks from it. All valves of such high-risk pipes are selected to be maintained.

To locate the "vulnerable regions" in the system, pipes with a significant rise in isolation risk as the corresponding isolation segment size increases are filtered out from the risk map. The DBSCAN clustering algorithm is applied to the chosen pipes, and all valid output clusters from the algorithm are identified as vulnerable regions of the network. Since the size of clusters can vary significantly, an adaptive sampling strategy is used to select valves to be maintained. Specifically, when the number of valves inside a vulnerable region is small, all the valves are selected to be maintained. However, for regions with a large number of valves, maintaining all valves in them may be impractical due to resource constraints. In such cases, only a portion of valves is randomly selected to be maintained. Randomly maintaining valves effectively decreases the risk of forming large-size segments in the region, reducing the probability of potential pipe isolation risk increase due to the blockage of water pathways from the isolation of large segments.

8.3 Case Study

This chapter uses the same case study networks introduced in Chapter 5 to test the effectiveness of the proposed valve maintenance method. The first case study considers the EBMUD Alameda Island WDN, a well-connected medium-size network. Next, the EBMUD Round Hill WDN, a small-scale sparsely connected network, is tested to illustrate the proposed method's effectiveness on networks with different scales and connectivity patterns.

Alameda island network

The EBMUD Alameda network is a medium-sized WDN with 2508 internal nodes and 3028 pipes for around 194 km. Currently, there are 1999 valves installed for the system. The network is well-connected with pipes of different lengths and materials. See Chap 6 Section 6.1 for details of the network description.

Centrality measures

The Segment-Valve representation of the Alameda Island WDN is constructed using the method described in 5 Section 5.1. With the constructed Segment-Valve graph, each valve's edge betweenness centrality measure is computed. Figure 8.6 shows the histogram of the edge betweenness measure of valves in the network. The majority of the valves (98.5%) have trivial edge betweenness centrality (< 0.1) in the Segment-Valve graph. Only 1.5% of valves have

significant edge betweenness centrality (> 0.1). The lack of high edge betweenness centrality valves in the network is partly due to the densely connected nature of the Alameda Island WDN. Since the majority of the network exhibits grid-like structures, there are multiple paths between one pipe (isolation segment) to another pipe (isolation segment). Consequently, only a few valves will be included multiple times in the shortest paths of the Segment-Valve graph. Hence, only a small amount of valves has a significant edge betweenness centrality measures in the network.

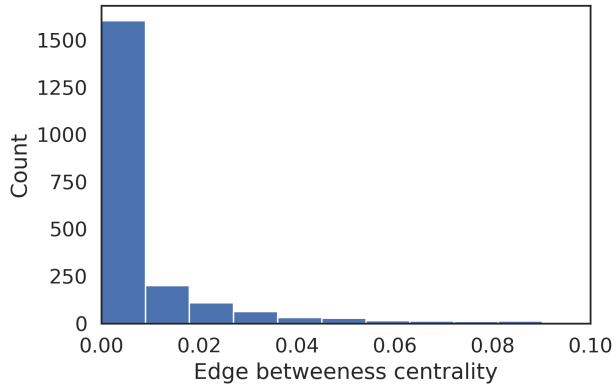


Figure 8.6: Distribution of edge betweeness centrality measures for valves in the Alameda island WDN

Figure 8.7 shows that the map of edge betweenness measure of valves for the Alameda Island WDN. Note the edge betweenness value of an valve is mapped to the corresponding pipe for visualization purposes. One primary observation from Figure 8.7 is that the valves located at the center of the network tend to have more prominent edge betweenness centrality than the valves at the other network parts. This is because connecting isolation segments located at the opposite edges of the network requires passing through isolation segments at the center. Therefore, the valves located at the center of the network are more likely to be included in multiple paths in the Segment-Valve graph comparing the valves at the other parts of the system. Note that not all central valves have a high edge betweenness centrality. The actual distribution of the centrality measure depends on the spatial distribution of the installed valves, which determines the connection pattern of the Segment-Valve graph.

Segment-wise pipe isolation risk

Figure 8.8a shows the generated multi-scale pipe isolation risk map of the Alameda Island WDN. To visualize the isolation risks of pipes regarding various isolation segment sizes, pipes are categorized into three classes. The pipes with high isolation risk (bigger than the average per pipe risk) when isolating with small segment size (0-5 pipes) are marked in red (first-degree vulnerable pipes). The pipes with high risk only when the corresponding

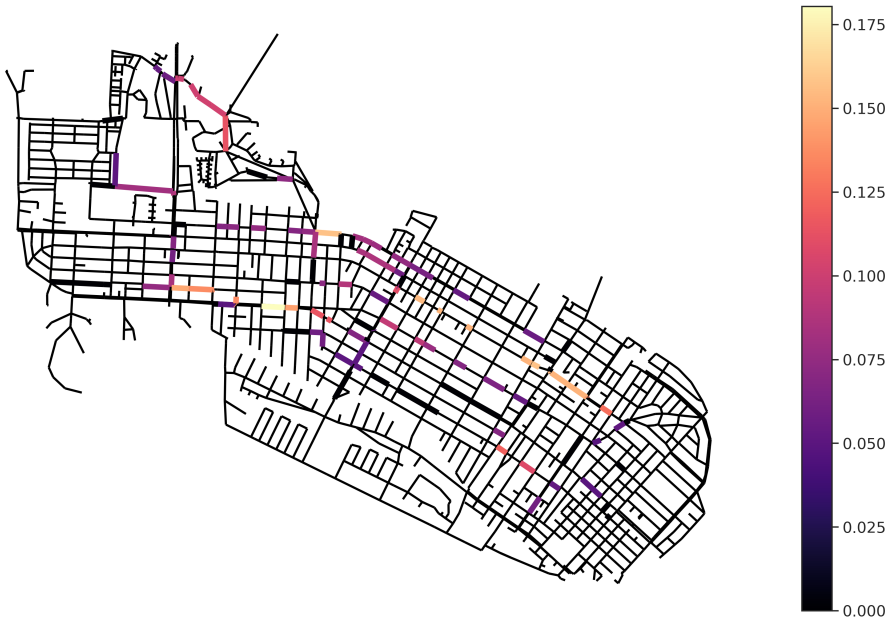


Figure 8.7: Edge betweeness centrality measure for each valve of the Alameda island WDN (mapped to the corresponding pipe for visualization purposes)

isolation segment size is large (5-30 pipes) are marked as yellow (second-degree vulnerable pipes). Pipes that have no risk regardless of tested segment sizes are colored in black. There are 45 first-degree vulnerable pipes in the system, with 32 corresponding valves.

The DB-scan algorithm is performed on all vulnerable pipes to find vulnerable regions of the network. The clustering result is shown in Figure 8.8b. Comparing Figure 8.8a with Figure 8.8b, all groups of vulnerable pipes in the network are successfully clustered by the DB-SCAN algorithm. Figure 8.9 shows the distribution of the size (number of pipes) of the identified clusters in the network. The network has 13 vulnerable pipe clusters in total. The largest cluster consists of 152 pipes, and the smallest one has only 6 pipes. On average, around 30 pipes are contained in a cluster. The size distribution of clusters is left-skewed. About half of the vulnerable regions contain very few pipes (< 10). Only three vulnerable regions contain an exceptionally large number of pipes (> 50)

Effectiveness of the maintenance strategy

The proposed valve maintenance strategy is tested on the Alameda Island WDN. The first step is to maintain valves with high edge centrality measures in the Valve-Segment graph. As illustrated in Figure 8.6, only a small amount of valves has a significant value of edge centrality measure in the Valve-Segment graph. Therefore, after ranking all the installed valves using the calculated edge centrality measure, only the top 5% are selected to be

(a) Pipe isolation risk map for the network

(b) Vulnerable regions of the network (using DB-SCAN clustering)

Figure 8.8: Pipe isolation risk map and the corresponding vulnerable regions for the Alameda Island WDN

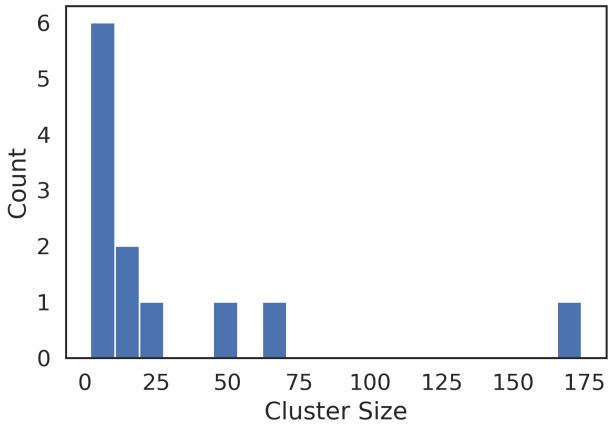


Figure 8.9: Distribution of the size (number of pipes) of vulnerable regions in the Alameda Island WDN

maintained in priority (100 valves). To control the segment-wise pipe isolation risk across the network, all valves that are installed on the first-degree vulnerable pipes are selected to be maintained (32 valves). The valves in the vulnerable regions are maintained using the adaptive sampling method. All valves in a small region (less than 20 valves) are chosen to be maintained. For large size vulnerable regions, only 20% of valves are randomly selected for maintenance. In total, 172 valves are selected for segment-wise pipe isolation risk control purposes. Figure 8.10 shows all pipes with valves for maintenance based on the proposed valve selection procedure.

Simulation is performed on the network to evaluate the effectiveness of the proposed valve maintenance strategy. Specifically, Monte Carlo (MC) simulation is used to estimate the pipe isolation risk of the maintained network with various valve failure rates. Valves are randomly selected from the pool of unmaintained valves to fail at each simulation run using the given valve failure rate. All maintained valves are assumed to be infallible (the probability of failure is 0%) during the simulation process. Hydraulics simulation is then conducted on the damaged network (with failed valves) to estimate the system's direct and indirect pipe isolation risk. The risk estimation process is repeated multiple times for a given valve failure rate scenario, with each run has distinct failed valve selections. The mean value across all the realizations (different sets of failed valves) is used as the pipe isolation risk of the network for a given valve failure rate.

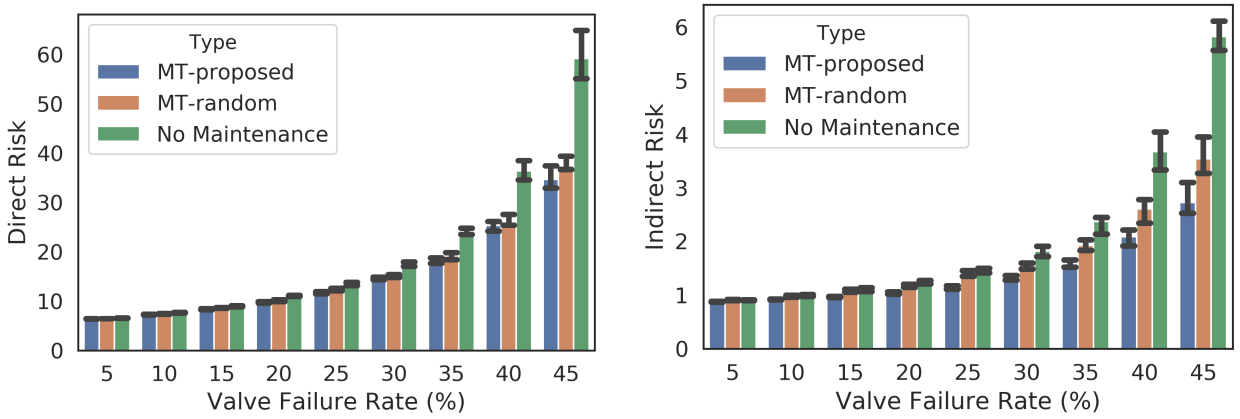
To measure the effectiveness of the proposed valve maintenance strategy, an alternative strategy is needed for comparison purposes. In this study, simulation result from a simple valve maintenance strategy is used as the benchmark for comparison. Instead of maintaining valves deliberately, the benchmark maintenance strategy randomly maintains the same amount of valves as the proposed strategy for the test WDN. MC simulation is then conducted on the valve maintained network to estimate the pipe isolation risks at various valve

Figure 8.10: Pipes with valves to maintain for the Alameda Island WDN

failure rate levels. Comparing the simulated pipe isolation risks between the proposed and benchmark strategies quantifies the effectiveness of the proposed one.

Figure 8.11 shows the simulated system pipe isolation risks of the Alameda Island WDN under the proposed valve maintenance strategy, benchmark valve maintenance strategy, and no valve maintenance. The statistical measures of the MC simulation results are summarized in Table 8.5.

