Water Distribution Network Sectorisation

Saeed Hajebi
School of Computer Science and Statistics
Trinity College Dublin

A thesis submitted to the University of Dublin, Trinity College
in fulfillment of the requirements for the degree of
Doctor of Philosophy

Trinity College Dublin August 2015

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

Saeed Hajebi

August 2015

Permission to Lend and/or Copy

Saeed Hajebi

August 2015

To

my loving parents,

my lovely daughter,

and

my love.

Acknowledgements

I would like to express my deepest appreciation to all those who provided me the possibility to complete this thesis.

First, I would like to thank my supervisors, Prof. Siobhán Clarke and Stephen Barrett, for their guidance and scientific advice, patience and kind words.

Thanks also to Lero, the Irish Software Research Centre, for giving me the opportunity to conduct research in Ireland and sponsoring me. More over, I thank IBM Software Ireland Lab. for their industry collaboration.

Many thanks goes to Ehsan Roshani, Mohsen Zandi, Alireza Moghaddam, and Mohammad Mortazavi Naeini for their vivid discussions about the hydraulics aspects of the work. Additionally, I would like to thank Moez Draief, Oliver Mason, Douglas Leith, Josep Domingo-Ferrer, and Mohammad Karzand for their knowledgeable discussions and comments about the graph theoretic aspects of the work.

I would also like to thank the members of my research group, DSG, for their support. In particular, Nicolás Cardozo, Suzy Temate, Hui Song, Vivek Nallur, Amit Raj, Eamonn O'Toole, and Nanxi Chen.

Thanks to all of my friends in Ireland who always reminded me that there is a life besides the PhD: Ali Akbari, Elham Esmaeeli, Aref Zarin, Tahmineh Esfandani, Rouzbeh Razavi, Elham Rasouli, Mojtaba Akbari, Ausieh Afsar, Leila Afsar, Mohammad Mortazavi, Ali Shamaei, Mina Shamaei, Javad Hosseinian, Afsaneh Kaghazgaran, Hadi Zamini, Hoda Akbari, Hamid Akbari, Hessan Feghhi, Saman Feghhi, Pooyan Jamshidi, and Hourieh Hamrah. In addition, two special friends who never stopped their constant support during the past four years: Mahmood Hassani and Jalil Alemasoom.

Thanks to my wonderful parents, Ali and Kobra, who gave me all their love and support in every moment of my life. Additionally, my beloved brothers and sisters: Alireza, Fahimeh, Najmeh, Meysam, and Narges.

Thanks to my lovely daughter, Saba, from her I learned a lot about curiosity, creativity, perseverance, being ambitious, and never giving up.

Last, but most importantly, Fatemeh, my love! Thank you for your constant support, love, and patience. It would have been impossible without you.

Saeed Hajebi
August 2015

Abstract

A water distribution network (WDN) supplies drinking water to homes and businesses, and links water sources to consumers. Such networks are typically complex and dynamic, consisting of thousands of nodes interconnected by thousands of elements. Control and management of a WDN as a whole is challenging as there is no granular information about different parts of the network to support location of issues. Partitioning a WDN into smaller sectors is a strategy to manage its complexity, as advised by the International Water Association. Advantages include (a) enhanced leakage and burst detection and management, (b) improved contamination spread control (associated with *water security*), (c) a capacity to provide different pressure levels, (d) enhanced demand management, and (e) better rehabilitation and work planning.

Water security (*i.e.*, safeguarding the network from intentional or unintentional water quality intrusion/problems), in particular, has become increasingly important, and requires the division of the distribution system into isolated sectors. All water that enters into an isolated sector is metered and is consumed within it. In the event of a contamination incident, such sectors limit the number of people who might be exposed to the threat, and minimise the number of pipes and their lengths that need to be decontaminated. The process of partitioning a WDN into a set of isolated sectors is referred to as water network sectorisation (WNS), and is the focus of this thesis.

There are three categories of requirements in WNS:

- Structural requirements, which are related to the structure of the network and include direct access to at least one source for each sector (*i.e.*, the path from the sector to a water source must not contain any nodes in other sectors), sector isolation (*i.e.*, there

should be no flow exchange between sectors), sector size limitations, sector size balance, connectedness of the partitioned network, and the requirement that sectors should cross as few mains as possible. The first two requirements are specifically related to sectorisation, while the latter ones are general in WDN partitioning.

- Hydraulic requirements, which are hydraulic objectives and constraints including customer demands satisfaction, pressure requirements, network reliability, energy efficiency, limited water velocity in pipes, minimum nodal elevation differences within sectors, and water quality. These requirements must be satisfied in all types of network design and re-design, including partitioning and sectorisation, and must also be considered during the WNS process.
- Economic requirements, which are related to the cost of sectorisation.

A significant challenge with WNS is that closing links in the network to address sector isolation deteriorates hydraulic behaviour of the network. In particular, network reliability, energy efficiency, and water quality are negatively affected. Graph theory can be used to address the structural requirements, but existing graph-theory techniques cannot address direct access to a source for each sector and sector isolation. Current solutions for WDN partitioning do not address all the requirements of sectorisation. In particular, direct access to a source for each sector and sector isolation are not addressed in most of the existing approaches. Those approaches that do address direct access and sector isolation, cannot handle large networks, as the number of sources is a limiting factor for the number of sectors. The only scalable approach that addresses direct access and sector isolation, does not address some other sectorisation requirements, *e.g.*, pressure requirements during different consumption scenarios, energy efficiency, limited water velocity in pipes, minimum nodal elevation differences within the sectors, and the requirement that sectors should cross as few mains as possible are not addressed.

This thesis explores WNS and proposes a WDN sectorisation method, called WDN-PARTITION, to address various requirements (specially water security requirements) while minimising the negative impact on the other structural and hydraulic requirements.

WDN-PARTITION first satisfies water security-related constraints and generates a collection of feasible solutions (*i.e.*, solutions that address structural requirements of sectorisation) using a novel heuristic graph partitioning algorithm. Then, the best solutions in this initial collection are identified using a many-objective optimisation procedure, based on the hydraulic requirements. The proposed method works well for both small and large networks, as the number of sources is not a limiting factor for the number of sectors.

WDN-PARTITION has been implemented and integrated with a hydraulic network simulator (*i.e.*, EPANET). The simulation-based evaluation assesses the results of the proposed method on a series of publicly available real world water distribution networks. The largest available network has been used as a benchmark to compare the proposed method with three baseline approaches, *i.e.*, a manual engineering approach (Murray et al. 2009) and two automated solutions (Diao et al. 2013 and Ferrari et al. 2014). The results show that WDN-PARTITION is a good alternative for the existing approaches, as it achieves its design objectives to partition a WDN into isolated sectors, satisfying all structural and hydraulic requirements with less than 1% deterioration in network resilience index (a metric for network reliability) and water age (a metric for water quality), comparing to the original network.

Publications related to this Ph.D.

- Hajebi, S., Roshani, E., Cardozo, N., Clarke, A. and Clarke, S. 2016. Water Distribution Network Sectorisation Using Graph Theory & Many-Objective Optimisation, Journal of Hydroinformatics. 18, 1 (2016), 77–95.
- Hajebi, S., Temate, S., Barrett, S., Clarke, A. and Clarke, S. 2014. Water Distribution Network Sectorisation Using Structural Graph Partitioning and Multi-objective Optimization. Procedia Engineering. 89, (2014), 1144–1151.
- Hajebi, S., Barrett, S., Clarke, A. and Clarke, S. 2013. Multi-Agent Simulation to Support Water Distribution Network Partitioning. 27th European Simulation and Modelling Conference - ESM'2013, Lancaster, UK (Lancaster, United Kingdom, 2013).
- Hajebi, S., Song, H., Barrett, S., Clarke, A. and Clarke, S. 2013. Towards a Reference Model for Water Smart Grid. International Journal of Advances in Engineering, Science and Technology. 2, 3 (2013), 310–317.
- Song, H., Raj, A., Hajebi, S., Clarke, A. and Clarke, S. 2013. Model-based cross-layer monitoring and adaptation of multilayer systems. Science China Information Sciences. 56, 8 (Aug. 2013), 1–15.
- Song, H., Raj, A., Hajebi, S., Clarke, S. and Clarke, A. 2013. Model driven engineering of cross-layer monitoring and adaptation. International Conference on Model-Driven Engineering and Software Development, MODELSWARD 2013 (2013).

Table of contents

List of figures

List of tables

Acronyms

ADD
: Average Day Demand.

DMA
: District Metered Area.

FD
: Fire Demand.

ICT
: Information and Communications Technology.

iDMA
: Isolated District Metered Area.

IWA
: International Water Association.

MDD
: Maximum Day Demand.

MHD
: Minimum Hourly Demand.

PA
: Partitioned network.

PHD
: Peak Hourly Demand.

SCADA
: Supervisory Control And Data Acquisition.

WDN
: Water Distribution Network.

WNS
: Water Network Sectorisation.

Glossary

Background leakage

Background leakage is characterised by small non-visible leaks that occur mostly at joints and fittings, not generating sufficient noise to be detected by existing equipment [30].

CAPEX

Capital Expenditure. In the context of this thesis, CAPEX is an expense incurred to create *i*DMAs, *e.g.*, expenditure on assets like pipes, valves, and meters, and the installations costs.

Cut

In graph theory, a *cut* is a partition of the nodes of a graph into two disjoint subsets. Any cut determines a cut-set, the set of edges that have one endpoint in each subset of the partition [14].

Dead-end

A dead-end is the end of a single pipe [119], created by closing off pipes or valves.

Direct access

In this thesis, *direct access* to a water source for each sector means that the path from a node or a link in the sector to a source does not contain any nodes in other sectors. Therefore, if the path contains nodes which are not assigned to any sector, it is still considered as a direct access.

District Metered Area

District Metered Area. A discrete area of a distribution system, in which the quantities of water entering and leaving the area are metered.

EPANET

A free and public-domain software developed by the US Environmental Protection Agency (EPA). It is capable of hydraulic simulation and analysis of a WDN. The computational engine of EPANET is used by many software that developed more powerful, proprietary packages. EPANET is widely used in academy and industry and can be arguably considered as a de facto standard for hydraulic simulations.

Graph

Graphs are mathematical structures used to model pair-wise relations between elements [39]. A graph G is an ordered pair $G = (V, E)$ comprising a set V of vertices or nodes together with a set E of edges or links, which are 2-element subsets of V (*i.e.*, an edge is related with two vertices, and the relation is represented as an unordered pair of the vertices with respect to the particular edge).

Hydraulic head

The total energy associated with a fluid per unit weight of the fluid is called *hydraulic head* (H). The kinetic energy is called velocity head, the potential energy is called elevation head, and the internal pressure energy is called pressure head. While typical units for energy are foot-pounds (Joules), the units of total head are feet (meters) [119].

***i*DMA**

Isolated District Metered Area. A discrete area of a distribution system, in which the quantity of water entering the area is metered and would be consumed within the area. There is no flow exchange between different *i*DMAs. Additionally, each *i*DMA has a direct access to at least one water source; *i.e.*, the flow path from at least one source to each *i*DMA does not pass other *i*DMAs.

Many-objective optimisation

A many-objective optimisation problem is a multi-objective optimisation problem with more than three objectives [28].

MATLAB

Matrix Laboratory. It is a multi-paradigm numerical computing environment and fourth-generation programming language. MATLAB is developed by MathWorks. It allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages.

Multi-objective optimisation

A Multiobjective Optimisation Problem (also called multi-criteria optimisation, multi-performance optimisation, vector optimisation, multi-attribute optimisation, or Pareto optimisation problem) can be defined (in words) as a problem of finding [88]: 'a vector of decision variables which satisfies constraints and optimises a vector function whose elements represent the objective functions. These functions form a mathematical description of performance criteria which are usually in conflict with each other. Hence, the term *optimise* means finding such a solution which would give the values of all the objective functions acceptable to the decision maker.' Multi-objective optimisation is used where optimal decisions need to be taken in the presence of trade-offs between two or more conflicting objectives [20].

NP-complete

In simple words, NP-complete problems are problems with no known fast solution; *i.e.*, the time required to solve the problem using any currently known algorithm increases very quickly as the size of the problem grows. However, any given solution to an NP-complete problem can be verified quickly (*i.e.*, in polynomial time) [51].

OPEX

Operational Expenditure. In the context of this thesis, OPEX is the required day-to-day operation cost of the network, like valve operations, pump operations, maintenance, and repairs.

Optimisation

Optimisation ‘is the act of obtaining the best result under given circumstances’ [101]; in other words, it is ‘the process of determining the best design’ [91].

Pareto front

In multi-objective optimisation, *Pareto front* is the set of all *non-dominated solutions*. Solution A is said to *dominate* solution B if and only if A is no worse than B in terms of all the objective functions, and A is exactly better than B in terms of at least one objective function. Typically, in the optimisation procedure all solutions are compared against each other and if one solution dominates another one, the dominated solution is removed from the candidate solution set. Therefore, all the non-dominated solutions will remain, which build the Pareto front [20, 29].

Resilience

Resilience is defined as the inherent capacity of a water network to manage unexpected failures, and it is measured as ‘the ratio between the surplus of power delivered to users and the maximum power that can be dissipated in the network when meeting exactly the design criteria’ [58].

Subgraph

A subgraph, H, of a graph, G, is a graph whose vertices are a subset of the vertex set of G, and whose edges are a subset of the edge set of G [39].

Water age

Water age is defined as ‘the cumulative residence time of water in the system’ [119], which is regarded as a reliable surrogate measure for water quality. Water age at sources assumed to be zero, so water age can be defined as the time from when water enters the system to when it reaches a point.

Water Distribution Network

Water Distribution Network. The infrastructure that supplies drinking water to homes and businesses, linking water sources and consumers. A WDN is typically a complex systems consisting of thousands of nodes (*i.e.*, *sources* including *reservoirs* and *tanks*, and *consumption nodes* and/or *junctions*), which are interconnected by thousands of links (*i.e.*, *pipes*, *pumps*, and *valves*). A WDN is an important infrastructure to fight fires in the cities as well [119].

Water distribution network reliability

The degree to which the system is able to provide consumers with a minimum acceptable level of supply (in terms of pressure, availability, and water quality) at all times under a range of operating scenarios, under both normal and abnormal conditions [4].

Water network partitioning

Partitioning a WDN into smaller sub-networks is the division of the network into smaller sub-networks [79]. Each sub-network, called a district metered area (DMA) [15], is defined as a discrete area of the distribution system, in which the quantity of water enters and leaves the area is metered [79, 115].

Water network sectorisation

The process of partitioning a WDN into a set of *i*DMAs is called water network sectorisation (WNS) [35]. An *i*DMA is defined as an area of the network in which all the water that enters is metered and is consumed within it; consequently, there is no flow exchange between *i*DMAs. Therefore, sectorisation is defined as a specific type of partitioning in which the resulted divisions (sectors) are isolated from each other; *i.e.*, there is no flow exchange between different sectors, and there is a direct path from each sector to at least one source.

Water network segmentation

Planned (*e.g.*, regular maintenance) and unplanned interruptions (*e.g.*, pipe burst) happen often in water distribution networks, which makes it inevitable to isolate pipes at some occasions. To isolate a pipe in the network, it is required to close off some valves which isolate a portion of the network. Using isolation valves to separate a portion of the water distribution network for management purposes is called WDN segmentation [54]. The main difference between segmentation and partitioning is that in segmentation the existing isolation valves are being used. This is not easy to achieve in most of the cases, as the valve system is not typically designed to isolate each pipe separately [54].

Water security

Water security is defined as the capacity of a population to safeguard sustainable access to adequate quantities of acceptable quality water for sustaining livelihoods, human well-being, and socio-economic development, for ensuring protection against water-borne pollution and water-related disasters, and for preserving ecosystems in a climate of peace and political stability [100, 116].

Chapter 1

Introduction

A water distribution network (WDN) is the infrastructure that supplies drinking water to homes and businesses, linking water sources and consumers. A WDN is also an important infrastructure to support fighting fires in cities. A WDN is typically a complex system consisting of thousands of nodes (*i.e.*, *sources*, including *reservoirs* and *tanks*, *demand nodes*, *e.g.*, consumers, and *no-demand nodes*, *e.g.*, junctions). These nodes are interconnected by thousands of elements (*i.e.*, *pipes*, *pumps*, and *valves*), which may have complex behaviour [119]. The role of a WDN is to supply high-quality water at the required pressure at different times of the day catering for various consumption scenarios for each consumer node.

The layout of a WDN is not typically the result of a sound design; they usually develop to respond to increasing demands that emerge as a result of urbanisation and population growth [45]. Control and management of a WDN as a whole is challenging as there is no granular information about different parts of the network to support location of issues. Additionally, as all the network is connected, in the event of a contamination that spreads, the affected area and the number of people who might be exposed to the threat are potentially large [80].

Partitioning a WDN into smaller sub-networks is a strategy to manage its complexity, as advised by the International Water Association (IWA) [79]. Each sub-network, called a district metered area (DMA) [15], is defined as a discrete area of the distribution system, in which the quantity of water enters and leaves the area is metered [79, 115]. A DMA is usually created by closure of valves or complete disconnection of pipes. The concept of DMA management was introduced to the UK water industry in the 1980's [115]. The main goal was improving the detection and management of leakages; however, there are other advantages, including [45]:

(a) improved contamination spread control, which is associated with *water security* [1],

(b) enhanced demand management,

(c) improved sensor placing,

(d) better rehabilitation and work planning, and

(e) definition of different pressure levels, which helps in the establishment of a permanent pressure control system (known as pressure zones [119]).

1.1 Water security and network sectorisation

Recently, water security (*i.e.*, safeguarding the network from intentional or unintentional water quality intrusion/problems) has gained a lot of attention [82, 98] [2]. Water security requires the division of the water distribution network into isolated DMAs, also known as *i*DMAs. An *i*DMA is defined as an area of the network in which all the water that enters is consumed within it; therefore, there is no flow exchange between *i*DMAs. In the event of a contamination incident, these sectors limit the number of people who are exposed to the threat and minimise the number and the total length of pipes that need to be decontaminated [80]. The process of partitioning a WDN into a set of *i*DMAs is called water network sectorisation (WNS) [35].

[1] Water security is defined as 'the capacity of a population to safeguard sustainable access to adequate quantities of acceptable quality water for sustaining livelihoods, human well-being, and socio-economic development, for ensuring protection against water-borne pollution and water-related disasters, and for preserving ecosystems in a climate of peace and political stability' [100, 116].

[2] The US National Research Council has suggested improving water security as an opportunity for research, which is yet to be addressed (*cf.*, Chapter 2): 'Research on innovative mechanisms to isolate or divert contaminated water in drinking water and wastewater systems would be useful' [82].

This thesis focuses on water network sectorisation. For more clarity, a summary of the definitions of partitioning, sectorisation, and also segmentation is given in the following:

Water network partitioning. Partitioning a WDN into smaller sub-networks is the division of the network into smaller sub-networks [79]. Each sub-network, called a district metered area (DMA) [15], is defined as a discrete area of the distribution system in which the quantity of water enters and leaves the area is metered [79, 115].

Water network sectorisation. The process of partitioning a WDN into a set of *i*DMAs is called water network sectorisation (WNS) [35]. An *i*DMA is defined as an area of the distribution system in which all the water that enters is metered and is consumed within it; consequently, there is no flow exchange between *i*DMAs. Therefore, sectorisation is defined as a specific type of partitioning in which the resulted divisions (sectors) are isolated from each other; *i.e.*, there is no flow exchange between different sectors, and there is a direct path from each sector to at least one source.

Water network segmentation. Planned (*e.g.*, regular maintenance) and unplanned interruptions (*e.g.*, pipe burst) happen often in water distribution networks, which makes it inevitable to isolate pipes at some occasions. To isolate a pipe in the network, it is required to close off some valves which isolate a portion of the network. Using isolation valves to separate a portion of the water distribution network for management purposes is called WDN segmentation [54]. The main difference between segmentation and partitioning is that in segmentation the existing isolation valves are being used.

In the rest of this thesis whenever *partitioning* is used, it refers to the division of the network into smaller sub-networks, whether isolated or not (therefore, *partitioning* is considered as a superset that involves *sectorisation*); however, when the intention is to refer to the creation of isolated sectors, *sectorisation* will be used explicitly. WDN segmentation is not studied in this thesis.

1.2 WNS challenges

Water security requires that the identified sectors must be isolated from each other. A significant challenge with sectorisation is that closing links in the network to address sector isolation deteriorates the hydraulic behaviour of the network [1, 47, 58]. This challenges does exist in WDN partitioning as well, however, it is more severe in sectorisation. This thesis explores the challenge of partitioning a WDN into isolated sectors (*i.e.*, sectorisation) and proposes a WDN sectorisation method, called WDN-PARTITION, to address water security requirements while minimising the negative impact on the other structural and hydraulic requirements.

As it will be detailed in Chapter 2 (*cf.*, Section 2.1), there are three categories of requirements in WNS: (a) structural requirements, (b) hydraulic requirements, and (c) economic requirements, each with its own challenges.

1.2.1 Structural requirements

Structural requirements are related to the structure of the network after sectorisation (*cf.*, Section 2.1). *Sector isolation* requires that there should be no flow exchange between different sectors. The sector isolation requirement implies that each sector must have *direct access* to at least one water source; *i.e.*, the path from the sector to at least one water source must not contain any nodes in other sectors. *Connectedness* of the initial WDN and the identified sectors after partitioning the WDN is another structural requirement. *Sector size* must be within a predefined boundary, usually between 500 and 5000 customer connections [79]. There are also some less-restricting structural requirements, including *sector size balance*, *minimum cut-set* [3] size, *minimum cut-set weight*, and finally, the requirement that *sectors should cross as few mains as possible*.

Sector isolation and direct access requirements are specific to WNS (*i.e.*, to address water security). The other structural requirements are general in WDN partitioning,

[3] In graph theory, a *cut* is a partition of the nodes of a graph into two disjoint subsets. Any cut determines a cut-set, the set of edges that have one endpoint in each subset of the partition [14].

including sectorisation.

Structural challenges

Abstracting out from the hydraulic complexities, the underlying problem of WDN partitioning can be modelled by a graph-theory problem. Graphs are mathematical structures used to model pair-wise relations between elements [39]. Graph theory techniques (*e.g.*, graph partitioning, graph search algorithms, clustering, and community detection) proved to be helpful in addressing problems that deal with networks of connected elements, like water distribution networks [11, 39].

A water distribution network can be modelled by a graph $G = (nodes, links)$ [31]. *nodes* consist of *source* nodes (including *reservoirs* and *tanks*), *consumer* nodes, and *junctions*. *links*[4] consist of *pipes*, *pumps*, and *valves*. Some approaches take advantage of existing knowledge of graph theory to deal with structural aspects of the WNS problem by modelling the WDN with a graph [1, 26, 36, 37, 47, 53, 54, 56, 63, 93, 114].

However, existing graph-theory techniques are not adequate for handling water security related requirements. After abstracting out the hydraulic complexities, the WNS problem can be reduced to three different classic graph theory problems: balanced graph partitioning, constrained clustering, and graph edge deletion. These classic graph-theory problems are all NP-complete, therefore, there is no known exact polynomial-time solution for them; but more importantly, existing techniques for solving them do not address water security related requirements, *i.e.*, *sector isolation*, and *direct access* to sources for each sector. This challenge will be discussed in detail in Chapter 2 (*cf.*, Section 2.1.1).