One critical observation from Figure 8.11 is that both the proposed and benchmark valve maintenance strategy are effective in reducing the system pipe isolation risk compared to the unmaintained base case. Among the two, the proposed valve maintenance strategy is more effective in reducing the system's direct and indirect pipe isolation risk under all tested valve failure rates than the benchmark strategy. Specifically, the proposed valve maintenance strategy performs well on reducing the indirect pipe isolation risk for the network. For instance, at the 25% valve failure rate level, the proposed valve maintenance strategy reduces the indirect pipe isolation risk of the system from 1.46 GPM to 1.15 GPM, a 21% reduction from the unmaintained case. On the other hand, the benchmark valve maintenance strategy only reduces the indirect pipe isolation risk to 1.41 GPM, a mere 3.4% decrease from the base case. However, the risk deduction effect of the proposed strategy is less prominent for



(a) System direct pipe isolation risk

(b) System indirect pipe isolation risk

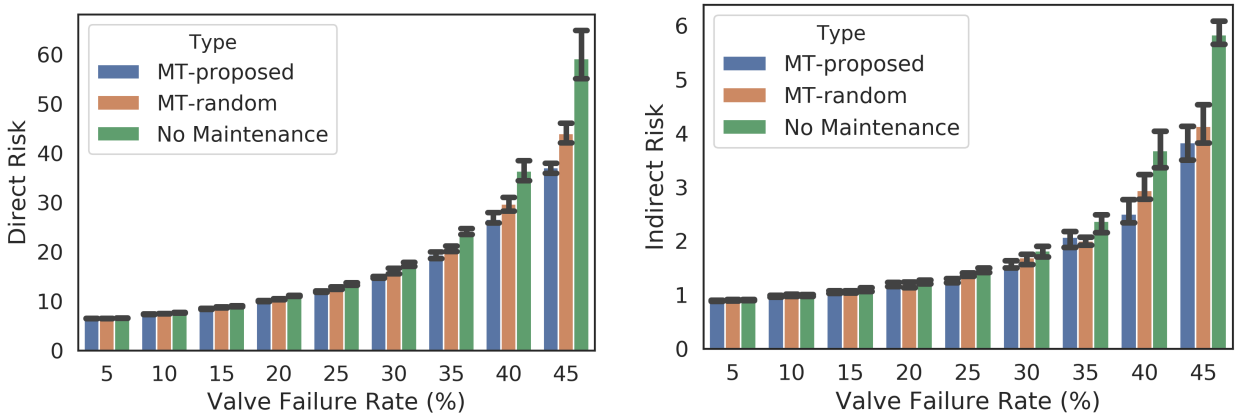
Figure 8.11: Comparison of the system pipe isolation risks under different maintenance strategies for the Alameda Island WDN

the direct pipe isolation risk of the system. The insignificant risk reduction is because most of the tested network has grid structures. In other words, the topology differences across the network are small. Therefore, maintaining valves on different parts of the system has a similar effect on controlling the size of isolation segments across the WDN, which is related to the direct pipe isolation risk of the system. Consequently, the benchmark strategy, which maintains valves randomly, has a similar effect on reducing the direct pipe isolation risk of the system compared to the proposed strategy.

As the valve failure rate increases, the differences in risk reduction for the two valve maintenance strategies become more significant. For example, at the 10% valve failure rate level, the proposed valve maintenance strategy reduces the indirect pipe isolation risk of the system from 0.99 GPM to 0.92 GPM, a 7.1% reduction. At the same valve failure rate level, the benchmark strategy reduces the indirect pipe isolation risk to 0.97 GPM, a 2.0 % reduction. The risk reduction difference is 5.1 % for the two methods at the 10% valve failure rate level. As the valve failure rate level increases to 45%, the difference between risk reduction of the two strategies becomes larger. Specifically, the proposed valve maintenance strategy reduces the indirect pipe isolation risk of the system from 59.19 GPM to 34.63 GPM, a 41.5% reduction. On the other hand, the benchmark valve maintenance strategy produces 40.07 GPM indirect pipe isolation risk (32.0 % reduction). The difference of risk reduction is 9.2 % at the 45% valve failure rate level, almost doubled the 5.1 % reduction rate at the 10% valve failure rate.

Table 8.5: Statistics of system pipe isolation risks under different valve maintenance strategies with various valve failure rates for the Alameda Island WDN

Valve failure rates	Valve maintenance strategy	Mean direct risk (GPM)	Direct risk STD	Mean indirect Risk (GPM)	Indirect risk STD
5%	Proposed	6.44	0.13	0.88	0.03
	Benchmark	6.47	0.16	0.91	0.05
	No Maintenance	6.56	0.16	0.91	0.05
15%	Proposed	8.34	0.43	0.97	0.06
	Benchmark	8.69	0.51	1.08	0.1
	No Maintenance	9.02	0.52	1.1	0.15
25%	Proposed	11.67	0.87	1.15	0.17
	Benchmark	12.39	1.11	1.41	0.21
	No Maintenance	13.5	1.25	1.46	0.27
35%	Proposed	18.21	2.47	1.57	0.34
	Benchmark	19.68	2.94	1.93	0.5
	No Maintenance	24.17	3.92	2.37	2.64
45%	Proposed	34.64	7.69	2.73	1.29
	Benchmark	40.07	11.08	3.55	3.01
	No Maintenance	59.19	21.1	5.83	4.01



(a) System direct pipe isolation risk

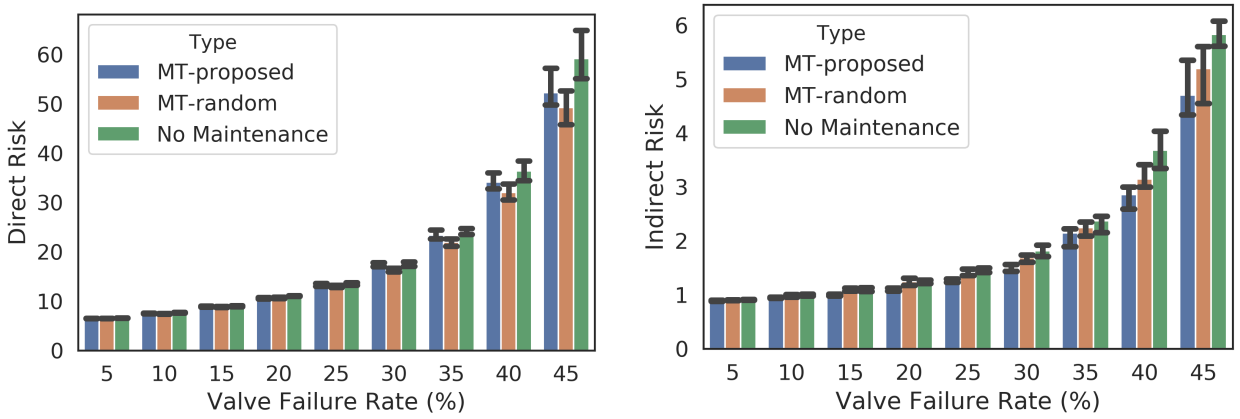
(b) System indirect pipe isolation risk

Figure 8.12: Comparison of the system pipe isolation risks under different maintenance strategies for the Alameda Island WDN (only valves with high edge betweeness centrality measure are maintained)

Contribution of the risk reduction The proposed valve maintenance strategy prioritizes the maintenance of two types of valves: valves with high edge betweenness centrality (in the Segment-Valve graph) and valves in the identified vulnerable region. This section explores the risk reduction contributions of the two valve types through controlled simulation experiments.

The first case considers the network with only large edge betweenness centrality valves maintained. Figure 8.12 shows the corresponding simulation result. Maintaining only valves with high edge betweenness centrality measures reduces direct and indirect system pipe isolation risk compared to both the non-maintenance and benchmark cases. Comparing Figure 8.12 with Figure 8.11, maintaining the valves with high centrality only has almost the same effect on reducing the direct pipe isolation risk as maintaining the full set of valves. However, the reduction degree on the indirect pipe isolation risk is significantly smaller than the full set case. For instance, at the valve failure rate of 45%, maintaining valves with high centrality measure only reduces system indirect pipe isolation risk to around 3.81 GPM. In contrast, the indirect risk of maintaining the whole set of identified valves is 2.73 GPM.

Figure 8.13 shows the simulation result when maintenance is only performed on the valves in the identified vulnerable region (92 valves). It is clear that maintaining such valves only does not help in reducing the direct pipe isolation risk for the system. For instance, the system's direct pipe isolation risk is even a bit higher than the benchmark strategy when the valve failure rate level is large ($> 30\%$). The performance drop is because maintaining valves in the vulnerable region restricts the formation of large-size isolation segments inside the selected regions only. Large size isolation segments can still form on other network parts, especially under high valve failure rates scenarios, resulting in large direct risk for the system. On the other hand, maintaining valves randomly effectively controls the isolation



(a) System direct pipe isolation risk

(b) System indirect pipe isolation risk

Figure 8.13: Comparison of the system pipe isolation risks under different maintenance strategies for the Alameda Island WDN (only pipes with valves in the vulnerable regions are maintained)

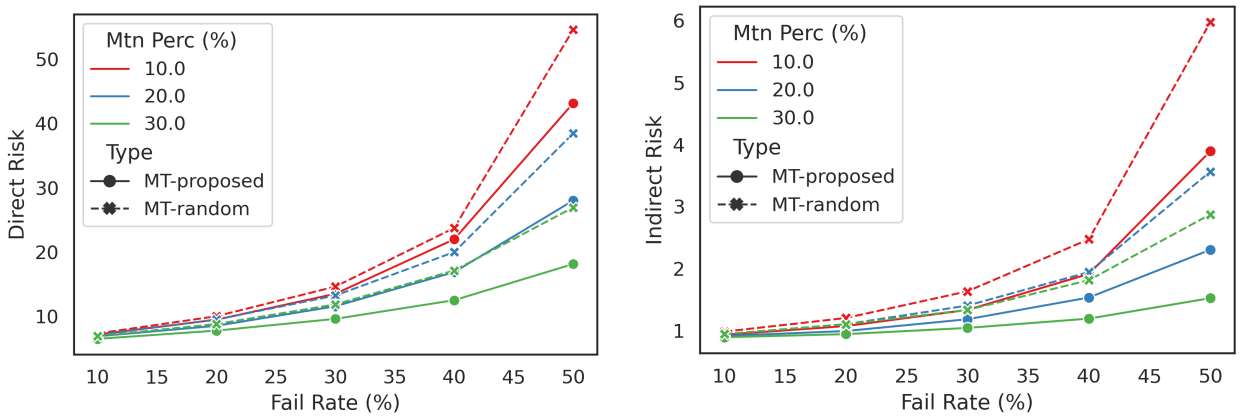
segment size globally, thus causing a reduction in the direct pipe isolation risk of the system. Nevertheless, maintaining valves in the vulnerable region reduces the indirect pipe isolation risk for the system, as shown in Figure 8.13b. For instance, the simulated indirect pipe isolation risk is about 15% less than the benchmark case at the 45% valve failure rate. In conclusion, maintaining valves in the vulnerable region reduces the indirect pipe isolation risk with the small cost of increasing the system's direct pipe isolation when the valve failure rate is high.

Sensitivity on number of selected valves One of the essential variables of the proposed maintenance strategy is the number of valves that need to be exercised. Water utilities often have limited resources to conduct valve maintenance programs. Therefore, understanding how the proposed method behaves regarding the number of exercised valves is critical for utilities to balance the valve maintenance cost and potential system risk. To answer such an question, sensitivity tests concerning the number of maintained valves (valve maintenance rate) are conducted in the following.

Performance of the proposed method under various valve maintenance rates are tested for the Alameda Island WDN. Given a specific valve maintenance rate, the total number of valves needing maintenance can be calculated. Half of the required valves are selected using the centrality criteria (valves with high betweenness centrality), and the other half is randomly chosen from the identified vulnerable regions. When there are not enough valves in the vulnerable regions (all valves in the vulnerable regions are selected, which happens when the valve maintenance rate is high), valves with the highest edge-betweenness centrality measure are chosen for the remainder. For instance, a maintenance rate of 20% requires 450

valves to be exercised in total, which requires 225 valves from the vulnerable region. However, the total number of isolation valves in the identified vulnerable regions is 92. Consequently, these 92 valves and the top 358 high edge-betweenness centrality valves are selected to be exercised for the 20% valve maintenance case.

Monte Carlo (MC) simulation is used to estimate the system pipe isolation risk (direct and indirect) with various valve failure rates for each maintenance scenario. It is assumed that the selected valves will not fail (0% valve failure rate). To benchmark the performance of the proposed valve maintenance method, a simple valve maintenance method, which is to randomly maintain the same number of valves as the proposed one, is also simulated for each tested case.



(a) System direct pipe isolation risk

(b) System indirect pipe isolation risk

Figure 8.14: Comparison of the system pipe isolation risks under maintenance strategies with different number of exercised valves for the Alameda Island WDN

Figure 8.14 shows the simulation results, and the detailed statistics is listed in Table 8.6. The first observation from Figure 8.14 is that increasing valve maintenance rates decrease both direct and indirect risk of the system at all valve failure levels. Furthermore, the system maintained using the proposed strategy has a smaller risk (both direct and indirect) than the benchmark strategy for all tested cases. The effectiveness of the proposed method becomes especially significant under high valve failure and maintenance rates. For example, at valve failure rate 10%, the indirect risk of maintaining 10% of the valves using the benchmark method is 0.98GPM. Under the same condition, the indirect risk reduces to 0.93GPM, a 5.1% deduction, when maintenance is performed using the proposed method. The direct risk decreases as well, with the reduction rate around 2.1% (7.32GPM to 7.16GPM). At the same valve failure level (10%), increasing the maintenance rate to 30% raises the risk reduction rate to 8.5% (0.94GPM to 0.86GPM) for indirect risk, 6.4% (6.92GPM to 6.48GPM) for the direct risk. The risk reduction effect becomes even greater under high valve failure rate cases. When 50% of valves fail in the network, maintaining 10% of valves using the

proposed strategy reduces indirect system risk by 34.8% (5.91GPM to 3.89GPM), and direct risk by 20.9% (54.59GPM to 43.15GPM). The indirect risk reduction ratio raises to 47.0% (2.87GPM to 1.52GPM) when 30% of valves are maintained in the system using the proposed method. The direct risk reduction rate is also significant, around 32.7% (26.92GPM to 18.12GPM). In conclusion, the proposed strategy's effectiveness (system risk reduction against the benchmark) improves as the number of maintained valves increases, and the risk reduction degree extends as the valve failure rate increases.