1.2.2 Hydraulic requirements

Hydraulic requirements are hydraulic objectives and constraints that need to be maintained in water networks, including conservation of mass, conservation of energy, limited

[4] Although pumps and valves are network elements that are not basically links, they are considered as links for modelling purposes.

pressure at nodes [5], limited water velocity in pipes, limited tank levels, customers demand satisfaction or nodal pressure requirements [6], network reliability, energy efficiency, minimum nodal elevation differences within the sectors, water quality, and background leakage (*cf.*, Section 2.1). Violation of one of these requirements could result in damages to the network infrastructure or make problems in service provisioning. These requirements must be satisfied in all types of water distribution network design and re-design, including partitioning and sectorisation [61].

Hydraulic challenges

A significant challenge with WNS is that closing links in the network to address sector isolation deteriorates the hydraulic behaviour of the network. In particular, network reliability, energy efficiency, customers demand satisfaction or nodal pressure requirements, and water quality are negatively affected.

Reduction of network reliability. Creation of highly looped networks is the strategy that is typically applied in water distribution systems design to guarantee high reliability [7] [119]. By closing some pipes using isolation valves or disconnection of pipes, with the purpose of creating DMAs or *i*DMAs, the loop structure of the network reduces. This may cause a decrease in network reliability and make problems in guaranteeing pressure requirements for some nodes [1].

Reduction in energy efficiency. WDN partitioning may cause more energy to be dissipated at the dead-ends [8] which are created as a result of partitioning [1, 36]. This can result in decreased network efficiency, and may cause problems in ensuring pressure requirements at some nodes.

[5] Limited pressure at nodes refers to the minimum and maximum admissible pressure for all nodes (*cf.*, Section 2.1).

[6] Nodal pressure requirements refers to the nodal minimum required pressure at different times during different consumption scenarios (*cf.*, Section 2.1).

[7] Water distribution *network reliability* is defined as the degree to which the system is able to provide consumers with a minimum acceptable level of supply (in terms of pressure, availability, and water quality) at all times under a range of operating scenarios, under both normal and abnormal conditions [4].

[8] A *dead-end* is the end of a single pipe [119]. They are created by closing off pipes or valves.

Water quality issues. Water may remain at the dead-ends created by cutting pipes and/or closing valves to isolate sectors. This can cause water quality issues, especially, *water ageing* [9] [2, 7, 58].

Hydraulic requirements of water distribution networks impose considerations far beyond the abstractions of graph theory techniques. For example, pressure in each node is a function of pressures in other connected nodes. Additionally, direction of flow is determined by hydraulic rules, which could vary during different demand scenarios. Therefore, in applying graph theory algorithms for partitioning water distribution networks, hydraulic requirements and connectivity structure of the network must be considered [47].

Additionally, the partitioned network must also supply a certain amount of high-quality water at the required pressure at different times during various consumption scenarios for each consumer node.

1.2.3 Economic requirements

The total cost of sectorisation should be minimised. Total cost includes both CAPEX (Capital Expenditure) and OPEX (Operational Expenditure). CAPEX is an expense incurred to create *i*DMAs, *e.g.*, expenditure on assets like pipes, valves, and meters, and the installation costs. OPEX is the required day-to-day operation cost of the network, like valve operations, pump operations, maintenance, and repairs. The total running cost of the sector creation and operation (CAPEX + OPEX) should be minimised in WDN partitioning projects.

1.3 Existing solutions

Existing solutions for WDN partitioning do not address all the requirements of WNS. In particular, *sector isolation* and *direct access* to sources for each sector are not addressed

[9] *Water age* is defined as 'the cumulative residence time of water in the system' [119], which is regarded as a reliable surrogate measure for water quality. Water age at sources assumed to be zero, so water age can be defined as the time from when water enters the system to when it reaches a point.

in [2, 5, 13, 16, 26, 27, 33, 34, 34, 37, 56, 57, 76, 79, 81, 93, 94], therefore, they cannot address water security. Those approaches that do address direct access and sector isolation, *i.e.*, Tzatchkov et al. (2006) [114], Fernández et al. [45, 65, 66] and Di Nardo et al. [36], cannot handle large networks. In theses three approaches, the number of sectors that can be created equals the number of main sources, therefore, for large networks some of the resulting sectors may be larger than the maximum allowable size (especially if the size of the network divided by the number of sources is larger than the maximum allowable sector size). The only scalable approach that addresses direct access and sector isolation, *i.e.*, Ferrari et al. [47, 104], does not address some other WNS requirements, *e.g.*, pressure requirements during different consumption scenarios, energy efficiency, limited water velocity in pipes, minimum nodal elevation differences within the sectors, and the requirement that sectors should cross as few mains as possible are not addressed. Chapter 2 will discuss the state of the art in WDN partitioning in detail.

1.4 Thesis approach

Based on valuable insights from the existing solutions [47, 104], this thesis proposes a novel heuristic WDN partitioning technique, called WDN-PARTITION, to partition a water distribution network into isolated sectors guaranteeing that each sector is of a good size and is directly connected to a water source, while minimising the negative impact on hydraulic requirements as far as possible.

Observation. As discussed, WNS is a multifaceted problem embracing graph-theoretic, hydraulic, and economic aspects. Existing techniques in graph theory cannot handle structural challenges; therefore, there is a need to develop an algorithm that takes the structural requirements of the problem into account. Additionally, besides the structural requirements, WNS has hydraulic and economic requirements (*i.e.*, constraints that must be satisfied and objectives that should be optimised), therefore, it is a constrained multi-objective optimisation problem [10].

[10] Optimisation 'is the act of obtaining the best result under given circumstances' [101]; in other words, it is

Strategy. In this thesis, a set of WNS *requirements* and *metrics* are first defined to reflect WNS structural, hydraulic, and economic requirements. The requirements are considered in the sectorisation process, and the outcome of sectorisation is evaluated using the defined metrics. This thesis develops a heuristic technique, which is specific to water network sectorisation, and takes full advantage of the particularities of the problem.

Basic idea. To achieve the best results under the given circumstances, WDN-PARTITION first satisfies water security specific requirements (*i.e.*, sector isolation and direct access) and generates a collection of structurally-feasible solutions using a novel heuristic graph partitioning algorithm. Then, the best solutions are identified considering the values for their objective functions (*i.e.*, the metrics for evaluation). Structurally-feasible solutions are the solutions that address structural requirements of sectorisation, especially sector isolation, direct access, connectedness, and sector size.

WDN-PARTITION first identifies sources and major flow paths (the main links that are directly connected to sources) [11], then identifies groups of nodes that are connected to the major flow paths but are isolated from the rest of the network, and finally checks for the size of the identified groups, and partitions them if needed. In partitioning a large group, WDN-PARTITION selects nodes which are already on the major flow paths or identifies nodes which are connected to the major flow paths directly, and grows subgraphs (using the BFS algorithm [96]) starting from the selected nodes. If all the subgraphs are of a good size, they are considered as sectors. This guarantees that each identified sector is directly connected to a source. The details of the proposed method is discussed in Chapter 4.

1.5 Thesis contribution

This research contributes to the body of knowledge by providing:

'the process of determining the best design' [91].

[11] In this thesis, when it is said a sector is *directly connected* to a water source, it means that the path from a node or a link in the sector to at least one water source does not contain any nodes or links in other sectors. Therefore, if a path contains nodes or links which are not assigned to any sector, it is still considered as a direct path.

Extension to WNS requirements and metrics. To perform sectorisation correctly, a general set of WNS requirements must be applied in the process. Additionally, a general set of metrics is needed to quantify the different aspects of the sectorisation results and to support evaluations. Existing design criteria and metrics do not cover all aspects of WNS (*cf.*, Chapter 2). This study extends the current design criteria [79] and proposes a more general set of *i*DMA design criteria (WNS requirements), which are derived from the structural and hydraulic requirements of water network sectorisation. In addition to the WNS requirements, this study extends the metrics to quantify the performance of a partitioning scheme, so enhancing the measures to compare different approaches [12]. The problem statement (*cf.*, Chapter 3) complements this contribution by providing a mathematical formulation of the WNS problem as a constrained multi-objective optimisation problem, which can be tackled by other researchers or professionals.

Extension to existing WDN partitioning methods. Current solutions for WDN partitioning do not address all the requirements of WNS. In particular, water security is not addressed in most of the existing approaches as comprehensive as it is addressed in this thesis (they cannot address sector isolation and direct access to a water source for each sector). Those approaches that do address water security cannot handle large networks (as the number of sources is the limiting factor for the number of sectors), and/or do not address other WDN partitioning requirements (*e.g.*, pressure requirements during different consumption scenarios, energy efficiency, limited water velocity in pipes, minimum nodal elevation differences within the sectors, and the requirement that sectors should cross as few mains as possible). WDN-PARTITION partitions a water network into isolated sectors guaranteeing that each sector is of a good size and is directly connected to a water source, while minimising the negative impact on the hydraulic requirements.

Extension to graph partitioning techniques. As discussed briefly in Section 1.2.1 and will be detailed in Chapter 2 and 3, water network sectorisation can be reduced to a graph partitioning problem after removing the hydraulic details and complexities. However, no

[12]SSI (sector size imbalance), ALnkExp (average pipe length per sector), and MLnkExp (maximum pipe length in a sector), and ED (elevation differences within sectors) are proposed in this thesis.

graph partitioning method that addresses component/sector [13] isolation and direct access to some specific nodes (*sources*) for each component/sector was found in the course of this research. A heuristic digraph partitioning technique is proposed in this thesis that partitions a digraph with two types of nodes (*i.e.*, *sources* and *consumers*), which guarantees component isolation and direct access to at least one *source* for each component, while minimising the cut-set size, cut weight, and sector size imbalance (*cf.*, Chapter 4). However, this method has not been designed for general graphs/networks and only targeted for specific digraphs with two types of nodes (*i.e.*, *sources* and *consumers*) that have the discussed structural characteristics and requirements.

1.6 Thesis scope

This thesis studies only water distribution networks; general networks/graphs are not studied, and generalisation of the proposed method to other types of networks is out of the scope of this thesis. Additionally, this study assumes a previously designed WDN as input, and suggests a set of near-optimal partitioning solutions for the input network as output (that can be ordered in terms of costs, pressure violations, or other criteria). Design and calibration of the networks are also out of the scope of this thesis. In addition, the required pressure is assumed to be given for each node.

In the sectorisation method, only closure of links (pipes and valves) is studied. It might be useful in some situations to add new components to the network; however, it is out of the scope of this thesis. Additionally, dynamic pump and/or valve operations optimisation might prove to be effective in DMA management; nonetheless, they are not studied in this research. Finally, infrastructure conditions (*e.g.*, the layout of roads, streets, bridges, etc.) are also not studied in this thesis.

[13] A *component* is the term used in the graph theory literature [14] for the smaller subgraphs resulted after partitioning a graph. A *sector* is the same notion, which is a smaller sub-network resulted after partitioning a water distribution network.

1.7 Thesis structure

State of the art Chapter 2 analyses how the state of the art in water distribution network partitioning methods address the requirements of WDN partitioning and WNS. It first elaborates the WNS problem and proposes a general set of WNS requirements, which are derived from the structural, hydraulic, and economic requirements of the problem. Then it analyses the existing approaches and discuses how well they address the WNS requirements.

Problem statement Chapter 3 returns to the characteristics of water network sectorisation (WNS) problem and corresponding challenges to frame the problem addressed in this thesis. Based on the discussed requirements in Chapter 2, this chapter introduces design objectives and constraints for WNS. Then, it formulates the WNS problem as a constrained many-objective optimisation problem and explains decision variables, constraints, and objectives. Decision variables are the specific elements in the problem that can be controlled. Constraints are the situations that should be maintained in the system. Objectives are mathematical representation of the proposed metrics, which quantify how good a solution is in terms of the corresponding criteria.

Design and implementation Chapter 4 first discusses design decisions related to different aspects of the WNS problem to motivate the need for the novel method proposed in this thesis. Then, it explains how the proposed solution (*i.e.*, WDN-PARTITION) integrates the design decisions and addresses the requirements and challenges of the problem. This chapter also explains the algorithms that compose the proposed solution, how they are implemented, and their integration with the hydraulic simulator (EPANET [103]) enabling the assessment of the candidate solutions from the hydraulic perspective.

Evaluation Chapter 5 evaluates how well the proposed approach achieves its objectives in partitioning a WDN into a collection of *i*DMAs with minimal impact on the hydraulic requirements. It first describes the experimental setup of the simulation-based evaluations. Then, it presents and analyses the results of applying WDN-PARTITION on seven real WDNs as case studies. Finally, WDN-PARTITION is compared to the existing WDN partitioning methods using the largest network available in the literature. The analysis shows WDN-PARTITION is a suitable alternative for the existing approaches for WDN partitioning.

Conclusion and future work Chapter 6 summarises the thesis and its achievements. It then discusses important findings with regard to the proposed sectorisation method and highlights potential areas for future work.

1.8 Chapter summary

Water distribution network sectorisation problem is characterised by structural and hydraulic requirements that make existing graph partitioning techniques inadequate for finding a good solution. Specifically, sector isolation and direct access to at least one source for each sector are not addressed. Existing solutions for WDN partitioning do not address all the requirements of water network sectorisation. In particular, structural requirements are not addressed in most of the existing approaches, therefore, they cannot handle the WNS problem.

This chapter gave an overview of the whole thesis by introducing the problem and the related challenges, thesis approach, contributions, and its scope. The following chapters review the state of the art, formulate the WNS problem as a many-objective optimisation problem, and describe how WDN-PARTITION solves the problem of partitioning a WDN into isolated sectors with minimum negative impact on the hydraulic requirements of the network.

Chapter 2

State of the art

Water distribution network partitioning is a strategy to manage its complexity, enhance leakage detection, and improve water security. Water security imposes the division of a distribution system into isolated sectors, or *i*DMAs. All water that enters a sector is consumed within it, *i.e.*, there is no flow exchange between different sectors. A specific type of WDN partitioning in which all the identified sectors are isolated from each other (*i.e.*, no flow exchange between sectors), is called water network sectorisation (WNS).

This chapter reviews the state of the art in water distribution network partitioning and sectorisation. It first gives an overview of the problem and justifies why the current graph partitioning methods are not capable of solving it. Then, it details the WNS requirements, which were introduced briefly in Section 1.2, and uses them as a framework to analyse the existing approaches.

2.1 WNS requirements

This section gives a detailed discussion of WNS requirements and proposes a general list covering different requirements that are considered in different approaches. Then, this list is utilised to analyse the state of the art.

To perform sectorisation efficiently, a general set of WNS requirements needs to be considered in the process [79]. Additionally, a general set of metrics is needed to quantify the different aspects of the sectorisation results and to support evaluations. Existing approaches each address a specific set of WNS requirements and there is no agreed general set of requirements. This thesis extends the current requirements [79] and proposes a more general set of WNS requirements [61], which are derived from the structural, hydraulic, and economic requirements of water network sectorisation, using the following methodology:

1. First, the International Water Association (IWA) guidelines on DMA design [79] are considered; including size (geographical area and the number of customer connections), elevation of nodes, pressure requirements, number of pipes to be cut, number of meters to be installed, infrastructure conditions, and minimum number of mains crossed by sectors.

2. Then, the initial list of requirement was complemented with those discussed in the state of the art in WDN partitioning, *i.e.*, [2, 5, 13, 16, 17, 17, 26, 27, 34, 36, 36, 41, 45, 53, 56–58, 65, 69, 76, 79–81, 93, 94, 104, 106, 114].

3. Finally, the list of requirements was discussed and modified with experts and practitioners in WDN partitioning.

The outcome of this process is the following set of *WNS requirements* [61], which are as comprehensive as possible within the scope of this research:

1. *Water security* requirements;

   1.1. *Sector isolation* (SR1); *i.e.*, no flow exchange between different sectors is allowed.

   1.2. *Direct access* (SR2) to a water source for each sector must be guaranteed; *i.e.*, the path from a sector to at least one source must not contain any nodes in other sectors.

2. *Connectedness* (SR3); *i.e.*, the partitioned sectors in the WDN should be connected to guarantee that all the nodes have access to water. There must be a flow path from (at

least) a source to all the nodes in an identified sectors, and there should be no isolated node in the network after partitioning.

3. *Conservation of mass* (HR1) and *conservation of energy* (HR2); two fundamental hydraulic constraints that govern the physics of hydraulic networks and should be considered in all water distribution network design and re-design problems. *Conservation of mass* can be regarded as continuity of flow constraint, which dictates that the fluid mass entering any pipe will be equal to the mass leaving the pipe (since fluid is typically neither created nor destroyed in hydraulic systems). *Conservation of energy*, can be regarded as hydraulic equilibrium equations, which imposes that the difference in energy between two points in a network must be the same irrespective of the flow path [119].

4. *Pressure requirements* (HR3 and HR4); *i.e.*, the pressure at nodes should not exceed certain limits (HR3) [119], *i.e.*, a minimum of 28 (m) and a maximum of 100 (m) [117]. Additionally, minimum nodal pressure requirements should be guaranteed (HR4). Both of these requirements should consider the following scenarios [79, 119]:

   - Average Day Demand (ADD),
   - Minimum Hourly Demand (MHD),
   - Peak Hourly Demand (PHD),
   - Maximum Day Demand (MDD), and
   - Fire Demand (FD).

   It should be noted that from the hydraulic point of view, whenever a system is able to satisfy MDD it will be able to satisfy MHD, PHD and ADD as well. For water quality modelling mainly ADD is used. Therefore, MHD, PHD, MDD+FD cover all the five scenarios, as FD is modelled with MDD to ensure the capability of the system to provide fire demand at the maximum day demand, and a system with such a capacity will definitely satisfy ADD [61] [1].

5. *Sector size*s must be balanced and within a predefined boundary (SR4) [79]. Boundaries like (300-2500), (500-3000), and (500-5000) customer connections can be found in the literature [32, 45, 79]; however, (500-5000) is the most common one. Additionally, it is preferable to have sectors of roughly the same size (SR5), to minimise the difference between the large and the small sectors.

6. *Customer demand satisfaction*; *i.e.*, enough water should be supplied to satisfy different customer demands at all times during the different consumption scenarios [38]. If pressure requirements (HR4) are satisfied, this requirement will be satisfied as well.

7. *Elevation* of the nodes in a sector should be within a specific range (HR5) [79], to remove or minimise the need for installation of pumps and/or pressure reducing valves (PRVs).

8. Limited *water velocity* (HR6); *i.e.*, the minimum and maximum velocity of water in pipes should be within a certain limits [56]. Typically a minimum of 1 m/s (3 ft/s) for large zones, and a maximum of 3 m/s (10 ft/s) are recommended [119].

9. Minimum *cut-set size* (SR6); *i.e.*, the number of pipes that must be cut (and equipped with flow meters or valves) during the partitioning process should be minimised [79].[2]

10. Minimum *cut-set weight* (SR7); *i.e.*, the sum of diameters of the pipes that must be cut should be minimised.

11. *Tank level limitations* (HR7); *i.e.*, the level of water in tanks should be within a boundary. The maximum allowable level is typically the top of the tank, and the

[1] It should be noted that 'limited pressure at nodes' (HR3) refers to global constraints that applies to all the nodes uniformly, regardless of the node, time, and scenario; *e.g.*, the minimum of 28 (m) and the maximum of 100 (m) for all nodes; however, 'nodal pressure requirements' (HR4) refers to the specific requirements of the specific nodes at specific times in specific scenarios, *e.g.*, a factory (which is a node) may need 80 (m) at some specific time during its pick consumption scenario, and 30 (m) in it's low demand hours.

[2] In graph theory, a *cut* is a partition of the nodes of a graph into two disjoint subsets. Any cut determines a cut-set, the set of edges that have one endpoint in each subset of the partition [14].

minimum allowable water level is typically above the bottom of the tank to provide some residual storage for potential fire defeat events. Additionally, the difference between the water level at the start and the end of simulation period should be limited (usually less than 10 percent of the tank height).

12. Maximum *network reliability* (HR8); *i.e.*, the reliability of the network should be maximised. Network reliability is defined as the degree to which the system is able to provide consumers with a minimum acceptable level of supply (in terms of pressure, availability, and water quality) at all times under a range of operating scenarios, under both normal and abnormal conditions [4]. Resilience index [113] is a measure of network reliability, and defined as 'the ratio between the surplus of power delivered to users and the maximum power that can be dissipated in the network when meeting exactly the design criteria' [58]. In fact, resilience index quantifies the inherent capacity of a water network to manage unexpected failures [80].

13. Maximum *network efficiency* (HR9); *i.e.*, the energy efficiency of the network should be maximised. *Dissipated power* is a measure of energy efficiency [35], and is defined as the power dissipated in the pipes during the time water flows from the sources to the users; *i.e.*, the total available power at the entrance of the distribution network minus the power that is delivered to the users. Dissipated power should be minimised to have the maximum energy efficiency in the network.

14. *Water quality* considerations (HR10) [80]. *Water age* is usually considered as a reliable surrogate measure for water quality [119], which should be minimised.

15. Minimum *background leakage* (HR11). Background leakage is characterised by small non-visible leaks that occur mostly at joints and fittings, not generating sufficient noise to be detected by existing equipment [30]. It should be minimised in the partitioned network [17, 46].

16. Sectors should *cross as few mains as possible* (SR8) [79]. The sector boundaries

should follow the natural geographic and hydraulic boundaries.

17. Minimum *costs*: CAPEX (ER1) which is capital expenditure, is an expense incurred to create sectors, *e.g.*, expenditure on assets like pipes, valves, meters, and installations. OPEX (ER2) which is operational expenditure is the required day-to-day operation cost of the network, like valve operations, pump operations, maintenance, and repairs. The total cost of sector creation and operation (CAPEX + OPEX) should be minimised.

The priorities for these requirements may depend on the local situations, policies and legislations. However, as this thesis focuses on water network sectorisation, water security requirements (SR1 and SR2) are highly prioritised.

These requirements can be classified into three different categories:

(a) **Structural requirements**, which are related to the structure of the network after partitioning, including sector isolation (SR1), direct access (SR2), connectedness (SR3), sector size (SR4), sector size balance (SR5), minimum cut-set size (SR6), minimum cut-set weight (SR7), and finally, the requirement that sectors should cross as few mains as possible (SR8). Of these structural requirements, SR1 and SR2 are specific to WNS (*i.e.*, to address water security); while the other ones are general in WDN partitioning, including sectorisation.

(b) **Hydraulic requirements**, which are hydraulic constraints and objectives that need to be maintained in water networks, including conservation of mass (HR1), conservation of energy (HR2), limited pressure at nodes (HR3), limited water velocity in pipes (HR6), limited tank level (HR7), customers demand satisfaction or guaranteeing nodal pressure requirements (HR4), network reliability (HR8), energy efficiency (HR9), minimum nodal elevation differences within the sectors (HR5), water quality (HR10), and background leakage (HR11). Violation of one of these requirements could result in damage to the network infrastructure or create problems in service provisioning. These requirements must be satisfied in all types of water distribution network design and re-design, including partitioning and sectorisation.

(c) **Economic requirements**, including minimising CAPEX (ER1) and OPEX (ER2).

Table 2.1 summarises the requirements and their coding used in this thesis.