Table 8.6: Statistics of system pipe isolation risks (at the 30% valve failure rate) for maintenance strategies with various number of exercised valves in the Alameda Island WDN

Valve maintenance rates	Valve maintenance strategy	Mean direct risk (GPM)	Mean indirect Risk (GPM)
10%	Proposed	13.49	1.33
	Benchmark	14.61	1.63
20%	Proposed	11.55	1.17
	Benchmark	13.22	1.40
30%	Proposed	9.59	1.04
	Benchmark	11.84	1.33
40%	Proposed	8.60	0.94
	Benchmark	10.01	1.21
50%	Proposed	7.80	0.89
	Benchmark	9.28	1.12

Round hill network

EMBUD Round Hill WDN is a small branch-like WDN serving a community in the East San Francisco Bay Area (City of Alamo). It has 126 internal nodes, 126 pipes, and 48 valves. Pipes in the network have various lengths, materials, and demands. See Chap 6 Section 6.2 for more details of the network.

Centrality measures

The edge betweenness centrality measure of each valve is computed for the corresponding Segment-Valve graph of the system. Figure 8.15 shows the histogram of the result. Compared to the case of the Alameda Island WDN (Figure 8.6), the Round Hill WDN shows a

broader range of edge betweenness centrality measure across valves. Over 61% of valves have significant edge betweenness centrality (> 0.1), distributed relatively evenly in the range of 0.1 to 0.5. Failure of such valves may create large isolation segments in the system, increasing the system's pipe isolation risk.

The right-skewed distribution is due to the Round Hill network's sparsely connected feature and relatively few amount of valves. Since pipes are sparsely connected in the network, the corresponding Valve-Segment graph is sparsely connected as well. Consequently, the number of available paths between any two segments for the graph is limited, resulting in large values of edge betweenness centrality for most valves in the system. Lack of valves in the system aggravates the connectivity sparseness of the Valve-Segment graph (even fewer links in the graph), increasing the magnitude of edge betweenness for certain valves in the system.

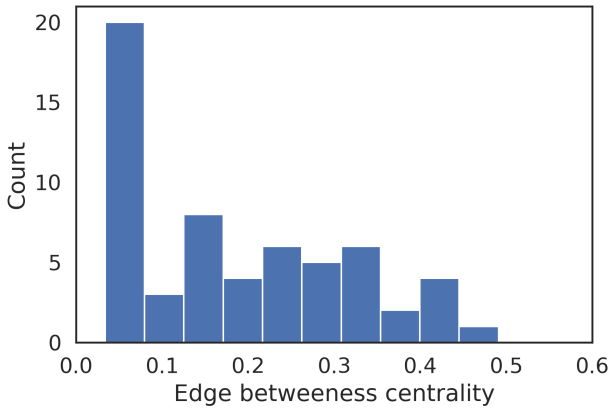


Figure 8.15: Distribution of edge betweeness centrality measures for valves in the Round Hill WDN

Figure 8.16 shows the map of edge betweenness measure for the Round Hill Network. Note that the centrality value of the valves are mapped to their corresponding pipes for visualization purposes. Similar to the Alameda Island WDN case, the valves located at the center of the network tend to have a higher value of edge betweenness centrality than the valves on the periphery pipes. Since merging isolation segments across the network requires the failure of central valves, they tend to have high centrality values. Note that the actual distribution of the centrality measure also depends on the spatial distribution of installed valves, which determines the connection pattern of the Segment-Valve graph. Therefore, not all central pipes have high centrality values in Figure 8.16 as some pipes have no valves installed.

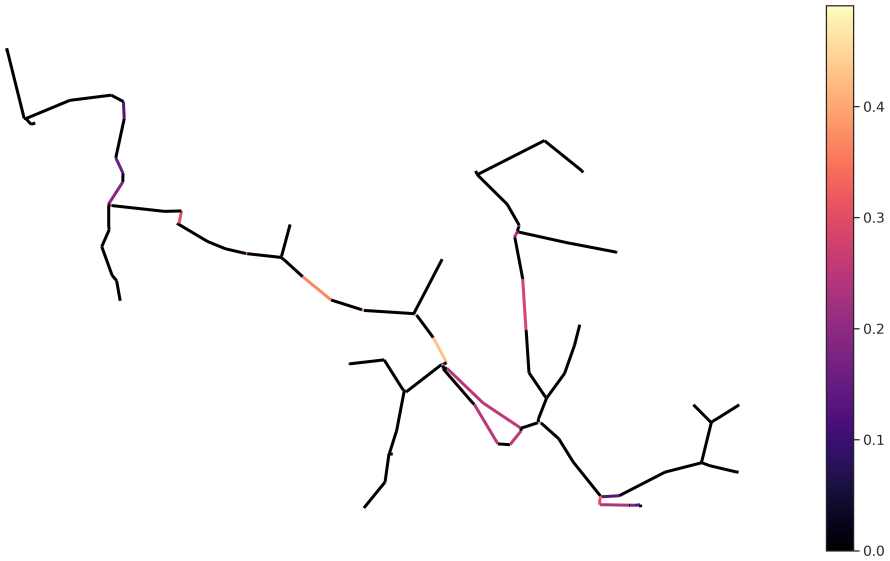


Figure 8.16: Edge betweeness centrality measure for each valve of the Round Hill WDN (mapped to the corresponding pipe for visualization purposes)

Segment-wise pipe isolation risk

Figure 8.17 shows the generated multi-scale pipe isolation risk map for the Round Hill WDN. The pipes with high demand risk (bigger than the system averaged per pipe risk) when isolating with small segment size (0-5 pipes) are marked in red (first-degree vulnerable pipes). The pipes with high risk only when the corresponding isolation segment size is large (5-30 pipes) are marked as yellow (second-degree vulnerable pipes). The pipes that have no risk regardless of tested segment sizes are colored in black. From Figure 8.17, the Round Hill WDN is dominated by the first-degree vulnerable pipes due to the sparse connectivity of pipes (See Chap 6 Section 6.2 for detailed discussion). Hence, the isolation of most pipes in the network will cause pipes in the other network parts to lose water supply. As a result, the whole Round Hill WDN can be treated as a large vulnerable region. In other words, all valves in the Round Hill WDN can be selected for maintenance based on the Segment-wise pipe isolation risk criteria.

Effectiveness of the maintenance strategy

Simulation is performed on the Round Hill WDN to test the effectiveness of the proposed valve maintenance strategy. As discussed in the previous section, the whole network can be viewed as a large vulnerable region due to the sparse connective pattern of pipes. Therefore, selecting valves in the vulnerable region to be maintained is equivalent to the benchmark valve maintenance strategy, which randomly maintain valves from the network. As a result, only valves with high edge betweenness values are maintained in this case. Following the

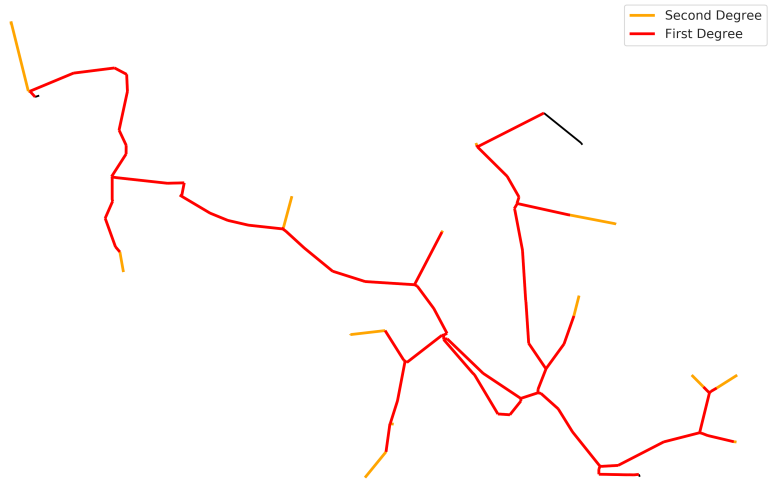


Figure 8.17: Pipe isolation risk map for the Round Hill WDN

case for the Alameda Island WDN study, valves with significant edge betweenness values (> 0.1) are selected to be maintained in priority, shown in Figure 8.18 (valves are mapped to the corresponding pipes for visualization purposes). In total, the total number of valves to be maintained is 30 in this case.

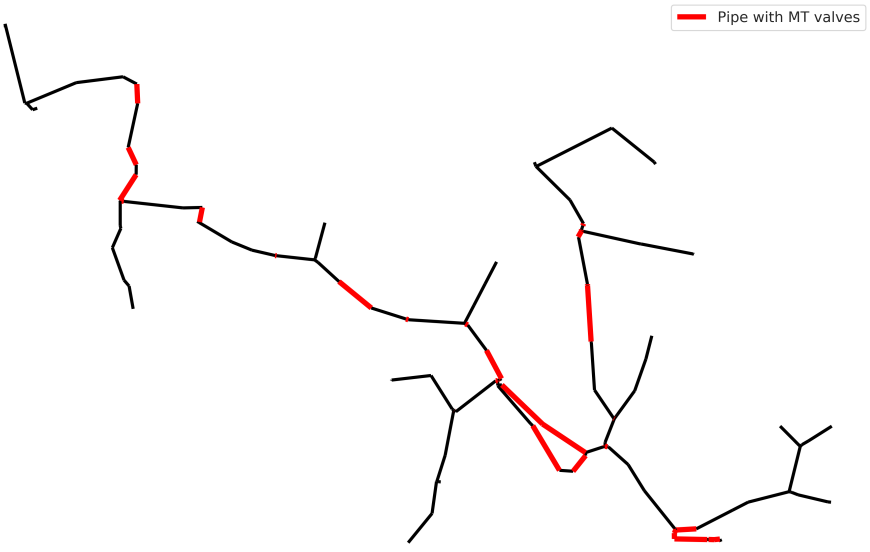
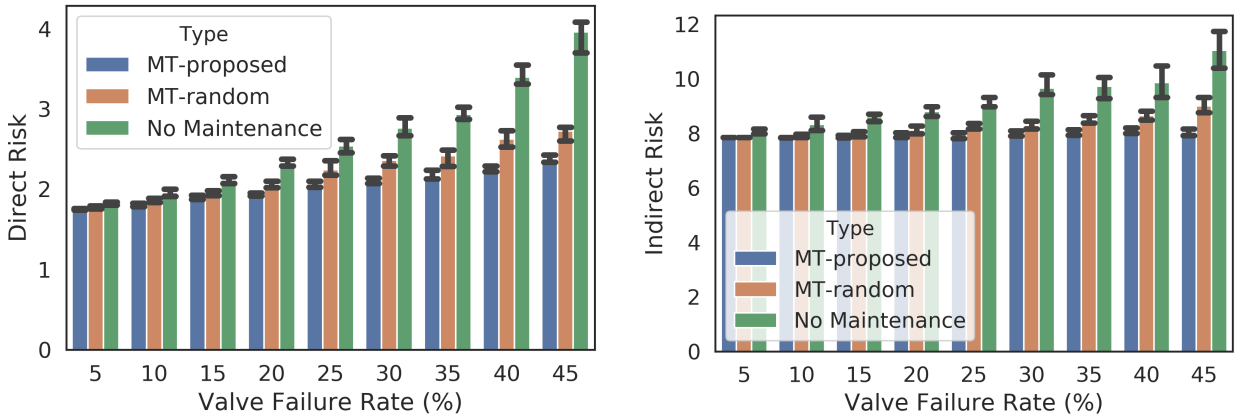


Figure 8.18: Pipes with valves to maintain for the Round Hill WDN



(a) System direct pipe isolation risk

(b) System indirect pipe isolation risk

Figure 8.19: Comparison of the system pipe isolation risks under different maintenance strategies for the Round Hill WDN

Same as the Alameda Island WDN study, an alternative valve maintenance strategy, which is to randomly select the same amount of valves to maintain for the tested network, is used as the benchmark strategy to demonstrate the effectiveness of the proposed maintenance strategy. The pipe isolation risks of the system are estimated using the Monte Carlo (MC) simulation under scenarios of no valve maintenance, the benchmark maintenance strategy, and the proposed maintenance strategy at various levels of valve failure rates. Figure 8.19 shows the simulation result, and the statistical measures are summarized in Table 8.7.

Both the proposed and benchmark valve maintenance strategies effectively reduce the system pipe isolation risk for the Round Hill WDN under all tested valve failure rate conditions. Among the two, the proposed valve maintenance strategy is more effective in restricting direct and indirect pipe isolation risk than the benchmark strategy. Specifically, both the risk level (mean) and uncertainty (standard deviation) are smaller for the network under the proposed valve maintenance strategy than the values under the benchmark case for all tested valve failure rates. For instance, at the 25% valve failure rate level, the proposed valve maintenance strategy reduces the direct pipe isolation risk of the system from 2.54 GPM to 2.06 GPM, a 19% reduction. On the other hand, the benchmark valve maintenance strategy only reduces the direct pipe isolation risk to 2.24 GPM, a 12% reduction. For indirect pipe isolation risk, the proposed method reduces the risk by 13%, where the benchmark strategy results in a 9 % risk reduction.

As the valve failure rate increases, the benefit of adopting the proposed valve maintenance strategy becomes more evident. As shown in Figure 8.19, both indirect pipe risk and direct pipe risk increase slower for the network under the proposed strategy compared to the case of the benchmark strategy. Note that the indirect risk of the network remains relatively stable across all tested valve failure rate scenarios, especially under the proposed valve maintenance

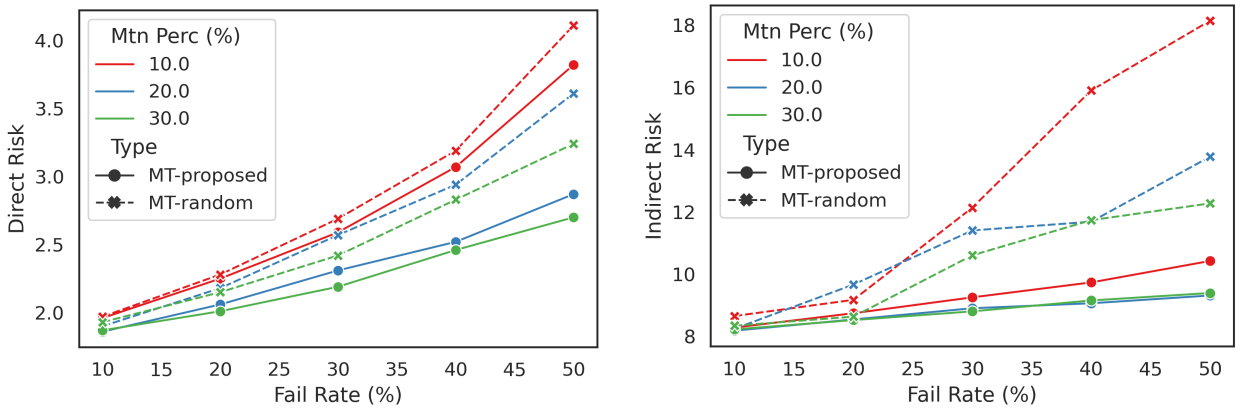
strategy. Specifically, the indirect risk of the system increases from 7.85 GPM to 8.10 GPM, a minor 3% increase, under the proposed valve maintenance strategy as valve failure rates rise from 5% to 45%. On the other hand, the indirect risk increases by 98% for the system with no valve maintenance and 15% for the network under the benchmark valve maintenance strategy. In conclusion, the proposed valve maintenance strategy, even with only part of it implemented (maintain valves with large edge betweenness measure), can effectively reduce the pipe isolation risk of a WDN.

Table 8.7: Statistics of system pipe isolation risks under different valve maintenance strategies with various valve failure rates

Valve failure rates	Valve maintenance strategy	Mean direct risk (GPM)	Direct risk STD	Mean indirect Risk (GPM)	Indirect risk STD
5%	Proposed	1.75	0.05	7.85	0.12
	Benchmark	1.77	0.07	7.85	0.21
	No Maintenance	1.82	0.11	8.03	0.42
15%	Proposed	1.90	0.1	7.86	0.31
	Benchmark	1.96	0.16	7.94	0.51
	No Maintenance	2.11	0.21	8.61	0.95
25%	Proposed	2.06	0.13	7.95	0.43
	Benchmark	2.24	0.25	8.24	0.58
	No Maintenance	2.54	0.31	9.1	15.92
35%	Proposed	2.17	0.15	8.03	0.52
	Benchmark	2.42	0.3	8.48	0.7
	No Maintenance	2.94	0.46	9.73	22.3
45%	Proposed	2.37	0.17	8.10	0.75
	Benchmark	2.72	0.35	9.01	1.16
	No Maintenance	3.96	0.84	11.05	15.91

Sensitivity on number of selected valves Unlike the densely connected Alameda Island WDN, the sparsely connected Round Hill WDN has no particular vulnerable regions, where an increasing number of isolation pipes (due to rising valve failure rate) leads to a dramatic indirect risk increase. More precisely, the whole Round Hill network can be considered a vulnerable region due to its sparse connective nature. Consequently, valve maintenance priority is ranked solely based on their centrality measures. Therefore, sensitivity tests are only performed regarding the number of exercised valves in this case.

Like the Alameda Island WDN case study, various valve maintenance rates are tested for the Round Hill network. For each valve maintenance case, Monte Carlo (MC) simulation is used to estimate the system pipe isolation risk (direct and indirect) under various valve failure rates. Again, it is assumed that the selected valves will not fail (0% valve failure rate). The effectiveness of the proposed valve maintenance method is illustrated through the comparison to the system performance under the benchmark maintenance strategy, which is to randomly maintain the same number of valves as the proposed one.



(a) System direct pipe isolation risk

(b) System indirect pipe isolation risk

Figure 8.20: Comparison of the system pipe isolation risks under maintenance strategies with different number of exercised valves for the Round Hill WDN

Figure 8.20 shows the simulation results. Increasing valve maintenance rates decreases system direct risk at all valve failure rate levels for both maintenance methods. Patterns of indirect risk are complicated, though. In most cases, increasing the valve maintenance rate for the benchmark strategy decreases the indirect risk, especially when the valve failure rate is high. On the other hand, the indirect risk stays almost intact for the proposed method as the valve maintenance rate passes a certain threshold. Table 8.8 shows that at valve failure rate 30%, increasing valve maintenance rate from 20% to 50% reduces the indirect risk by 4.5% (8.90GPM to 8.50GPM) for the proposed method. In comparison, the benchmark method reduces the risk by 23.2% (11.40GPM to 8.76GPM).

The stability of the indirect risk under the proposed valve maintenance strategy is due to the sparse connective nature of the Round Hill WDN. Sparse connectivity implies that the magnitude of total indirect risk depends on the status of certain critical pipes inside the system. For example, isolation of the root pipe for a branch starves the whole downstream network, dramatically increasing the total system indirect cost. Consequently, as long as valves around such critical pipes are maintained, the total indirect risk of the system will be effectively limited regardless of valve failure levels. Since the proposed method ranks the maintenance priority of valves based on edge betweenness centrality measures, valves on such critical pipes obtain high priority ratings (see Figure 8.17). In other words, by considering the topological nature of the network, the proposed method requires much fewer valves to maintain compared to the benchmark to achieve the same indirect risk level, especially under high valve failure rate cases.

Both Figure 8.20 and Table 8.8 shows that the proposed method performs better than the benchmark for both direct and indirect system risk at all tested valve maintenance and failure rates. Moreover, the risk reduction effect of the proposed method becomes more prominent as the valve failure rate increases. For instance, maintaining 10% of valves using the proposed method reduces the direct risk by 5.5% (2.18GPM to 2.06GPM), indirect risk by 8.5% (9.12GPM to 8.34GPM), compared to the benchmark at valve failure rate 10%. In extreme cases where half amount of valves are failed (50% valve failure rate), the proposed method performs significantly better than the benchmark at the same 10% valve maintenance rate. Specifically, the direct risk is reduced by 11.1% (4.21GPM to 3.74GPM), and indirect risk is reduced by 43.4% (17.88GPM to 10.12GPM). At 30% valve maintenance level, the direct risk is further reduced by 17.4% (3.16GPM to 2.61GPM). However, the indirect risk reduction becomes less prominent as the valve maintenance ratio increases. Specifically, the indirect risk reduction shrinks to 25.3% (11.98GPM to 8.95GPM) at maintenance rate 30%. The decrease of risk reduction is because the difference of the selected exercised valves between the proposed and benchmark methods decreases as the maintenance rate increases.

8.4 Summary

This chapter presents a novel valve ranking algorithm to assign maintenance priorities to isolation valves. The proposed method quantifies valve maintenance importance considering both the number and location uncertainties of malfunctioning valves for a WDN.

The valve ranking process is developed based on the findings from Chapter 5), which states that the system risk depends on the mean and variance of isolation segments across the network and the spatial distribution of the isolation cost. Therefore, maintenance of valves whose failures lead to the increase of such factors should be prioritized.

Network analysis is performed on the Segment-Valve graph to identify valves that are key to the mean and variance of network segments. Edge betweenness centrality measures are computed, which measures the frequency of the edge being included in the shortest path between a pair of nodes. Therefore, valves with high edge betweenness centrality are critical

Table 8.8: Statistics of system pipe isolation risks (at the 30% valve failure rate) for maintenance strategies with various number of exercised valves in the Round Hill WDN

Valve maintenance rates	Valve maintenance strategy	Mean direct risk (GPM)	Mean indirect Risk (GPM)
10%	Proposed	2.59	9.25
	Benchmark	2.69	12.13
20%	Proposed	2.31	8.90
	Benchmark	2.57	11.40
30%	Proposed	2.19	8.80
	Benchmark	2.42	10.60
40%	Proposed	2.11	8.67
	Benchmark	2.23	9.18
50%	Proposed	2.09	8.50
	Benchmark	2.08	8.76

to the segment merging process in the network, thus being labeled with high maintenance priority.

The inhomogeneous spatial distribution of isolation costs across the network is addressed by the multi-size pipe isolation risk map introduced in Chapter 5. An unsupervised machine learning technique, the Density-Based Spatial Clustering (DBSCAN) algorithm, is used to automatically identify vulnerable regions (subnetworks experience rapid pipe isolation risk increase as the segment sizes increase) from the risk map. Valves contained inside the detected vulnerable areas are selected as valves to maintain with priority.

The proposed valve maintenance strategy is tested on two real-life WDNs, the densely connected large scale Alameda Island WDN and the sparsely connected small scale Round Hill WDN. The effectiveness of the proposed method is measured by the risk reduction from the system configured with a benchmark maintenance strategy, which is to randomly choose the same amount of valves as the proposed one to exercise. Simulation shows that applying the proposed valve maintenance strategy effectively reduces both the direct and indirect risk for the two testing networks, especially under high valve failure rate cases.

Sensitivity analysis has been conducted on the two studies as well. For the Alameda Island WDN, it is found that prioritizing the maintenance of valves with a high edge betweenness centrality measure (in the Segment-Valve graph) reduces both the direct and indirect pipe isolation risks of a WDN. Maintaining valves in the identified vulnerable regions reduces

the indirect pipe isolation risk with the cost of slightly increased direct pipe isolation risk under high valve failure rate conditions. Maintaining valves from both criteria provides the system with the most pipe isolation risk reduction. However, if only a few valves can be maintained due to resource constraints, valves with high edge centrality measures should be prioritized to effectively reduce both the system's direct and indirect pipe isolation risks under all valve failure rate conditions. On the other hand, the entire Round Hill WDN is considered vulnerable due to its sparse connective nature. Therefore, the maintenance priority of valves in Round Hill WDN is solely ranked by the edge betweenness criteria.

Number of valves to maintain influences the performance of the proposed method for both networks. For the Alameda Island WDN, having more valves maintained leads to greater risk reduction against the benchmark (randomly maintain the same number of valves). The risk reduction effect increases as the valve failure rate rises. For example, maintaining a significant portion of valves (30%) using the proposed method makes the system more resilient (47% risk reduction rate) under extreme events (50% valve failure rate) than less severe conditions (6.4% risk reduction rate at 10% valve failure rate). On the other hand, valve maintenance benefits quickly saturated as the maintenance rate increased. Specifically, increasing the number of maintained valves does not significantly decrease the indirect risk after a portion of valves has been maintained (around 15%). Such a stable indirect risk response is due to the sparse connective nature of the Round Hill WDN. As long as the critical pipes (e.g., the root of a branch) are secured, the indirect risk of the network is bounded. Since the proposed method ranks the valves based on the network's topology, valves around the critical pipes obtain high maintenance priority. In other words, the proposed method prioritizes the maintenance of valves around pipes with high indirect risk. Since the indirect risk is much larger (around 10GPM) than the direct risk (around 3GPM), it suggests that the most cost-effective way of maintaining valves in the Round Hill WDN is to exercise a small portion (around 15%) of them following the maintenance priorities assigned by the proposed method.

Chapter 9

Summary and Conclusions

This chapter provides a summary of major research findings associated with the study. The sections that follow are organized to present an overview of research findings that correspond to the five objectives of the research, 1) aging water infrastructure and the consequences; 2) an efficient WDN hydraulic simulator and its application to EBMUD systems' post-earthquake response assessment; 3) risk formulation for WDNs with isolation valve condition uncertainty; 4) an optimal isolation valve placement algorithm that achieves minimum pipe isolation risk for a system given the number of valves to install; 5) a cost-effective isolation valve maintenance method that efficiently increases WDN resilience through deliberate valve maintenance/exercise. The final section discusses the limitation of the current research and future research directions.

9.1 Aging Water Infrastructure and the Consequences

WDNs in the United States are aging and underfunded. Across the country, many water infrastructure systems are serving beyond their intended design life, and the maintenance/replacement processes are generally insufficient. Aging and inadequate maintenance cause frequent WDN component failures. Water pipes, which are often buried underground, are the most frequent failed components for a WDN. According to the American Society of Civil Engineering (ASCE), there are around 250,000 to 300,000 pipe failures in the U.S. annually [45, 9]. Pipe may fail due to one of the following reasons: a) Material degradation over time and subsequent loss of structural integrity; b) Design defects or construction errors that weaken the system over time; c) Fatigue loading and subsequent localized structural damage; d) Adverse environments the system is exposed to; e) External environment impacts; f) Improper operating schemes; g) Lack of proper maintenance.