	Requirement	**Coding**
Structural Requirements	Sector Isolation	SR1
	Direct Access	SR2
	Connectedness	SR3
	Sector Size	SR4
	Sector Size Balance	SR5
	Cut Size	SR6
	Cut Weight	SR7
	Sectors Cross Few Mains	SR8
Hydraulic Requirements	Conservation Of Mass	HR1
	Conservation Of Energy	HR2
	Limited Pressure	HR3
	Pressure Requirements	HR4
	Elevation Differences	HR5
	Limited Velocity	HR6
	Limited Tank Levels	HR7
	Network Reliability	HR8
	Energy Efficiency	HR9
	Water Quality	HR10
	Background Leakage	HR11
Economic Requirements	CAPEX	ER1
	OPEX	ER2

Table 2.1 Requirements coding.

2.1.1 Why graph theory is inadequate for WNS

Abstracting out from the hydraulic complexities, the underlying problem of WDN partitioning can be modelled as a graph-theory problem. Graphs are mathematical structures used to model pair-wise relations between elements [39]. Graph theory techniques (*e.g.*, graph partitioning, graph search algorithms, clustering, and community detection) proved to be helpful in addressing problems that deal with networks of connected elements, like water distribution networks [11, 39].

A water distribution network can be modelled as a graph $G = (nodes, links)$ [31]. *nodes* consist of *source* nodes (including *reservoirs* and *tanks*), *consumer* nodes, and *junctions*.

links[3] consist of *pipes*, *pumps*, and *valves*. Some approaches take advantage of existing graph theory knowledge to deal with structural aspects of the WNS problem by modelling the WDN as a graph [1, 26, 36, 37, 47, 53, 54, 56, 63, 93, 114].

However, existing graph-theory techniques are not sufficient for handling water security related requirements, *i.e.*, sector isolation (SR1), and direct access to a source for each sector (SR2). After abstracting out from the hydraulic complexities, the WNS problem can be reduced to three different classic graph theory problems:

Balanced graph partitioning. A balanced graph partitioning problem can be stated as: given a graph G with n nodes, in a (k,ν) balanced partitioning problem the aim is to partition G into k components of size $\nu \cdot (n/k)$ at maximum, while minimising the number of edges between different components [3]. For $\nu = 1$, the sizes of all the components are exactly equal, which is called a *perfect balance* partition [4]. There are also graph partitioning problem formulations in which *balance* is not directly set in the problem statement but incorporated into the objective functions [14]. While a balanced graph partitioning problem has overlaps with the structural requirements of the WNS problem, no balanced graph partitioning method that addresses direct access to some specific nodes (*sources*) for each component was found in the course of this research. Balanced graph partitioning is NP-complete [5] [3].

Constrained clustering. Constrained clustering aims at finding clusters that satisfy user-specified constraints. Although the constraints can be of any type in the general form of the problem, the existing constrained clustering methods mostly deal with two types of constraints: *must-link*, which imposes that two nodes must be in the same cluster, while *cannot-link* enforces that two nodes must not be in the same cluster

[3]Although pumps and valves are network elements that are not basically links, they are considered as links for modelling purposes.

[4]ν controls the maximum component size and is always greater or equal to 1. As an example, if the number of nodes in a graph is 1000 ($n = 1000$), $k = 4$, and $\nu = 1.1$, the maximum component size can be $1.1 * 1000/4 = 275$.

[5]NP-complete problems are problems with no known fast solution. In other words, the time required to solve the problem using any currently known algorithm increases very quickly as the size of the problem grows. However, any given solution to an NP-complete problem can be verified quickly (*i.e.*, in polynomial time) [51].

[22–24, 73, 118]. However, in a WNS problem, the structural constraints are different from these constraints, as they are not defined at the node level. The must-link constraint cannot completely reflect direct access to a source for each sector requirement (SR2), as it only enforces that two specific nodes must / must not be in the same cluster. Nonetheless, no constrained clustering method that addresses direct access to some specific nodes (*sources*) for each component (SR2) and sector isolation (SR1) was found in the course of this research. Constrained graph clustering problems with cannot-link constraints are NP-complete [73]. In many situations, determining whether there is a solution that satisfies all the constraints is also NP-complete [22]. Clustering is also known as community detection [50].

Graph edge deletion. The general edge deletion problem can be stated as follows: given a specific graph property, find the minimum number of edges whose deletion results in a subgraph satisfying the same property [123]. Edge deletion is not a partitioning problem. Although one may add partitioning constraints to it, no edge deletion problem was found in the course of this research that considers structural requirements of WNS, especially sector isolation (SR1) and direct access to some specific nodes (*sources*) for each component (SR2). Edge deletion, in its general form [6], is NP-complete [123].

These classic graph-theory problems are all NP-complete, therefore, there is no known exact polynomial-time solution for them; but more importantly, existing formulated problems and techniques for solving them do not address water security related requirements, *i.e.*, sector isolation (SR1), and direct access to some specific node (*sources*) for each sector (SR2). As discussed, besides these requirements, there are other structural, hydraulic, and economic requirements that should be addressed in WNS. A couple of approaches have been proposed to address these requirements, which will be discussed in the following section.

[6] If the given property belongs to a rather broad class of properties that are hereditary on induced subgraphs.

2.2 Water distribution network partitioning approaches

A number of approaches have been proposed for partitioning a WDN into a collection of sectors aiming at satisfying some specific requirements of the problem. The literature can be classified into three main categories: (a) *graph theory based*, (b) *complex networks based*, and (c) *multi-agent systems based*.

2.2.1 Graph theory approaches

Graph theory has been used as a central technique in most of the WDN partitioning approaches, including Tzatchkov et al. [114], Gomes et al. [56], Di Nardo et al. [36, 37], Ferrari et al. [47], Alvisi and Franchini [1], and De Paola et al. [26]. [7]

Tzatchkov et al. (2006)

Tzatchkov et al. [114] propose a method for partitioning a WDN into independent sectors (*i.e.*, sectors that are supplied exclusively from their own water sources, and they are not connected to other sectors in the network [114]). In particular, the depth-first search algorithm (DFS) [111] is used to find independent sub-networks in the network layout, and the breadth-first search algorithm (BFS) [96] is applied to examine disconnected nodes from water sources. Then, a source to node contribution analysis is performed, using flow direction in the network pipes, by applying an algorithm similar to breadth-first search.

Of the WNS requirements, SR1, SR2, and HR10 are addressed in this approach; while SR4, SR5, SR6, SR7, SR8, HR3, HR4, HR5, HR6, HR7, HR8, HR9, HR11, ER1 and ER2 are not addressed.

Gomes et al. (2012)

Gomes et al. [56] propose an approach to divide a WDN into a number of sectors. The method uses two operational models and a hydraulic simulator to recognise the optimal

[7] Giustolisi et al. [54] study the application of graph theory in WDN segmentation, but as segmentation is not the focus of this thesis, it will not be discussed in this section.

number of sectors, their entry points, and their boundary valves. The first model uses the Floyd–Warshall algorithm [49] to split the network into appropriate sectors with the following steps: first, the WNS requirements including maximum sector size (SR4), maximum elevation difference within a sector (HR5), and implicit / explicit constraints (user-defined settings that prevent the extension of a sector in a given direction) are characterised. The implicit constraints are associated with network elements like reservoirs, tanks, PRVs, and pumps. The explicit constraints are related to the natural geographic or hydraulic boundaries, flow paths, and pipe flow capacities. Next, the flow paths between each source and any node of the WDN are identified, using the Floyd-Warshall algorithm and the flow direction in the daily peak flow. Then, the reference nodes to grow each sectors are selected. After that, the sectors are extended in all possible directions from downstream to upstream (guided by the peak flow paths and the design criteria for the sectors). Then, if possible, sectors are grouped to decrease the number of entry points and boundary valves. Finally, the sector boundaries are manually adjusted if it is necessary [8]. The second model takes advantage of a simulated annealing algorithm [72] to detect the optimal number and the position of the entry points and the boundary valves, to minimise the cost of partitioning.

Despite its advantages, *i.e.*, addressing sector size (SR4), minimum cut-set weight (SR7), limited water velocity (HR6), normal daily pressure requirements (only ADD in HR4), nodal elevation (HR5), and costs of DMA implementation or CAPEX (ER1), the most important WNS requirements, *i.e.*, sector isolation (SR1) and direct access (SR2), along with SR5, SR6, SR8, HR3, HR4 (for MHD, PHD, MDD, and FD), HR7, HR8, HR9, HR10, HR11, and ER2 are not addressed in this method.

Di Nardo et al. (2013b)

Di Nardo and Di Natale [32] propose a design support methodology which is based on a multilevel recursive bisection algorithm [71] to identify the location of flow meters and boundary valves required to describe sectors. In another work [37], this methodology has

[8] It is not clear in this work that in what situations it is necessary to manually adjust the boundaries.

been tested on a real case study using some performance criteria (*i.e.*, energy index, pressure index, and the flow deficit index). The methodology enables characterization of the optimal water network partitioning harmonious with the level of service that the users need.

The proposed methodology is as follows: first, a hydraulic simulation is performed in the peak hourly demand (PHD) scenario. Next, the number of sectors (k) is chosen, and a multilevel recursive bisection (MLRB) method is applied to find a k-way partitioning. After a set of edge-cuts (or boundary pipes) is found, the method chooses the number of boundary pipes that must be equipped with either gate valves or flow meters (boundary pipes = gate valves + flow meters). Then, a heuristic optimisation technique based on genetic algorithms is used to find the pipes that should be equipped with gate valves or meters by minimising the total dissipated power in the network. Then, three performance indices are computed: (a) *energy index*, characterised by the resilience index (*cf.*, Chapter 3), by comparing the dissipated power and the maximum power required to fulfil the nodal demand requirements, and the resilience deviation index, by comparing the resilience indices of the original and the partitioned networks; (b) *pressure index*, which is conventionally characterised by mean, maximum, minimum, and standard deviation of node pressure (HR3); and (c) *flow deficit index*, which is calculated using a pressure driven analysis (PDA) approach. A partitioning arrangement (*i.e.*, a solution for partitioning) is selected based on the preferred level of service for the consumers by repeatedly changing the number of sectors (k) or decreasing the number of flow meters [37].

Of the WNS requirements, minimum cut-set size (SR6), limited pressure HR3, normal daily pressure requirements (PHD in HR4), network reliability (HR8), and energy efficiency (HR9) are considered in this method. Therefore, the most important WNS requirements, *i.e.*, sector isolation (SR1) and direct access (SR2), along with SR4, SR5, SR7, SR8, HR4 (for ADD, MHD, MDD, and FD), HR5, HR6, HR7, HR10, HR11, ER1 and ER2 are not addressed in this method (SR6 can be considered as a proxy for ER1).

Di Nardo et al. (2013c)

Di Nardo et al. [36] proposed a method to sectorise a WDN based on graph theory techniques (*i.e.*, a DFS method to search independent branches of the WDN), and energy concerns to minimise dissipated power. The proposed method finds independent sectors using a depth first search (DFS) algorithm, which allows for the identification of all possible independent sectors starting from each source node (root nodes) in the network. Then it determines the hierarchical level (HL) of the graph related to each source. Then, the independent and common node collections for each hierarchical level of the graph are found. Finally, the necessary controlling gate valves are identified to isolate sectors, using a heuristic genetic algorithm optimisation technique which minimises the dissipated power.

This method addresses water security requirements, *i.e.*, sector isolation (SR1) and direct access (SR2); however, as the number of sectors created equals the number of main sources, if it is applied to a large network, some of the resulting sectors may be larger than the maximum allowable size, especially if the size [9] of the network divided by the number of sources is larger than maximum allowable sector size. Additionally, network reliability (HR8) and energy efficiency (HR9) are also addressed. However, SR4, SR5, SR6, SR7, SR8, HR3, HR4, HR5, HR6, HR7, HR10, HR11, ER1 and ER2 are not addressed in this method.

Ferrari et al. (2014, 2015)

Ferrari et al. [46, 47, 104] proposed a method to partition a WDN into isolated sectors, which allows for setting the required number of sectors. The method first does a preliminary analysis of the WDN, to identify the transmission mains [10] using the size of the pipes, the independent districts (*i.e.*, groups of nodes linked to the transmission main with no connections with any other group), and the number of sectors to create. The

[9] *cf.*, Chapter 3, Section 3.3.1, page 42.