Pipe failures cause both direct economic and indirect social impacts for a community. The direct financial cost is related to WDN properties, including water loss, regional supply suspension, and repair costs. The indirect social impact of the events is often more severe

than the direct cost, which includes property damage, traffic disturbance, and public health issues. To assess the indirect social consequences of water main break events in an objective way, this study utilizes data mining techniques to collect news articles related to water main break events across the country. In total, 7674 news articles were collected from the internet over a year of the collection period. Due to the volume of the collected data, Natural Language Processing (NLP) techniques, the Latent Dirichlet Allocation (LDA) method in specific, is applied to extract the prevailing topics of the collected news automatically. Although the data analysis reveals no dominant topic for the collected data, some topics are more frequent than other topics. Specifically, health and traffic are the most common themes among all articles, validating the commonly held assumption that water quality deterioration and local traffic disturbance are the most severe social damages from water main break events.

9.2 Hydraulic Simulation for Water Distribution Networks

Natural hazard, such as earthquakes, often causes tremendous damage to aged WDNs. This study developed an efficient computer simulation package, HydrauSim, to understand the hydraulic behavior of WDNs before and after a disruptive hazard event such as an earthquake. HydrauSim has the following main functionalities: a) Generate water network models from scratch or from existing EPANET formatted water network model input (EPANET INP) files; b) Modify network structure by adding/removing components and changing component characteristics; c) Modify network operation by changing initial conditions and component settings; d) Simulate network hydraulics (steady-state) using either a demand-driven mode (DD) or a pressure-driven mode (PDD); e) provides tools to read, configure, and analyze the impact of isolation segments concerning different isolation valves configurations and valve failure scenarios. HydrauSim is faster than most existing simulators, making it ideal for hydraulic analysis on large-scale networks or optimization procedures that require many simulations.

Following the U.S. Geological Survey's (USGS) HayWired Report [65], earthquake impacts (M7.05 Hayward fault) on WDNs in the San Francisco Bay Area were studied. Compared to the HayWired case study, which used empirical models for post-earthquake WDN status estimation, this study adopted hydraulic simulation based methods for hazard impact assessment. Thanks to the computation efficiency of HydrauSim, post-earthquake hydraulic responses for a large-scale WDN under different ground-motion scenarios are estimated, revealing valuable patterns that may benefit hazard planning and mitigation.

This study adopted a probabilistic seismic hazard analysis (PSHA) as the general approach for simulating earthquake risk for the study area, the San Francisco Bay Area. A set of M7.05, Hayward-Rodgers Creek HN+HS earthquake rupture events from UCERF2 [144] were simulated with spatial correlation considerations on earthquake intensity mea-

sures (IMs).

The generated IMs were used as input to estimate the earthquake's impact on the WDN, which were modeled herein as follows: (1) Using the simulated earthquake ground-motion IMs (PGV), the probability of failure for each network component was estimated using pipeline fragility curves; (2) Based on the estimated failure probability, certain components were sampled as the failed component with corresponding degrees of damage; (3) A pressure-dependent hydraulic simulation was performed on the damaged network to estimate the water supply and degree of shortages anticipated for the case study. Monte Carlo (MC) simulation was used to deal with uncertainties about pipe failure type and severity.

The simulation was performed on East Bay Municipal Utility District's (EBMUD) main gravity feed zone (65,700 pipes). Due to the large size of the WDN, uncertainties of earthquake epicenter greatly impact the number of potential leaked/broken pipes in the system. Around 200–800 pipes were estimated to break during the simulated earthquake events. Moreover, the relationship between the pipe failure number and the resulting total water shortage ratio is found to be nonlinear. Specifically, the total water supply shortage ratio increases faster when the number of pipe failures increases (convex nonlinearity). On average, 25% of demand nodes might experience insufficient water pressure levels, which can rise to 78% for the worst-case scenario.

9.3 Pipe Isolation Risks Under Valve Condition Uncertainties

Pipe failure events, whether due to internal reasons (e.g., aging-related mechanical fatigue) or external events (e.g., earthquake-induced ground movements), require the closure of isolation valves for loss constrain and repair. However, many water utilities do not install enough valves, and the installed valves can also fail at the time of usage. Hence, it is necessary to consider the valve conditions of a pipeline system when analyzing the impacts of pipe failure events for a WDN. By combining and extending the previous works on pipe isolation risk analysis, this study illustrate how network properties of the Segment-Valve graph dynamically influence the risk level of the WDN.

The first step of WDN risk analysis with valve failure consideration is to convert the conventional Junction-Pipe representation to Segment-Valve representation for a WDN, where isolation segments and valves become the graph vertices and edges, respectively. Representing a WDN using segment and valves helps us to deduce Equation 5.2, which provides a complete description of pipe isolation risk for a WDN with valve condition considerations. Equation 5.2 states that the system risk depends on a) average size of segments; b) variance of segment sizes across the network; c) spatial distribution of the isolation cost, which is related to both topological and hydraulic properties of the network.

Simulation is conducted to evaluate the properties of the segment-valve graph of a WDN under various valve failure scenarios. Both the mean and standard deviation of segment sizes

show the five phases phenomenon regarding valve failure rates changes. Initially (low valve failure rates), the number of multi-pipe segments (a segment that contains multiple pipes, which is related to both mean and standard deviation of segment sizes) stays small (Phase A) due to the valve redundancy. As the valve failure rate increases, it starts to increase as redundancy is depleted (Phase B) and reaches the equilibrium state (Phase C). The number decreases dramatically in Phase D when large-size segments dominate the network and stabilize in Phase E, where only a few segments are left.

A method of generating networks with different topologies is developed to validate the generality of the observed five phases phenomenon. The proposed network generation process can be considered a spatial version of the configuration model in network theory, which utilizes graph operations (e.g., edge contraction, etc.) to generate arbitrary network topology from a lattice network. The five phases phenomenon is observed on all the generated networks. In addition to synthetic networks, the proposed risk estimation method is also tested on real-life networks. The proposed five phases phenomenon is observed on both the EBMUD Alameda Island WDN, a densely connected large scale network, and EBMUD Round Hill WDN, a sparsely connected small scale network.

9.4 Valve Placement Strategy for Minimal System Pipe Isolation Risk

Since isolation valves play an essential role in the resilience of a WDN, where to place them is a critical issue many water utilities face. This study proposes an optimal valve placement algorithm to address the challenge. The proposed method led to a system with minimal pipe isolation risk for all valve placement configurations with the same number of installed valves, which can be proved using induction. Because of its simplicity and optimality, the proposed method also provides insights into the benefits and limitations of some common rule-of-thumb valve placement strategies.

The proposed process starts from a system with the ideal or safest valve placement configuration (e.g., two valves installed at the two ends of every pipe) and gradually removes the least impact valve from the system until the desired number of valves is reached. In addition to finding the best locations to install valves for a new system, the proposed method can be modified to find the best place to add new valves for an established system. Specifically, instead of considering all the installed valves at each removal step, the already deployed valves are excluded in the least-impact valve searching process. The other aspects (e.g., valve removal impact assessment) are the same for the two algorithms. In summary, the modified algorithm finds the minimal risk valve placement configuration under the constrain that the already deployed valves must be included.

The proposed method theoretically proves the effectiveness of some commonly used rule-of-thumb valve placement strategies. For example, the N valving strategy is the initial state (ideal case) of the proposed valve placement process. The N-1 placement strategy is the

minimal valves solution to the proposed valve placement process under the constraint that system risk remains at the lower bound (system risk at the initial condition). Assuming all pipes are equal-likely to fail, and the failure consequences are the same, the proposed process generates the "one valve per pipe" and the "one valve per pipe and one valve per node" valve placement configuration.

The naive implementation of the proposed method is not computationally feasible. An efficient implementation is developed to address this problem, which eliminates the redundant computation through graph operation and memorization. Using the improved implementation, the optimal valve placement configuration is generated for EBMUD Alameda Island WDN and EBMUD Round Hill WDN. Comparing to the already deployed valve placement configuration in the two systems, the proposed algorithm dramatically reduces the overall system risk, proving the effectiveness of the proposed method on real-life networks.

9.5 Valve Improvement Strategy for System Pipe Isolation Risk Reduction

Valves fail in real-life situations. Although it is best to add new valves into the system using the proposed valve placement strategy, utilities often find it hard to do so due to budget constraints. Instead, maintaining/exercising the existing valves is more practical than installing new ones. However, the large volume of the isolation valves inside a system (over tens of thousands for an extensive network) makes it unrealistic to maintain/exercise every valve. This study presents a method that ranks the maintenance priority of valves based on their failure consequence. Valves with significant failure consequences gain high maintenance priority, which helps utilities use the limited resources effectively to reduce overall system risk regarding pipe failure events.

The developed valve ranking algorithm is based on the insights provided by the system risk equation (Equation 5.2), which states that the system risk depends on the mean and variance of isolation segments across the network, and the spatial distribution of the isolation cost. Network analysis is used to find the valves that are key to the network's segment size properties (mean and variance). Specifically, valves with large edge betweenness centrality values (in the Segment-Valve graph) are ranked higher (to be maintained with high priority) than low-value ones. Since edge betweenness centrality measures the frequency of the edge being included in the shortest path between a pair of nodes, valves with high edge betweenness centrality are critical to the segments merging process in the network. Therefore, decreasing the failure probability of such valves through maintenance may prevent the spread of successive merges of segments in the network, leading to a network with smaller isolation segments than the non-maintained ones.

To account for the inhomogeneous spatial distribution of isolation costs across the network, a multi-size pipe isolation risk map is introduced to locate "vulnerable regions" of the system that experience rapid pipe isolation risk increase as valve failure rate increases. Specif-

ically, the multi-size pipe isolation risk map contains the isolation risk of every single pipe in the network with a set of predefined isolation segment sizes throughout simulation runs. An unsupervised machine learning technique, the Density-Based Spatial Clustering (DBSCAN) algorithm, is used to automatically identify vulnerable regions for a WDN. Valves that are contained inside the detected vulnerable regions are selected as valves to maintain in priority.

The proposed valve ranking algorithm is tested on real-life WDNs. For the densely-connected large-scale EBMUD Alameda Island WDN, the proposed valve maintenance strategy performs significantly better (especially for large valve failure rate cases) on reducing system risk compared to the baseline model (randomly maintaining the same number of valves). The proposed method also beats the benchmark for the sparsely connected small-scale Round Hill WDN, proving its effectiveness across networks with various sizes and topologies.

9.6 Limitations and Future Studies

This study has several limitations that can be improved in future works.

Regarding the social impact estimation of water main break events, this study only uses topic mining techniques (LDA) to list the prevailing topics of the collected news data. However, the collected data contains much richer information than topics alone. For example, some news may contain the time and location of the incident, which can be used to create a national water-main break hazard map. Time series analysis can also be applied to the collected data to reveal any hidden temporal structure (e.g., does water main break happen more frequently during winter than summer?). Lastly, advanced models can be developed to estimate the economic consequence from the collected data, which provides objective quantitative insights about the social impacts of such events.

The main drawback of the developed WDN hydraulic simulator is that it does not include extended-time hydraulic simulation nor water-quality simulation, which are available for other software such as EPANET and WNTR. Future studies can expand the code to include these functionalities.

The post-earthquake WDN performance assessment case study in the San Francisco Bay Area also has several limitations. First, the study area of WDNs herein does not cover the whole San Francisco Bay Area. Instead, the study focused on the Central Pressure Zone, which covers major populated cities in the East San Francisco Bay Area and provides water services for almost half of the EBMUD's customers. Network data outside the EBMUD management area has not been included. Future studies may include contacting other water utility companies to acquire network data to model the WDN systems for the whole Bay Area. Another limitation of this study is that only buried pipes were modeled as breakable. The reason for such simplification is that the fragility of sophisticated WDN components such as pumps and regulators depends not only on ground motions but also on the functionalities of other infrastructure systems, such as the power grid. In future work, the performance of

other critical infrastructure after earthquakes should be considered to correctly model the fragility of other critical WDN components other than pipelines.

In the current study, water sources such as reservoirs, tanks, and water treatment plants were modeled as robust (i.e., their failure wasn't considered). This is usually a valid assumption considering the strict building codes and regular maintenance for these structures. That said, severe earthquakes may damage such structures in rare cases, causing a complete water supply loss for the region. Due to the severity of the consequences should failure occur, the fragility of water-source structures may require further study. This study did not model WDN damages due to permanent ground deformation because of the region's lack of data for soil profiles. Hence, earthquake damage levels for the WDN may be underestimated. Future studies should explore the impacts of other earthquake properties on the WDNs, such as permanent ground deformation and aftershocks. Lastly, this study focused on quantifying the impact of an earthquake on the WDN. Modeling the restoration process after the earthquake for WDNs will be considered in future studies.

One can apply the proposed system risk analysis framework (with valve condition consideration) to improve the hazard response assessment of a WDN (e.g., earthquakes). In the current research, earthquake-induced pipe breaks are treated at pipe level (broken pipes can be isolated alone). However, as illustrated in later chapters, isolation valves need to be closed before pipe repair or replacement. Treating pipe-break events in the isolation segment level may provide a more realistic view of the earthquake impact for a WDN.

The current research uses supply deficiency (due to pressure drop) as the measure to quantify the risk of a WDN with uncertain valve conditions. More advanced measures can be developed in the future to account for other essential aspects of water supply for a WDN (e.g., hospitals are more critical than single-family houses). In addition to the risk measure, the biggest assumption made in this research regarding valve failure assessment is that every valve has the same probability of failure. Such an assumption does not hold in real-life situations. Future research may focus on creating a fragility equation for isolation valves based on valve type, age, surrounding soil type, etc. Nevertheless, the risk analysis framework proposed in this work can be applied to any distribution of valve failure rates.

The computation efficiency of the proposed valve placement and valve ranking algorithm can be further improved in future studies. Through analyzing the behavior of the proposed valve placement strategy, it may be possible to come up with heuristics to guide water utilities without running the actual simulation. Regarding valve maintenance strategies, network properties other than the edge betweenness centrality can be tested to improve the effectiveness of the ranking procedure. Moreover, the application limitations of the proposed two algorithms should be explored using large amount of real-life WDNs with various topological and demand distribution features.

Bibliography

[1] Brad T Aagaard et al. "Ground-motion modeling of Hayward fault scenario earthquakes, part I: Construction of the suite of scenarios". In: *Bulletin of the Seismological Society of America* 100.6 (2010), pp. 2927–2944.

[2] Brad T Aagaard et al. "Ground-motion modeling of Hayward fault scenario earthquakes, Part II: Simulation of long-period and broadband ground motions". In: *Bulletin of the Seismological Society of America* 100.6 (2010), pp. 2945–2977.

[3] BT Aagaard et al. *HayWired scenario mainshock ground motions, chap. C of Detweiler, ST, and Wein, AM, eds., The HayWired earthquake scenario—Earthquake hazards: US Geological Survey Scientific Investigations Report 2017–5013–A–H, 126 p.* 2017.

[4] Rania Albalawi, Tet Hin Yeap, and Morad Benyoucef. "Using topic modeling methods for short-text data: A comparative analysis". In: *Frontiers in Artificial Intelligence* 3 (2020), p. 42.

[5] Alliance, AL. *Seismic fragility formulations for water systems, Part I-Guideline*. 2001.

[6] José M Alonso et al. "Parallel computing in water network analysis and leakage minimization". In: *Journal of Water Resources Planning and Management* 126.4 (2000), pp. 251–260.

[7] F Alvarruiz, Fernando Martınez-Alzamora, and AM Vidal. "Improving the efficiency of the loop method for the simulation of water distribution systems". In: *Journal of Water Resources Planning and Management* 141.10 (2015), p. 04015019.

[8] Fernando Alvarruiz, Fernando Martınez Alzamora, and Antonio M Vidal. "Improving the performance of water distribution systems' simulation on multicore systems". In: *The Journal of Supercomputing* 73.1 (2017), pp. 44–56.

[9] ASCE. "2021 infrastructure report card". In: *American Society of Civil Engineers*. 2021.

[10] Tirusew Asefa et al. "Performance evaluation of a water resources system under varying climatic conditions: Reliability, Resilience, Vulnerability and beyond". In: *Journal of Hydrology* 508 (2014), pp. 53–65.

[11] Gregory M Baird. "Managing Assets: When Going With the Flow Doesn't Save Money". In: *Journal-American Water Works Association* 103.9 (2011), pp. 18–23.

[12] Jack W Baker. "An introduction to probabilistic seismic hazard analysis (PSHA)". In: *White paper, version* 1 (2008), p. 72.

[13] Donald B Ballantyne and Craig Taylor. "Earthquake loss estimation modeling of the Seattle water system using a deterministic approach". In: *Lifeline Earthquake Engineering*. ASCE. 1990, pp. 747–760.

[14] Marc Barthélemy. "Spatial networks". In: *Physics Reports* 499.1-3 (2011), pp. 1–101.

[15] Luigi Berardi et al. "Assessing mechanical vulnerability in water distribution networks under multiple failures". In: *Water Resources Research* 50.3 (2014), pp. 2586–2599.

[16] Michael W Berry and Malu Castellanos. "Survey of text mining". In: *Computing Reviews* 45.9 (2004), p. 548.

[17] David M Blei, Andrew Y Ng, and Michael I Jordan. "Latent dirichlet allocation". In: *the Journal of machine Learning research* 3 (2003), pp. 993–1022.

[18] Phillip Bonacich. "Power and centrality: A family of measures". In: *American journal of sociology* 92.5 (1987), pp. 1170–1182.

[19] David M Boore and Gail M Atkinson. "Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s". In: *Earthquake Spectra* 24.1 (2008), pp. 99–138.

[20] Stephen P Borgatti. "Centrality and network flow". In: *Social networks* 27.1 (2005), pp. 55–71.

[21] Stephen P Borgatti et al. "Network analysis in the social sciences". In: *science* 323.5916 (2009), pp. 892–895.

[22] F Bouchart and I Goulter. "Reliability improvements in design of water distribution networks recognizing valve location". In: *Water Resources Research* 27.12 (1991), pp. 3029–3040.

[23] Gregor Burger et al. "Quest for a new solver for EPANET 2". In: *Journal of Water Resources Planning and Management* 142.3 (2016), p. 04015065.

[24] Lee Cesario. *Modeling, analysis, and design of water distribution systems*. Amer Water Works Assn, 1995.

[25] Jeongwook Choi and Doosun Kang. "Improved Hydraulic Simulation of Valve Layout Effects on Post-Earthquake Restoration of a Water Distribution Network". In: *Sustainability* 12.8 (2020), p. 3492.

[26] Symeon E Christodoulou and Michalis Fragiadakis. "Vulnerability assessment of water distribution networks considering performance data". In: *Journal of Infrastructure Systems* 21.2 (2015), p. 04014040.

[27] R M Chung et al. "January 17, 1995 Hyogoken-Nanbu (Kobe) earthquake: Performance of structures, lifelines, and fire protection systems (NIST SP 901)". In: (1996).

[28] Carlo Ciaponi, Enrico Murari, and Sara Todeschini. "Modularity-based procedure for partitioning water distribution systems into independent districts". In: *Water resources management* 30.6 (2016), pp. 2021–2036.

[29] CNT. *The Case for Fixing the Leaks*. 2013.

[30] J Coate. "Water main shutdown application prevents costly mistakes". In: *ArcUser, Jan-Mar, ESRI, Redlands, CA* (2003), pp. 20–21.

[31] E Creaco, M Cunha, and M Franchini. "Using heuristic techniques to account for engineering aspects in modularity-based water distribution network partitioning algorithm". In: *Journal of Water Resources Planning and Management* 145.12 (2019), p. 04019062.

[32] Enrico Creaco, Marco Franchini, and Stefano Alvisi. "Optimal placement of isolation valves in water distribution systems based on valve cost and weighted average demand shortfall". In: *Water resources management* 24.15 (2010), pp. 4317–4338.

[33] John E Cromwell. *Costs of infrastructure failure*. American Water Works Association, 2002.

[34] Peter A Crous, Jakobus E van Zyl, and Yuko Roodt. "The potential of graphical processing units to solve hydraulic network equations". In: *Journal of Hydroinformatics* 14.3 (2012), pp. 603–612.

[35] Daniel A Crowl and Joseph F Louvar. *Chemical process safety: fundamentals with applications*. Pearson Education, 2001.

[36] Arne De Coninck et al. "Needles: Toward Large-Scale Genomic Prediction with Marker-by-Environment Interaction". In: 203.1 (2016), pp. 543–555. ISSN: 0016-6731. DOI: 10.1534/genetics.115.179887. eprint: http://www.genetics.org/content/203/1/543.full.pdf. URL: http://dx.doi.org/10.1534/genetics.115.179887.

[37] Arun K Deb et al. *Criteria for valve location and system reliability*. American Water Works Association, 2006.

[38] Kegong Diao and Wolfgang Rauch. "Controllability analysis as a pre-selection method for sensor placement in water distribution systems". In: *Water research* 47.16 (2013), pp. 6097–6108.

[39] Kegong Diao et al. "Speedup of water distribution simulation by domain decomposition". In: *Environmental modelling & software* 52 (2014), pp. 253–263.

[40] EPA. *Drinking Water Infrastructure Needs Survey and Assessment, Sixth Report to Congress*. 2018.

[41] Robert Epp and Alvin G Fowler. "Efficient code for steady-state flows in networks". In: *Journal of the hydraulics division* 96.1 (1970), pp. 43–56.

[42] FEMA, ed. *Hazus - MH 2.1 Hurricane Model Technical Manual*. en. Federal Emergency Management Agency, 2018. URL: http://www.fema.gov/plan/prevent/hazus.

[43] FEMA. *HAZUS-MH 2.1: Earthquake Model Technical Manual*. 2012.

[44] Edward H Field, Thomas H Jordan, and C Allin Cornell. "OpenSHA: A developing community-modeling environment for seismic hazard analysis". In: *Seismological Research Letters* 74.4 (2003), pp. 406–419.

[45] Steven Folkman. "Water main break rates in the USA and Canada: A comprehensive study". In: (2018).

[46] Steven Folkman. "Water Main Break Rates in the USA and Canada: A Comprehensive Study, April 2012". In: (2012).

[47] Peter E Gaewski and Frank J Blaha. "Analysis of total cost of large diameter pipe failures". In: *Proc. AWWA Research Symposium Distribution Systems: The Next Frontier, Reno, Nev*. 2007.

[48] Paolo Gardoni et al. "The interdependent networked community resilience modeling environment (IN-CORE)". In: *16th European Conference on Earthquake Engineering*. 2018.

[49] Johannes Gessler and Thomas M Walski. *Water Distribution System Optimization.* Tech. rep. ARMY ENGINEER WATERWAYS EXPERIMENT STATION VICKSBURG MS ENVIRONMENTAL LAB, 1985.

[50] O Giustolisi and L Ridolfi. "A novel infrastructure modularity index for the segmentation of water distribution networks". In: *Water Resources Research* 50.10 (2014), pp. 7648–7661.

[51] Orazio Giustolisi. "Water distribution network reliability assessment and isolation valve system". In: *Journal of Water Resources Planning and Management* 146.1 (2020), p. 04019064.

[52] Orazio Giustolisi and Dragan Savic. "Identification of segments and optimal isolation valve system design in water distribution networks". In: *Urban Water Journal* 7.1 (2010), pp. 1–15.

[53] Neil S Grigg. "Water main breaks: Risk assessment and investment strategies". In: *Journal of Pipeline Systems Engineering and Practice* 4.4 (2013), p. 04013001.

[54] Bluefield Research Group. *Underground Infrastructure: U.S. Water Wastewater Pipe Network Forecast, 2019-2028*. 2019.

[55] Gaël Guennebaud, Benoît Jacob, et al. *Eigen v3*. http://eigen.tuxfamily.org. 2010.

[56] Michele Guidolin, Zoran Kapelan, and Dragan Savić. "Using high performance techniques to accelerate demand-driven hydraulic solvers". In: *Journal of Hydroinformatics* 15.1 (2013), pp. 38–54.

[57] R Gupta et al. "Upgrading reliability of water distribution networks recognizing valve locations". In: *Procedia Engineering* 89 (2014), pp. 370–377.

[58] Aric Hagberg, Pieter Swart, and Daniel S Chult. *Exploring network structure, dynamics, and function using NetworkX*. Tech. rep. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[59] John Frederick Hall, William T Holmes, and Peter Somers. *Northridge earthquake of January 17, 1994: reconnaissance report*. Vol. 11. Earthquake Engineering Research Institute, 1995.

[60] John A Hartigan and Manchek A Wong. "Algorithm AS 136: A k-means clustering algorithm". In: *Journal of the royal statistical society. series c (applied statistics)* 28.1 (1979), pp. 100–108.

[61] Erika Hernandez and LE Ormsbee. "Application of segment based robustness assessment for water distribution networks". In: *WDSA/CCWI Joint Conference Proceedings*. Vol. 1. 2018.

[62] Erika Hernandez Hernandez and Lindell Ormsbee. "A Heuristic for Strategic Valve Placement". In: *Journal of Water Resources Planning and Management* 148.2 (2022), p. 04021103.

[63] JW Hoff. "Maintenance requirements of valves in distribution systems". In: *Annual conference proceedings AWWA, June*. Vol. 23727. 1996.

[64] Kenneth W Hudnut et al. *The HayWired earthquake scenario—We can outsmart disaster*. Tech. rep. US Geological Survey, 2018.

[65] Kenneth W Hudnut et al. *The HayWired earthquake scenario—We can outsmart disaster*. Tech. rep. US Geological Survey, 2018.

[66] Tamás Huzsvár, Richárd Wéber, and Csaba János Hős. "Analysis of the segment graph of water distribution networks". In: *Periodica Polytechnica Mechanical Engineering* 63.4 (2019), pp. 295–300.

[67] Hwee Hwang and Kevin Lansey. "Isolation valve impact on failure severity and risk analysis". In: *Journal of Water Resources Planning and Management* 147.3 (2021), p. 04020110.

[68] *Introducing JSON*. JSON ORG. URL: `%5Curl%7Bhttps://www.json.org/json-en.html%7D`.

[69] Nirmal Jayaram and Jack W Baker. "Correlation model for spatially distributed ground-motion intensities". In: *Earthquake Engineering & Structural Dynamics* 38.15 (2009), pp. 1687–1708.

[70] Sang-Soo Jeon and Thomas D O'Rourke. "Northridge earthquake effects on pipelines and residential buildings". In: *Bulletin of the Seismological Society of America* 95.1 (2005), pp. 294–318.

[71] Charles Job and Robert Barles. "USEPA's 2015 Drinking Water Infrastructure Needs Survey". In: *Journal-American Water Works Association* 107.5 (2015), pp. 84–86.

[72] Hwandon Jun and GV Loganathan. "Valve-controlled segments in water distribution systems". In: *Journal of Water Resources Planning and Management* 133.2 (2007), pp. 145–155.

[73] Hwandon Jun and GV Loganathan. "Valve-controlled segments in water distribution systems". In: *Journal of Water Resources Planning and Management* 133.2 (2007), pp. 145–155.