[10] A piping system is often categorised into *transmission mains* and *distribution mains*. Transmission mains consist of components that are designed to transport large amounts of water over large distances, typically between major facilities within the system. Individual customers are usually not served from transmission mains. Distribution mains are smaller in diameter than transmission mains, and typically follow the general topology and alignment of the city streets [119].

independent districts are identified using a BFS algorithm, starting from the nodes on the transmission mains. Then, for large independent districts, a recursive bisection algorithm is used to partition them and determine the sector boundaries, considering the design criteria, *i.e.*, sector size (SR4), direct access to a water source for each sector (SR2), and sector isolation (SR1). Finally, a hydraulic simulation is performed to examine if the minimum pressure requirements (HR4) for each node is satisfied. This method identifies the near-optimal solutions considering cut-set size (SR6), water quality (HR10), and network reliability (HR8) as objective functions. Cut-set (SR6) has been used as a surrogate for cost of creating DMAs (CAPEX).

Water security requirements (SR1 and SR2) are addressed in this method. Additionally, SR4, SR6, HR4 (for ADD), HR8, HR10, HR11, and ER1 (SR6 is considered as a proxy for ER1) [46] are addressed in this method. However, this work has some limitations. SR5, SR7, SR8, HR3, HR4 for four scenarios (MHD, PHD, MDD, and FD), HR5, HR6, HR7, HR9, and ER2 are not addressed.

Alvisi and Franchini (2014)

Alvisi and Franchini [1] propose a heuristic approach for WDN partitioning which applies BFS and Dijkstra [40] algorithms to find the shortest paths in a graph. The aim is to find a near optimal solution regarding (a) assigning the network nodes to a pre-defined number of sectors, (b) finding the links that should be equipped with flow meters, and (c) finding the links that should be equipped with isolation valves. The method first generates a broad set of partitioning solutions, *i.e.*, allocation of the nodes to different sectors of proper sizes (SR4), and placement of the flow meters and isolation valves. BFS algorithm is used for this purpose. Then, a narrow set of solutions are selected for hydraulic analysis. To this end, the Dijkstra algorithm is used to calculate the shortest weighted distance of each node from the supply points. The solutions with the lowest cumulative distance of the nodes to sources are selected. Finally, after hydraulic analysis, the optimal solutions are identified considering network reliability (HR8) as the objective function.

Although the pressure requirement (HR4) is not considered as a design requirement in this method, the authors claim that the solutions are good regarding to minimum pressures for the peak hourly demand (PHD) and fire demand (FD) situations (based on a case study). Additionally, SR4, SR7, and HR8 are addressed in this method. However, the most important WNS requirements, *i.e.*, sector isolation (SR1) and direct access (SR2), along with SR5, SR6, SR8, HR3, HR4 (for ADD, MHD, and MDD), HR5, HR6, HR7, HR9, HR10, HR11, ER1 and ER2 are not addressed in this method.

De Paola et al. (2014)

De Paola et al. [26] propose an approach similar to k-means clustering [64], combined with a multi-objective optimisation algorithm (*i.e.*, NSGA-II [29]) to partition a WDN into a predefined number of sectors. There are two objectives to minimise: (1) the total operative cost (ER2) for the partitioned network, which involves also the water leakage (HR11) costs and the energy consumed by pumps, if present; and (2) the *resilience deviation index* as proposed by [32], which quantifies the change in the network hydraulic reliability (HR8) before and after sectorisation. The constraints include: (a) minimum required pressure at demand nodes (HR4), (b) maximum budget (ER1), and (c) the sector sizes (SR4).

The approach has the following steps: first the desired number of sectors, k, is selected. Then, k nodes are selected at random as centroids, and the sectors are created by adding every node in the network to the sector of the closest centroid. The distances on the network graphs are computed as the lengths of the shortest paths between each pair of nodes which are calculated using the Floyd-Warshall algorithm [49]. In the next step, the neighbouring links are closed and the solution is sent for hydraulic simulations. Then, for the sectors which are isolated from all sources, the most appropriate link between them and other sectors will be opened (however, it is not clear what is the *most appropriate link*). Finally, a hydraulic simulation is performed on the updated network model, and based on the simulation results, the values for the two objective functions are calculated and the constraint violations are examined, the infeasible solutions are removed, and the non-dominated solutions are chosen.

This approach is one of the approaches that applies multi-objective optimisation in the process of designing a partitioning arrangement. SR4, HR4 (for ADD), HR5, HR8, HR11, ER1 and ER2 are addressed in this method. However, the most important WNS requirements, *i.e.*, sector isolation (SR1) and direct access (SR2), along with SR5, SR6, SR7, SR8, HR3, HR4 (for MHD, PHD, MDD, and FD), HR6, HR7, HR9, and HR10 are not addressed in this method.

2.2.2 Complex networks approaches

In network theory, a complex network is a graph (network) which has non-trivial topological characteristics that do not arise in simple networks, but frequently happen in real networks [83]; *i.e.*, a network 'whose structure is irregular, complex and dynamically evolving in time' [10]. Community detection [50] is an important concept in complex network theory for WDN partitioning. Given a graph $G = (V, E)$, a community (or cluster) is a subgraph $G' = (V', E')$, whose nodes are tightly connected. Since the structural cohesion of the nodes of G' can be quantified in several different ways, there are different formal definitions of community structure. A related definition to WDN partitioning is based on the relative frequency of links, in which communities are seen as collections of nodes within which connections are dense, and between which connections are sparse [10].

Four approaches applied complex networks theory to partition a WDN, namely, Scibetta et al. [105], Diao et al. [38], Campbell et al. [16], and Giustolisi et al [53].

Scibetta et al. (2013)

Scibetta et al. [105] propose a community detection method to identify sectors in a WDN, by finding a trade-off between the maximisation of *modularity* [11] and the reduction of the number of communities. This method uses an iterative algorithm for modularity evaluation that has been proposed by Blondel et al. [8].

Although there is some discussion about the hydraulic aspects of the water distribution

[11]Modularity is a measure of the structure of a network or a graph, and is defined as the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random [84]. Networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules [50].

network in this work, there is no hydraulic requirement involved in the method. It is claimed that the approach successfully identified communities in a case study and the division of the network into sectors is both useful to compute water balances and reduce pressures in the sectors for which the average pressure is excessively high [105]. However, there is no warranty that these results are generalisable as the hydraulic requirements are not considered in the method.

Of the WNS requirements, only SR4 and SR7 are addressed; therefore, the water security requirements (*i.e.*, SR1 and SR2) are not addressed, along with SR5, SR6, SR8, HR3, HR4, HR5, HR6, HR7, HR8, HR9, HR10, HR11, ER1 and ER2.

Diao et al. (2013)

Diao et al. [38] propose a methodology based on the decomposition theorem in complex networks, which can cluster a WDN into a collection of sectors. The method performs a hierarchical decomposition of the WDN into communities, where each community denotes a sector.

The method first maps the network into a weighted undirected graph. Then, the community structure of the network is identified using a procedure suggested by Clauset et al. [18], which uses *modularity* as an indicator to quantify the quality of the graph dissection into communities. An ideal partitioning is a situation in which the number of links within the communities is maximum while the number of links between communities is minimum. Next, the sector size (SR4) restriction is applied on the communities to create sectors. The number of connections in each community is calculated based on water-demand data from the hydraulic model. Then, the feed lines for each community are identified using an iterative selection process based on a sensitivity analysis. The aim of this step is to minimise the number of feed lines for each sector. In each community, the method selects the largest pipes if the pressure constraints (HR4) are satisfied for all nodes. The link that affects the pressure change in the community the most is selected as the feed line. As those links usually are the largest pipes, if using such a link as the feed line is not sufficient to meet the pressure constraints, it may indicate that more than one feed line is required for the community, therefore, more feed line(s) have to be selected. In this regard, the method tries to find the minimal (both in number and in size, *i.e.*, SR6 and SR7) additional feed line(s) that could ensure pressure requirements (HR4) in the community.

This method addresses SR4, SR6, SR7, HR4 (for ADD and FD), and HR10. However, the water security requirements (*i.e.*, SR1 and SR2) are not addressed, along with SR5, SR8, HR3, HR4 (for MHD, PHD, and MDD), HR5, HR6, HR7, HR8, HR9, HR11, ER1 and ER2(SR6 can be considered as a proxy for ER1).

Campbell et al. (2014)

Campbell et al. (2014) [16] propose a method to partition a WDN based on *centrality* and *community detection* concepts in complex networks theory. Centrality is a measure used to assess the importance of a given element in the network to interconnect two or more nodes. Node or link *betweenness* quantifies this measure, and defined as the number of shortest paths from all nodes to all others that pass through that node or link [85].

The approach has the following steps: first, the links on the transmission mains are identified using edge betweenness, flow analysis, and diameter. These links should not be included in any sectors. Then, the *walktrap* algorithm [12] [97] is applied to detect the communities with highest modularity. Finally, energy and hydraulic evaluations are performed to find out the best pipes to be equipped with meters and serve as entry points for each sector. Two indices have been used: *resilience index* (HR8) and *head reduction index*, which resembles *energy efficiency* (HR9) in other approaches [36].

Of the WNS requirements, HR4 (for PHD), and HR9 are addressed in this method. However, the water security requirements (*i.e.*, SR1 and SR2) are not addressed, along with SR4, SR5, SR6, SR7, SR8, HR3, HR4 (for ADD, MHD, MDD, and FD), HR5, HR6, HR7, HR8, HR10, HR11, ER1 and ER2.

[12] The idea behind the walktrap algorithm is that random walks throughout a graph tend to identify high density subgraphs as there are few links leading outside a given community [97].

Giustolisi et al. (2014)

Giustolisi et al. [53] modified and tailored the concept of modularity index (which is a measure of the strength of the network division into communities) considering the specificities of the hydraulic systems and used it as a metric to identify cluster of nodes.

In tailoring the modularity index to WDN, the following considerations were taken into account [53]:

(a) The structure of a WDN is strongly affected by a number of physical constraints (*e.g.*, two-dimensionality, urban structure and planning, demand locations) that make random network inadequate and misleading term of comparison.

(b) The capacity to have parallel pipes, which happens in WDNs.

(c) Considering pipe characteristics to account for the specific technical task of partitioning.

(d) The need for a cut-position-sensitive metric; *i.e.*, the devices segmenting networks (which is equivalent to cuts in the modularity index formulation) are usually installed close to the end nodes of pipes, whereas the classic modularity assumes that they are in the middle of pipes.

(e) Customer demands, and

(f) Background leakages along a pipe.

A genetic algorithm based multi-objective optimisation method is then presented to partition a WDN. The objectives of the optimisation are maximisation of the proposed (modified) modularity index and minimisation of the cost of the required devices to partition the network.

Of the WNS requirements, SR6, HR11 and ER1 are addressed in this work, while water security requirements (*i.e.*, SR1 and SR2) are not addressed, along with SR4, SR5, SR7, SR8, HR3, HR4, HR5, HR6, HR7, HR8, HR9, HR10, and ER2.

2.2.3 Multi-agent systems approaches

A multi-agent system (MAS) can be defined as a loosely coupled network of autonomous problem solvers (also called agents) that interact to solve common problems that are outside of the individual competencies or knowledge of each of them. These agents can be heterogeneous in their nature. The characteristics of multi-agent systems are that (a) each agent has imperfect information or capabilities for solving the problem and, therefore, has a limited and partial perspective; (b) there is no global control; (c) data are decentralised; and (d) communications are asynchronous. Multi-agent systems are ideal for problems that have multiple problem solving methods, multiple perspectives and/or multiple problem solving entities [109].

Fernández et al. [45] applied a multi-agent approach to partition a WDN into a set of sectors. Additionally, in the early stages of the work described in this theses, a multi-agent WDN partitioning approach was proposed [59].

Fernández et al. (2011)

Fernández et al. [45, 65, 66] propose a multi-agent approach to partition a WDN into isolated sectors. In this work, a given number of hydraulic sectors is assumed a priori, which is exactly the same as the number of sources. These sources will be the starting points for creating the corresponding sectors. Autonomous agents start from sources, consider themselves as a cluster, and examine if the adjacent nodes should be added to their cluster. This examination is based on the geographical distance of the node from the source of the cluster, demand of the node, and the elevation of the node.

This work addresses sector isolation (SR1) and direct access (SR2) to water sources for each sector. However, it is limited to small-sized networks or networks with a large number of main sources, as the number of sectors created equals the number of main sources; therefore, if it is applied to a large network, some of the resulting sectors may be larger than the maximum allowable size, especially if the size [13] of the network divided by the

[13] *cf.*, Chapter 3, Section 3.3.1, page 42.

number of sources is larger than maximum allowable sector size. Additionally, no hydraulic requirements are considered in this work. Therefore, of the WNS requirements, SR1, SR2, and HR5 are addressed, while SR4, SR5, SR6, SR7, SR8, HR3, HR4, HR6, HR7, HR8, HR9, HR10, HR11, ER1 and ER2 are not addressed in this method.

Hajebi et al. (2013)

In early stages of the work described in this theses, a method was proposed [59] for partitioning a WDN into a collection of sectors based on graph clustering and multi-agent simulations for sectors' boundary adjustment. The method is composed of two phases: first, the nodes are grouped into a predefined number of components using a k-means clustering algorithm, and then the nodes in the boundaries start negotiation based on their hydraulic characteristics. If a node finds that it is closer to another component in terms of elevation (HR5), it will leave its current component and will join the other one. In the negotiation, minimum cut size (SR6) is considered as a criterion. This method is not deterministic, as the k-means partitioning which is the base of it is not deterministic; so the results of different runs may differ.

Of the WNS requirements, SR4, SR6, and HR5 are addressed in this work, while SR1, SR2, SR5, SR7, SR8, HR3, HR4, HR6, HR7, HR8, HR9, HR10, HR11, ER1 and ER2 are not addressed (SR6 can be considered as a proxy for ER1).

2.3 Discussion

Table 2.2 demonstrates the WNS requirements and compares the existing approaches for WDN partitioning. The columns show the WNS requirements, and the rows show the existing approaches.

It should be noted that connectedness (SR3) of the network must be addressed in all methods. Additionally, conservation of mass (HR1) and conservation of energy (HR2) are physical constraints that govern hydraulic networks. Therefore, these requirements are satisfied in all methods, so they are not reflected in Table 2.2.

	SR1	SR2	SR4	SR5	SR6	SR7	SR8	HR3	HR4					HR5	HR6	HR7	HR8	HR9	HR10	HR11	ER1	ER2
									ADD	MHD	PHD	MDD	FD									
Tzatchkov et al. (2006)	✓	✓																	✓			
Gomes et al. (2012)			✓			✓			✓					✓	✓						✓	
Di Nardo et al. (2013b)					✓						✓						✓	✓			*	
Di Nardo et al. (2013c)	✓	✓															✓	✓				
Ferrari et al. (2014)	✓	✓	✓		✓				✓								✓		✓		*	
Alvisi and Franchini (2014)			✓			✓					✓		✓				✓					
De Paola et al. (2014)			✓						✓					✓			✓			✓	✓	✓
Scibetta et al (2013)			✓			✓											✓					
Diao et al. (2013)			✓		✓	✓			✓				✓						✓		*	
Campbell et al. (2014)											✓							✓				
Giustolisi et al. (2014)					✓															✓	✓	
Fernandez et al. (2011)	✓	✓												✓								
Hajebi et al. (2013)			✓		✓									✓							*	

* Cut size is considered as a proxy for cost.

Table 2.2 Comparison of the state of the art in addressing WNS requirements.

It can be seen in Table 2.2 that:

- Only four approaches address water security requirements, *i.e.*, sector isolation (SR1) and direct access to a water source (SR2): Tzatchkov et al. (2006) [114], Fernández et al. (2011) [45], Di Nardo et al. (2013c) [36], and Ferrari et al. (2014) [47]. In other words, only these approaches deal with water network sectorisation; the other methods deal with WDN partitioning.

- No approach addresses holistic pressure requirements (HR4) considering all scenarios; *i.e.*, ADD, MHD, PHD, MDD, and FD. Additionally, no approach addresses the following requirements: sectors should cross as few mains as possible (SR8), and limited pressure (HR3).

- Limited water velocity (HR6) is considered in just one approach: Gomes et al. (2012) [57].

- Background leakage (HR11) is considered in three approaches: De Paola et al. (2014) [26], Giustolisi et al. [53], and Ferrari et al. (2015) [46].

- Water quality (HR10) is considered as a requirement in three approaches: Tzatchkov et al. (2006) [114], Diao et al. (2013) [38], and Ferrari et al. (2014) [47].

- Elevation (HR5) is considered as a requirement in three approaches: Fernández et al. (2011) [45], Hajebi et al. (2013) [59], and De Paola et al. (2014) [26].

- Energy efficiency (HR9) is addressed in three approaches: Di Nardo et al. (2013b) [37], Di Nardo et al. (2013c) [36], and Campbell et al. (2014) [16].
- Cost is considered as a requirement in three approaches: Gomes et al. (2012) [57], Giustolisi et al. [53], and De Paola et al. (2014) [26]. However, cut-set size can be considered as a proxy for cost. Only De Paola et al. (2014) [26] studied both CAPEX and OPEX.

It should be noted that it is not claimed that the discussed approaches cannot address all the requirements that they do not address. Hydraulic requirements could be addressed in all the discussed approaches, however, if a full hydraulic analysis was performed (*i.e.*, taking all the hydraulic requirements into account), the solutions might have been considered as infeasible. In terms of the structural requirements, only four approaches [36, 45, 47, 114] address water security requirements (*i.e.*, SR1 and SR2). The other approaches cannot address these requirements. It is not impossible to address other structural requirements (*i.e.*, SR4, SR5, SR6, SR7, and SR8) by some modifications in the approaches; however, the current states of the approaches do not address them.

WDN partitioning is a multi-faceted engineering problem. The optimality of partitioning arrangements should be verified in the DMA design process. Only Gomes et al. (2012) [57], Di Nardo et al. (2013c) [36], Ferrari et al. (2014) [47], Alvisi and Franchini (2014) [1], and De Paola et al. (2014) [26] apply optimisation in their approaches. Among them, only Ferrari et al. (2014) [47] and De Paola et al. (2014) [26] apply multi-objective optimisation. For the optimisation algorithm, Gomes et al. (2012) [57] use simulated annealing [78], Di Nardo et al. (2013c) [36] use genetic algorithms [25], Alvisi and Franchini (2014) [1] use a local search, and De Paola et al. [26] use NSGA-II [29]. It is not clear what optimisation method is used in Ferrari et al. (2014) [47].

2.4 Chapter summary

This chapter analysed how the state of the art in water distribution network partitioning methods address the requirements of WDN partitioning and WNS. It first gave an overview of the problem and justified why the current graph partitioning methods are not capable of solving it. Then, the WNS requirements were discussed in detail. These requirements then were used as a framework to analyse the existing approaches. The existing approaches were classified into three categories and each method was analysed with regards to the discussed requirements.

Next chapter (*i.e.*, Problem statement) uses the requirements discussed in this chapter and formulates WNS as a constrained many-objective optimisation problem, which then is solved in Chapter 4.

Chapter 3

Problem statement

This chapter returns to the characteristics of water network sectorisation (WNS) and its corresponding challenges to frame the problem addressed in this thesis. Chapter 2 detailed different aspects of the WNS problem, and proposed a general set of WNS requirements. Based on the discussed requirements, this chapter introduces design objectives and constraints for WNS. Then, it formulates the WNS problem as a constrained many-objective optimisation problem.

3.1 Design objectives

Based on the WNS requirements discussed in Chapter 2 (*cf.*, Section 2.1), there are three sets of design objectives in this problem: structural, hydraulic, and economic objectives. Structural objectives are related to the structure of the network and include minimising sector size imbalance (to address SR5), minimising cut size (to address SR6), minimising cut weight (to address SR7), minimising average customer exposure, minimising maximum customer exposure, minimising average pipe length exposure, and minimising maximum pipe length exposure. The latter four objectives are solely related to water security and quantify SR1, SR2, SR4, and SR5. Hydraulic objectives are related to the hydraulic behaviour of the network after partitioning. These objectives are used as fitness measures to examine different solutions, and include: minimum pressure requirement violations (to address HR4), maximum network resilience (to address HR8), minimum dissipated power (to address HR9), minimum elevation difference within sectors (to address HR5), minimum average of nodal water age during the last 24 hours (to address HR10), and minimum background leakage (to address HR11). Economic objectives deal with the cost of creation and operation of sectors and include minimising CAPEX (ER1) and OPEX (ER2).

There are also three sets of constraints in this problem: structural constraints, hydraulic constraints, and explicit bound constraint. Structural constraints are related to the structure of the network after partitioning, *i.e.*, sector isolation (to address SR1), direct access to a source for each sector (to address SR2), connectedness of the partitioned network (to address SR3), and sector size constraints (to address SR4). Hydraulic constraints govern the hydraulics of the system and include conservation of mass (to address HR1), conservation of energy (to address HR2), pressure constraints (to address HR3), water velocity constraints (to address HR6), and tanks level constraints (to address HR7). Explicit bound constraint limits the domain of the decision variables: links status are binary variables, bounded to 1 (for open links) and 0 (for closed links).

Although some of the constraints are basic hydraulic requirements of a water distribution system, they should be taken into account for WDN partitioning and should be re-examined after the process. In some sources [17, 58, 80] the process of WDN partitioning (or DMA design) is referred to as *network re-design*. As the structure of the network will be modified and some of the hydraulic requirements may not hold after partitioning, it is important to re-examine all the requirements of network design.

3.2 WNS decision variables

Partitioning a water distribution network into a collection of isolated sectors (*i*DMAs) involves [1]:

(a) Assigning all nodes in the network into different sectors, ensuring each sector has direct access to a water source (SR2), and the size of all sectors are within a pre-defined boundary (SR4) and are as balanced as possible (SR5).

(b) Deciding how many (SR6) and which (SR7) links should be closed off to restrict the sectors and how many and which links need to be left open and be equipped with flow meters, in such a way that the sectors are isolated from each other (SR1), while minimising the negative impact on the hydraulic requirements (HR1 to HR11), with minimum cost (ER1 and ER2).

These measures can be achieved by identifying and closing sectors boundary links. In other words, the status of the links [1], *i.e.*, open or closed, determine sector boundaries. It is assumed that no other modifications to the network (*e.g.*, adding nodes, links, and/or other elements) is allowed. Therefore, water network sectorisation (WNS) can be formulated as a constrained multi- (many-) objective optimisation problem with a selection of links status as the only decision variable. The links layout and their connectivity, nodal demand, nodal elevation, and pressure requirements are assumed to be known.

In a general form, the WNS problem can be stated as: finding the best combination of the links status that partition a given WDN into a collection of *i*DMAs, optimising the related objectives while satisfying the related constraints. Constraints are the conditions that must be held in the system to have a feasible solution, and are derived from the WNS requirements. Objective functions reflect the metrics that are also derived from WNS requirements and quantify them. [2]

3.3 Problem formulation

Mathematically, the WNS problem can be formulated as:

Given a graph $G = (nodes, links)$, which is a graph representation of a WDN called *wdn*, where $nodes = \{n_1, n_2, \ldots, n_N\}$, $\forall$ n $\in$ $nodes$, label(n) $\in$ {source, consumer} and $links =$

[1] Here, pipes or valves.