[74] Sasan Karamizadeh et al. "An overview of principal component analysis". In: *Journal of Signal and Information Processing* 4.3B (2013), p. 173.

[75] R Karjalainen and K-K Hogstrom. "Emergency valve reliability: The intelligent solution". In: *VALVE MAGAZINE* 12.4 (2000), pp. 50–55.

[76] Michael Kerrisk. *CRON Command*. Linux man pages. URL: `%5Curl%7Bhttps://man7.org/linux/man-pages/man8/cron.8.html%7D`.

[77] Mahmoud Khater and Mircea Grigoriu. "Graphical demonstration of serviceability analysis". In: *Structural Safety and Reliability*. ASCE. 1989, pp. 525–532.

[78] Katherine Klise et al. *Water Network Tool for Resilience (WNTR) User Manual*. Tech. rep. Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2020.

[79] Thomas K Landauer, Peter W Foltz, and Darrell Laham. "An introduction to latent semantic analysis". In: *Discourse processes* 25.2-3 (1998), pp. 259–284.

[80] Chan Wook Lee, Hyuk Jae Kwon, Do Guen Yoo, et al. "Seismic Reliability Assessment of Water Supply Systems Considering Critical Paths". In: *Applied Sciences* 10.22 (2020), p. 8056.

[81] Daniel D Lee and H Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization". In: *Nature* 401.6755 (1999), pp. 788–791.

[82] Renee Lee and Anne S Kiremidjian. "Uncertainty and correlation for loss assessment of spatially distributed systems". In: *Earthquake Spectra* 23.4 (2007), pp. 753–770.

[83] Pei-Hao Li and Jehng-Jung Kao. "Segment-based vulnerability analysis system for a water distribution network". In: *Civil Engineering and Environmental Systems* 25.1 (2008), pp. 41–58.

[84] Haixing Liu et al. "Failure impact analysis of isolation valves in a water distribution network". In: *Journal of Water Resources Planning and Management* 143.7 (2017), p. 04017019.

[85] Jun Liu and Rui Han. "Spectral clustering and multicriteria decision for design of district metered areas". In: *Journal of Water Resources Planning and Management* 144.5 (2018), p. 04018013.

[86] Edward Loper and Steven Bird. "Nltk: The natural language toolkit". In: *arXiv preprint cs/0205028* (2002).

[87] Le Val Lund. *TCLEE Pipeline Failure Database*. Technical Council on Lifeline Earthquake Engineering, American Society of Civil Engineers, 1992.

[88] I Markov, Mircea Grigoriu, and TD O'Rourke. "An evaluation of seismic serviceability of water supply networks with application to the San Francisco auxiliary water supply system". In: (1994).

[89] DW Martin and G Peters. "The application of Newton's method to network analysis by digital computer". In: *Journal of the institute of Water Engineers* 17.2 (1963), p. 115.

[90] JC Matthews, RJ Stowe, and ST Ariaratnam. "Main breaks: State of the science". In: *Water Research Foundation Project* 4374 (2015).

[91] John C Matthews. *Integrated, multi-attribute decision support system for the evaluation of underground utility construction methods*. Louisiana Tech University, 2010.

[92] Larry W Mays. *Water distribution system handbook*. McGraw-Hill Education, 2000.

[93] Fanlin Meng, Chris Sweetapple, Guangtao Fu, et al. "Placement of isolation valves for resilience management of water distribution systems". In: *WDSA/CCWI Joint Conference Proceedings*. Vol. 1. 2018.

[94] Microsoft. *Microsoft Bing News Search API*. Microsoft. URL: `%5Curl%7Bhttps://www.microsoft.com/en-us/bing/apis/bing-news-search-api%7D`.

[95] Austin Mohr. "Quantum computing in complexity theory and theory of computation". In: *Carbondale, IL* 1 (2014).

[96] Attilio Fiorini Morosini, Olga Caruso, and Paolo Veltri. "Management of water distribution systems in PDA condition with isolation valves". In: *Multidisciplinary Digital Publishing Institute Proceedings* 2.11 (2018), p. 672.

[97] National Infrastructure Advisory Council (NIAC). *Critical infrastructure resilience: Final report and recommendations*. National Infrastructure Advisory Council, 2009.

[98] Christopher M Navarro et al. "MAEviz: Exploring earthquake risk reduction strategies". In: *2008 IEEE Fourth International Conference on eScience*. IEEE. 2008, pp. 457–457.

[99] Water Infrastructure Network. "Water infrastructure now: Recommendations for clean and safe water in the 21st century". In: *Water Infrastructure Network (WIN), Washington, DC* (2001).

[100] Mark Newman. *Networks*. Oxford university press, 2018.

[101] Michael O'Rourke and Gustavo Ayala. "Pipeline damage due to wave propagation". In: *Journal of Geotechnical Engineering* 119.9 (1993), pp. 1490–1498.

[102] Congressional Budget Office. *Federal Support for Financing State and Local Transportation and Water Infrastructure*. 2018.

[103] Sukru Ozger and Larry W Mays. "Optimal location of isolation valves in water distribution systems: a reliability/optimization approach". In: *Water resource systems management tools* 7 (2004).

[104] Lina Sela Perelman et al. "Automated sub-zoning of water distribution systems". In: *Environmental Modelling & Software* 65 (2015), pp. 1–14.

[105] Kalyan R Piratla et al. "Empirical analysis of water-main failure consequences". In: *Procedia engineering* 118 (2015), pp. 727–734.

[106] KA Porter. "A new model of water-network resilience, with application to the HayWired scenario". In: *The HayWired earthquake scenario–Engineering Implications: chap. N* (2018).

[107] Juan Ramos et al. "Using tf-idf to determine word relevance in document queries". In: *Proceedings of the first instructional conference on machine learning*. Vol. 242. 1. Citeseer. 2003, pp. 29–48.

[108] HS Rao, DW Bree, and R Benzvi. "Extended period simulation of water distribution networks". In: *AVAILABLE FROM THE NATIONAL TECHNICAL INFORMATION SERVICE AS PB-230 148* 5.00 (1974).

[109] Radim Řehřek, Petr Sojka, et al. "Gensim—statistical semantics in python". In: *Retrieved from genism. org* (2011).

[110] Natalia Romero et al. "Seismic hazards and water supply performance". In: *Journal of Earthquake Engineering* 14.7 (2010), pp. 1022–1043.

[111] Lewis A Rossman et al. "Epanet 2 users manual, us environmental protection agency". In: *Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, OH* 45268 (2000).

[112] Lewis A Rossman et al. "EPANET 2: users manual". In: (2000).

[113] Lewis A Rossman et al. "EPANET users manual". In: (1994).

[114] GF Santonastaso, A Di Nardo, and E Creaco. "Dual topology for partitioning of water distribution networks considering actual valve locations". In: *Urban Water Journal* 16.7 (2019), pp. 469–479.

[115] S. Sarikelle and A.L. Cesario. "Developments in Water Distribution System and Analysis and Graphics Programs". In: *Proceedings of the AWWA Distribution System Symposium, Seattle, WA.* (1985).

[116] C Scawthorn. "Lifeline interaction and post-earthquake functionality". In: *Proc. 5th US-Japan Workshop on Earthquake Disaster Prevention for Lifeline Systems*. 1992, pp. 441–450.

[117] Charles Scawthorn, TD O'rourke, and FT Blackburn. "The 1906 San Francisco earthquake and fire—Enduring lessons for fire protection and water supply". In: *Earthquake Spectra* 22.2_suppl (2006), pp. 135–158.

[118] Erich Schubert et al. "DBSCAN revisited, revisited: why and how you should (still) use DBSCAN". In: *ACM Transactions on Database Systems (TODS)* 42.3 (2017), pp. 1–21.

[119] Uri Y Shamir and Charles DD Howard. "Water distribution systems analysis". In: *Journal of the Hydraulics Division* 94.1 (1968), pp. 219–234.

[120] S Shea. "Valve maintenance in Boston". In: *Proc., Resources Engineering and Operations for the New Decades, Annual Conf., AWWA, Denver*. 1991, pp. 561–566.

[121] P Shi, TD O'rourke, and Y Wang. "Simulation of earthquake water supply performance". In: *Proceedings of 8th US National Conference on Earthquake Engineering*. 2006.

[122] Peixin Shi. *Seismic response modeling of water supply systems*. Cornell University, 2006.

[123] Peixin Shi and Thomas D O'Rourke. *Seismic response modeling of water supply systems*. 16. Multidisciplinary Center for Earthquake Engineering Research Buffalo, NY, 2008.

[124] Sangmin Shin et al. "A systematic review of quantitative resilience measures for water infrastructure systems". In: *Water* 10.2 (2018), p. 164.

[125] Masanobu Shinozuka, A Rose, and RT Eguchi. "Engineering and socioeconomic impacts of earthquakes". In: *Buffalo: MCEER* (1998).

[126] Qing Shuang et al. "System reliability evaluation in water distribution networks with the impact of valves experiencing cascading failures". In: *Water* 9.6 (2017), p. 413.

[127] R Sitzenfrei et al. "Cascade vulnerability for risk analysis of water infrastructure". In: *Water Science and Technology* 64.9 (2011), pp. 1885–1891.

[128] Taronne Tabucchi, Rachel Davidson, and Susan Brink. "Simulation of post-earthquake water supply system restoration". In: *Civil Engineering and Environmental Systems* 27.4 (2010), pp. 263–279.

[129] Ah-Hwee Tan et al. "Text mining: The state of the art and the challenges". In: *Proceedings of the pakdd 1999 workshop on knowledge disocovery from advanced databases*. Vol. 8. Citeseer. 1999, pp. 65–70.

[130] Ezio Todini and S Pilati. "A gradient algorithm for the analysis of pipe networks". In: *Computer applications in water supply: vol. 1—systems analysis and simulation*. 1988, pp. 1–20.

[131] Ezio Todini and Lewis A Rossman. "Unified framework for deriving simultaneous equation algorithms for water distribution networks". In: *Journal of Hydraulic Engineering* 139.5 (2013), pp. 511–526.

[132] Jacob M Torres et al. "Exploring topological effects on water distribution system performance using graph theory and statistical models". In: *Journal of Water Resources Planning and Management* 143.1 (2017), p. 04016068.

[133] USEPA. *Factoids: Drinking Water and Ground Water Statistics for 2002*. 2003.

[134] Janet M Wagner, Uri Shamir, and David H Marks. "Water distribution reliability: simulation methods". In: *Journal of water resources planning and management* 114.3 (1988), pp. 276–294.

[135] David J Wald et al. "Relationships between peak ground acceleration, peak ground velocity, and modified Mercalli intensity in California". In: *Earthquake spectra* 15.3 (1999), pp. 557–564.

[136] Thomas M Walski. "Issues in providing reliability in water distribution systems". In: *ASCE EWRI Conference, Roanoke, VA*. 2002.

[137] Thomas M Walski. "Practical aspects of providing reliability in water distribution systems". In: *Reliability Engineering & System Safety* 42.1 (1993), pp. 13–19.

[138] Thomas M Walski. "Water distribution valve topology for reliability analysis". In: *Reliability engineering & system safety* 42.1 (1993), pp. 21–27.

[139] AR Whittaker and AW Arscott. "Development of an intelligent valve key to improve identification and control of distribution system valves". In: *1997 Annual Conf. Proc., AWWA*. 1997, pp. 229–237.

[140] C. Wilson and M. Garbark. "Manager to Manager, Message to CEOs: Focus on the Bottom Line." In: *American Water Works Association. Journal* 94.11 (2002), p. 26.

[141] Don J Wood. "User's manual-computer analysis of flow in pipe networks including extended period simulations". In: *Department of Civil Engineering, University of Kentucky, Lexington, KY* (1980).

[142] Don J Wood and Carl OA Charles. "Hydraulic network analysis using linear theory". In: *Journal of the Hydraulics division* 98.7 (1972), pp. 1157–1170.

[143] Don J Wood and AG Rayes. "Reliability of algorithms for pipe network analysis". In: *Journal of the Hydraulics Division* 107.10 (1981), pp. 1145–1161.

[144] Working Group on California Earthquake Probabilities(WGCEP). "The uniform California earthquake rupture forecast, version 2 (UCERF 2)". In: *US Geol. Surv. Open-File Rept. 2007-1437 and California Geol. Surv. Special Rept. 203, Version 1.1.* (2008).

[145] Zhijiang Yang et al. "Optimal Placement of New Isolation Valves in a Water Distribution Network Considering Existing Valves". In: *Journal of Water Resources Planning and Management* 148.6 (2022), p. 04022032.

[146] Sreeganesh R Yerri et al. "Empirical analysis of large diameter water main break consequences". In: *Resources, Conservation and Recycling* 123 (2017), pp. 242–248.

[147] Do Guen Yoo et al. "Seismic hazard assessment model for urban water supply networks". In: *Journal of Water Resources Planning and Management* 142.2 (2016), p. 04015055.

Appendix A

Supplement Materials for Water Main Break Events Text Mining

A.1 Data collection

This study uses the Bing News Search API to collect water main break news data. The Bing News Search API allows users to query online news data using search keywords. The query results include an authoritative image of the news article, related news, categories, provider info, article URL, and date added. The first step to using the Bing News Search services is registering on Microsoft Azure and requesting an API key, which is free for limited usages. With the API key, the connection can be made with the Bing search engine using Python programs (other programming languages work as well). The API requires query sentences to search news online. Since the study aims to obtain news related to water main break events, the query "water main break" is used. This study conducts daily searches on US news sources through a news collection program. The program is automatically executed daily using the CRON command on a Linux machine [76].

After the query, the Bing News Search Service returns a list of queried news objects. Each news object contains summarized information about the news article. Only the article headline, description, provider, and URL are extracted and stored among all returned information. If the discarded information such as image and video is required during the analysis phase, they can be recovered from the original news post using the saved URL. All the preprocessed, cleaned daily news are stored into a JSON file [68] with collection date and time. The status of the data collection program is monitored and notified to the administrator through text messages. The data collection process is summarized in Figure A.1.

A.2 Data preprocessing

Python library Natural Language Toolkit (NLTK) [86] is used as the preprocessing tool for the collected news data. NLTK provides stop-word removal, stemming, lemmatizing, tok-

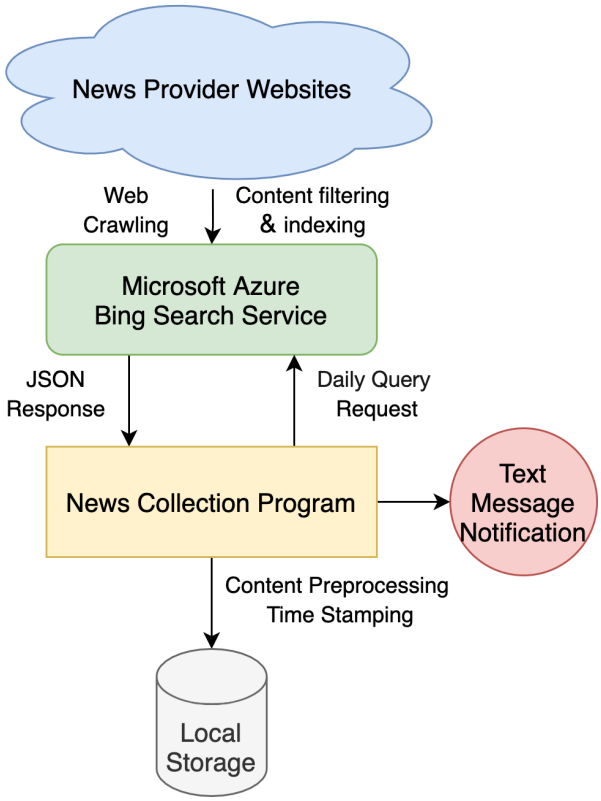


Figure A.1: Flowchart of the online water main break news collection program

enization, identifying n-gram procedures, and other data cleanings like lowercase transformation and punctuation removal. The preprocessing steps used in the study are summarized as the following:

- Identifying n-gram: find bigram (phrases containing two words) and trigram (phrases containing three words) words from the data and consider them as one word.
- Tokenizing: divide the text document into tokens like phrases, words. In other words, tokenization converts the text piece into a sequence of tokens.
- Stop-word elimination: remove the most common words from the English language. These are words such as prepositions, numbers (such as “the”, “a”, “an”, “in”) that do not contain meaningful information of the semantics of the text.
- Stemming: convert words into their root. For example word ”cared” is stemmed into ”care”. This study uses Snowball Stemmer algorithms for stemming.

- Lemmatizing: determine the lemma of a word based on its intended meaning. System's accuracy can be improved through the lemmatizing process. WordNetLemmatizer algorithm from NLTK library is used in this study for lemmatizing.

The preprocessed dataset is then converted to a digital dictionary (text to integer) using the Python gensim library [109]. Words with extremely high frequency such as "water", "main" are excluded from the dictionary to remove redundant information. Terms (words) in the preprocessed text data are weighted using the frequency-inverse document frequency (TF-IDF) method [107]. The TF-IDF weighting process uses frequency information to induce discriminative information for each term. A high weight TF-IDF term is essential inside a given document (has a high term frequency) but with a low document frequency in the whole collection of documents. Specifically, the TF-IDF weighting process requires the calculation of the frequency of word t in document d $TF(t, d)$ (term frequency TF), and the document frequency $DF(t)$, which is the number of documents that contain at least one t term. Opposite the document frequency, the inverse document frequency (IDF) is minor for frequently occurring terms in the entire document sets and large for terms that occur rarely. The TF-IDF score is calculated by the product of the TF term and IDF term:

$$\text{tfidf}(t, d, D) = tf(t, d)idf(t, D) \tag{A.1}$$

where t is the tested term, d is the document the term in in, and D is the total documents sets. $tf(t, d)$ and $idf(t, D)$ is defined as the following:

$$tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \tag{A.2}$$

where $f_{t\prime,d}$ is the count of term t in document d.

$$idf(t, D) = \log \frac{N}{d \in D : t \in d} \tag{A.3}$$

where N is the total number of documents in the corpus, and $\frac{N}{d \in D : t \in d}$ means the number of documents that contain term t.

A.3 Choosing TM models

Many state-of-art mature TM methods are available to extract topics for the collected dataset. Common used state-of-art TM methods are listed as below:

- PCA: principal component analysis (PCA) is a state-of-art dimension reduction algorithm that can decrease high dimensional feature vectors to a lower dimension while retaining most information. The effectiveness of PCA is widely proved in many studies across different areas. The PCA TM method found a d-dimensional subspace (topics subspace) that could capture as much of the dataset's variation as possible when applying to topic mining [74].

- LSA: Latent Semantic Analysis (LSA) is a commonly used algorithm in topic modeling. The LSA method is based on the assumption that terms with similar meanings are closer in terms of their contextual usage, assuming that words near their meaning show in the related parts of texts. LSA considers both the similarity terms of text and associated terms to generate more insights into the topic [79].

- LDA: Latent Dirichlet Allocation (LDA) is a probabilistic model for topic mining. It assumes that each document is made up of various topics, where each topic is a probability distribution over words. As a Bayesian statistical model, LDA uses Dirichlet priors for the document-topic and word-topic distributions [17]. LDA is considered the most popular TM algorithm because meaningful topics can be learned from the dataset without any prior knowledge [4].

- NMF: Non-Negative Matrix Factorization (NMF) is an unsupervised (no prior knowledge needed) matrix factorization (linear algebraic) method that performs both dimension reduction and clustering simultaneously. NMF decomposes high-dimensional vectors into a lower-dimensional, non-negative representation. The original dataset (articles by words matrix) is decomposed into a topic matrix (words of topics) and the coefficients (weights) for the extracted topics. Both topics and their weights to a given article are automatically found through the matrix factorization procedure on the corpus [81].

This study adopted LDA as the TM method for the collected news data. The primary reason for such a choice is that LDA can extract topics without prior knowledge of the underlying semantics in the data. In contrast, LSA needs to be trained with data that has topic labels, which are unavailable in this study. The underlying assumptions of the LDA algorithm also match the aim of the study. Specifically, LDA explicitly assumes that each article is composed of different topics (social impacts of water main break events). Each topic is represented by a collection of terms (words in the news). NMF, as a linear algebraic method, does not match the study goal compared to LDA (unsuitable assumptions). The final advantage of the LDA algorithm is that terms (words) are weighted and ranked for each topic based on their importance on the topic formulation, making the model outcomes easy to understand and interpret. On the other hand, since PCA only considers variances during the topic mining process, the produced outcomes are often hard to interpret [4].

Appendix B

Supplement Materials for HydrauSim

B.1 Isolation valves input format

To efficiently add leakage information to a WDN, the EPANet .inp file has been extended to allow users to input leakage information conveniently. HydrauSim can automatically parse the leakage information from the extended .inp file and run pressure-driven hydraulic simulation to quantify leaks. Each leak is represented by two properties. NID is the name of the node that may experience leakage. Diameter is the diameter of the leak, which is used for leak quantification with Equation (3.8). All leakage information can be grouped to the [LEAKS] section and appended to the .inp file. Figure B.1 shows an example of the [LEAKS] section.

Note that the leaks section should be appended after the [END] section in the original .inp file to allow the file to be compatible with EPANet or WNTR.

B.2 Isolation Valves Input Format

To efficiently add isolation valves to a WDN, the EPANet .inp file is extended to allow users to conveniently input the configuratoin of the isolation valves. HydrauSim can automat-

```
[LEAKS]
;NID                Diameter
node1                  3 ;
node2                  5 ;
node3                  8 ;
```

Figure B.1: Format of leaks information for the extended EPANet .inp file.

```
[ISOVALVES]
;ID                 Node            Pipe
V1                  2               P1  ;
V2                  4               P3  ;
V3                  2               P4  ;
V4                  3               P4  ;
V5                  6               P5  ;
```

Figure B.2: Format of isolation valves for the extended EPANet .inp file.

ically parse the isolation valves information from the extended .inp file and generate the corresponding segment-valves graph. Each isolation valve is represented by three properties. ID is the name of the isolation valves. Node is the closest node id to the isolation valve, and pipe is the pipe id that the isolation valve is placed on. All the isolation valve information can be grouped to the [ISOVALVES] section and appended to the .inp file. Figure B.2 shows an example of the [ISOVALVES] section.

Note that the isolation valve configuration section should be appended after the [end] section in the original .inp file to allow the file to be compatible with EPANet or WNTR.

B.3 Synthetic WDNs Generation Procedure

Synthetic WDNs are useful to profile the computation speed of hydraulic simulation programs as network attributes such as network size can be controlled by users. A simple yet effective algorithm is used in this study to generate synthetic water distribution networks that resemble real WDNs:

1. User inputs the desired size of the network, N. Let the square root integer of N to be n.
2. Generate a n by n grid graph.
3. Randomly remove edges with probability p.
4. Randomly assign node properties to nodes from the node property pool (uniform sampling)
5. Randomly assign pipe properties to remaining pipes from the pipe property pool (uniform sampling)
6. Connect reservoirs to the nodes at the four corners at higher elevations
7. Finish generating the synthetic network.

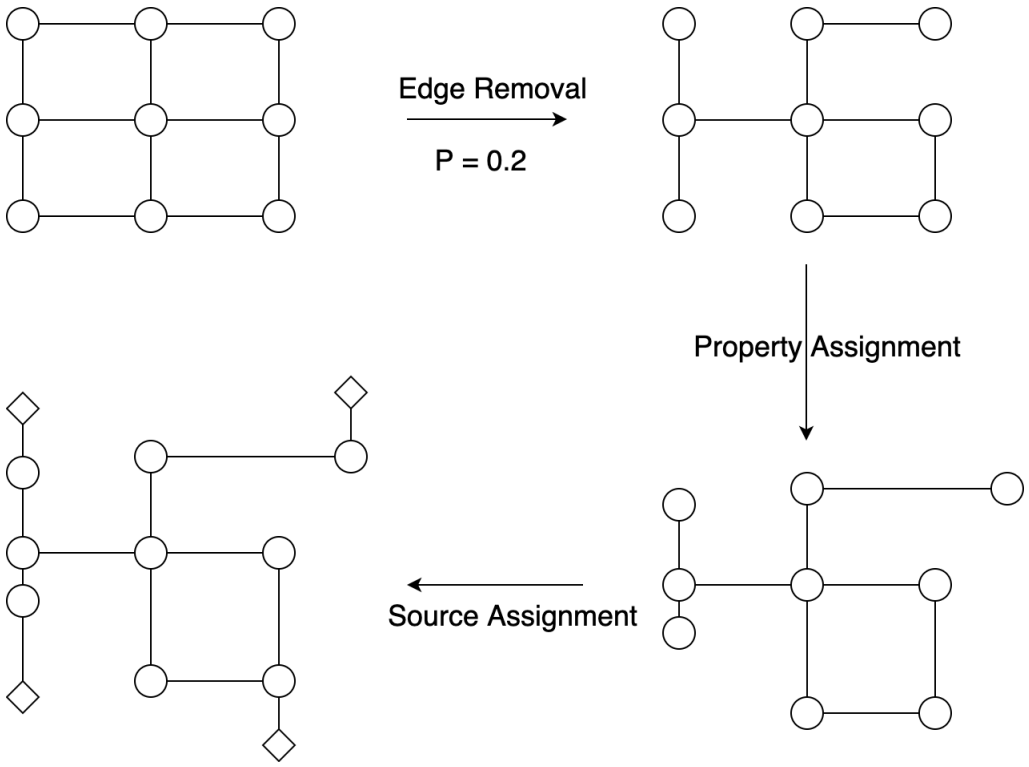


Figure B.3: The procedure to generate a synthetic network with 9 nodes.

From graph theory, the number of edges e for a n by n grid graph is given as:

$$e = 2n^2 - 2n \tag{B.1}$$

Hence, the ratio of number of edge e to the number of nodes n^2, r can be calculated as:

$$r = 2 - \frac{2}{n} \tag{B.2}$$

Since r for real-world networks is around 1.5, setting the probability of removal p to 0.2 leads to $r \approx 1.6$ for large networks. Generation of a synthetic WDN with 9 nodes is illustrated by Figure B.3.

The node property pool used in this study is shown by Table B.1. The pipe property pool used in this study is shown in Table B.2.

Note that synthetic networks are used for program profiling purposes; therefore, additional components such as pumps and valves are not included in the network generation procedure. All hydraulic simulations running on all of the generated synthetic networks with procedures and configurations described above converge.

Table B.1: Node properties used for creating synthetic WDNs.

	Min	Max
Elevation(ft)	400	900
Demand(GPM)	0	5
Head for sources(ft)	500	1000

Table B.2: Pipe properties used for creating synthetic WDNs.

	Min	Max
Length(ft)	200	800
Diameter(inch)	3	20
Roughness(unitless)	120	155