[2] The distinction between constraints and objectives is that a *constraint* is a design target that must be met for the solution to be acceptable; while an *objective* is a design target where more (or less) values for it is better [42]. In practice, a constraint might be translated into an objective in terms of a penalty function for the violations from the constraint, for which less values are better.

$\{l_1, l_2, \ldots, l_m\}$, $\forall\, l \in links$, label(l) $\in$ {pipe, valve, pump} [3], such that $l_j = \langle p, q \rangle$, where $p, q \in nodes$, $1 \leqslant j \leqslant m$. The edge weight is defined as a function ω such that $\omega : Links \rightarrow \mathbb{R}_{>0}$. A *partition* of G is obtained by deleting links from G and is defined here as $\overline{G} = \bigcup_{i=1}^{k} dma_i$ where $dma_i = (nodes_i, links_i)$, is a subgraph [4] of G (a *sector* or an *i*DMA of *wdn*) such that $\forall i \in \{1, ..., k\}$: $nodes_i \subset nodes$, $\quad links_i \subset links$, and also $nodes_i \cap nodes_j = \emptyset$ and $\quad links_i \cap links_j = \emptyset$.

The goal is to find the optimal partition of G considering the following objectives and constraints.

3.3.1 Objectives

The following objectives are defined for this optimisation problem, based on the requirements discussed in Chapter 2 (*cf.*, Section 2.1).

Structural objectives

Structural objectives are related to the structure of the network and include minimising cut size (to address SR6), minimising cut weight (to address SR7), minimising sector size imbalance (to address SR5), minimising average customer exposure, minimising maximum customer exposure, minimising average pipe length exposure, and minimising maximum pipe length exposure (the latter four objectives address SR1, SR2, SR4, and SR5).

Minimum cut size: Minimise the number of links that should be cut and equipped with a valve or a meter, to address SR6; *i.e.*, *minimise*:

$$\text{CS} = |Cuts| \,. \tag{3.1}$$

where,

[3] Although pumps and valves are network elements that are not basically links, they are considered as links for modelling purposes.

[4] A subgraph, H, of a graph, G, is a graph whose vertices are a subset of the vertex set of G, and whose edges are a subset of the edge set of G [39].

- $Cuts = \bigcup_{j=1}^{k} Cut_j$ such that $Cut_j = l_j$ for $l_j \subset links$ with exactly one end-node in dma_j, where $dma_j \in \overline{G}$.

Minimum cut weight: Minimise the sum of diameters of the pipes that must be cut, to address SR7; *i.e.*, *minimise*:

$$\text{CW} = \sum_{l \in Cuts} \text{diameter}(l) \tag{3.2}$$

where,

- $Cuts$ is defined as in formula 3.1, and
- diameter is a function that maps a link to its physical diameter if it is a pipe, otherwise to 0.

It should be noted that the relationship between cost and pipe diameters is generally non-linear [90], but it is considered as a linear relationship in this formulation for simplicity.

Minimum sector size imbalance: The sizes of components (sectors) must be balanced and in a predefined boundary, to address SR5. Therefore, one objective is to *minimise*:

$$\text{SSI} = 1 - \frac{\min_i(\text{size}\,(dma_i))}{\max_j(\text{size}\,(dma_j))} \tag{3.3}$$

where,

- $dma_i, dms_j \in \overline{G}$,
- $\text{size}\,(dma_i)$ is defined as the number of customer connections in dma_i. [5]

Minimum average customer exposure: The average number of people who might be

[5] Alternatively, $\text{size}\,(dma_i)$ could be defined as $\text{size}\,(dma_i) = \max_t \sum_{C \in nodes_i} demand_C^t$ where $demand_C^t$ is the demand of the consumer node C at time $t \in T$ while T is the operational time period.

exposed to a possible contamination should be minimised. Therefore, *minimise*:

$$\text{AExp} = \text{mean}(\text{size}(dma)) \tag{3.4}$$

where $dma \in \overline{G}$, and size is defined as in 3.3. This objective is related to addressing SR1, SR2, SR4, and SR5.

Minimum maximum customer exposure: The maximum number of people who might be exposed to a possible contamination should be minimised. Therefore, *minimise*:

$$\text{MExp} = \max(\text{size}(dma)) \tag{3.5}$$

where $dma \in \overline{G}$, and size is defined as in 3.3. This objective is also related to addressing SR1, SR2, SR4, and SR5.

Minimum average pipe length exposure: The average pipe length in sectors should be minimised to minimise the average effort to decontaminate pipes after a possible contamination. Therefore, *minimise*:

$$\text{ALnkExp} = \underset{dma_i \in dma}{\text{mean}} \Big(\sum_{\forall l \in dma_i} \text{length}(l) \Big) \tag{3.6}$$

where,

- $dma \in \overline{G}$,
- $dma_i \in dma$,
- l is a link in dma_i, and
- length is a function that maps a link to its physical length if it is a pipe, otherwise to 0.

This objective is also related to addressing SR1, SR2, SR4, and SR5.

Minimum maximum pipe length exposure: The maximum pipe length in sectors should be minimised to minimise the maximum effort to decontaminate pipes after a possible contamination. Therefore, *minimise*:

$$\text{MLnkExp} = \max_{dma_i \in dma} \Big(\sum_{\forall l \in dma_i} \text{length}(l) \Big) \tag{3.7}$$

where,

- $dma \in \overline{G}$,
- $dma_i \in dma$,
- l is a link in dma_i, and
- length is a function that maps a link to its physical length if it is a pipe, otherwise to 0.

This objective is also related to addressing SR1, SR2, SR4, and SR5.

Hydraulic objectives

Hydraulic objectives quantify how a partitioned network behaves from hydraulic point of view.

Minimum pressure requirements violations: Provisioning of minimum system pressure requirements (to address HR4) is an important objective in WDN partitioning. A certain amount of water must be supplied at the required pressure for each node during different consumption scenarios (*i.e.*, ADD, MHD, PHD, MDD, and FD) (*cf.*, Section 2.1). This objective is defined as, *minimising*:

$$\text{PV} = \sum_{s \,\in\, scen} \sum_{j=1}^{N} C_{s,j}^{pen} \times max\left(0, H_{s,j}^{min} - H_{s,j}\right) \tag{3.8}$$

where,

- $s \in scen$ = {ADD, MHD, PHD, MDD, FD},
- N is the number of $nodes$,

- $C_{s,j}^{pen}$ is the penalty cost of pressure deficit for node j in scenario s [6],
- $H_{s,j}^{min}$ is the minimum admissible pressure at node j in scenario s (which is assumed to be given), and
- $H_{s,j}$ is the pressure for node j in scenario s obtained from simulation results.

Minimum dissipated power: *Dissipated power*, is the power that is dissipated in the pipes during the time water flows from the sources to the users; *i.e.*, the total available power at the entrance of the distribution network minus the power that is delivered to the users [113]. It is a good measure for network energy efficiency (which addresses HR9). This objective can be formulated as [36], *minimise*:

$$\mathrm{DP} = \gamma \sum_{i=1}^{m_s} Q_i \Delta H_i. \tag{3.9}$$

where,

- γ is the specific weight of water,
- $m_s = m - y$, where y is the number of boundary links, and m is the total number of links in the network,
- Q_i = flow rate, and
- ΔH_i= head loss for each link i in network.

Minimum elevation difference within sectors: Variation in nodal elevation (ground level) is very important in water distribution networks. Nodes with high differences in their elevations should be in different sectors. If there is a high elevation difference in a sector, it must be compensated by the use of pressure reducing valves (in case of high pressure) or by using pumps (in case of low pressure), which is not desirable because of their installations and operations costs [79]. Therefore, the following objective function is

[6]Penalty cost is used to give different weights for pressure violations to different nodes; *e.g.*, pressure violations for an important node like a hospital will be treated differently than the pressure violations of a park, by assigning different penalty costs to them.

defined as, *minimise*:

$$\mathrm{ED} = \sigma_{el} = \sum_{d=1}^{k} \sigma_d. \tag{3.10}$$

where,

- $\sigma_d = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \left(z_j - \bar{z}\right)^2}, \forall\, dma_d \in \overline{G},$
- n is the number of nodes in dma_d,
- z_i is the elevation of node $i \in dma_d$, and
- $\bar{z}$ is the mean of elevation of all nodes in dma_d.

This objective addresses HR5.

Maximum network resilience: Resilience index, which is a measure for network reliability, measures the amount of energy redundancy in a water network. *Resilience* is defined as the inherent capacity of a water network to manage unexpected failures, and it is measured as 'the ratio between the surplus of power delivered to users and the maximum power that can be dissipated in the network when meeting exactly the design criteria' [58]. It is formulated as [113], *maximise*:

$$\mathrm{ANR} = \underset{T}{\mathrm{mean}}\left(\frac{\sum_{i=1}^{N} D_i \left(H_i - H_i^*\right)}{\sum_{r=1}^{n_r} Q_r H_r + \sum_{j=1}^{n_p} P_j/\gamma - \sum_{i=1}^{N} D_i H_i^*}\right) \tag{3.11}$$

where,

- T is the operational time period (the mean is calculated over all operational time periods),
- N is the number of *nodes* in the network,
- D_i is the demand at node i in (m^3/sec),
- H_i and H_i^* are available and minimum required head at node i in (kPa), respectively,
- n_r is the number of reservoirs in the network,
- Q_r and H_r are the discharge in (m^3/sec) and head in (kPa) at reservoir r, respectively,

- n_p is the number of pumps in the system,
- P_j is the is the power introduced into the network by the jth pump in (kW), and
- γ is the specific weight of water in (N/m^3).

This objective addresses HR8. There exist other measures for network reliability, *e.g.*, the ones proposed in [99] and [68]; however, the above formula is used most frequently in the literature [6].

Minimum water age: Water age is usually considered as a reliable surrogate measure for water quality [119]. The average water age in the network in the last 24 hours should be minimised. This objective is defined as, *i.e.*, *minimise*:

$$\mathrm{WA} = \frac{\sum_{i=1}^{N} \sum_{t=LT-24}^{LT} wa_i^t}{24} \tag{3.12}$$

where,

- i is the subscript of the ith node in the network,
- LT is the last time step in the operational time period, and
- wa_i^t is the water age at node i at time t in (hour).

This objective addresses HR10.

Minimum background leakage: Background leakage is characterised by small non-visible leaks that occur mostly at joints and fittings, not generating sufficient noise to be detected by existing equipment [30], which should be minimised [17]; *i.e.*, *minimise*:

$$\mathrm{SL} = \sum_{l=1}^{m} d_l^{leaks}(P_{l,mean}) \tag{3.13}$$

where,

- $d_l^{leaks}(P_{l,mean}) = \beta_l L_l P_{l,mean}^{\alpha_l}$ if $P_{l,mean} > 0$, and it is zero otherwise,

- l is the subscript of the lth pipe; $P_{l,mean}$ is the model mean pressure along the lth pipe in (m),
- d_l^{leaks} is the background leakages outflow along the lth pipe in (m^3/sec),
- α_l and β_l are the model parameters, where α is unit-less and β is in ($m^{2-\alpha}$/sec), and
- L_l is the length of the lth pipe in (m) [52].

This objective addresses HR11.

3.3.2 Constraints

There are three types of constraints that should be satisfied in this problem: structural, hydraulic, and explicit bound constraints.

Structural constraints

There are four structural constraints in this problem: sector isolation (to address SR1), direct access to a source for each sector (to address SR2), connectedness of the partitioned network (to address SR3), and sector size limitations (to address SR4).

Sector isolation: This constraint enforces that no flow exchange between different sectors is allowed.

$$\text{Given } u \in nodes_i, v \in nodes_j, \nexists \text{ a } u - v \text{ flow path in the PN} \tag{3.14}$$

where,

- PN is the partitioned network, and
- a $u - v$ flow path in the PN is a directed path between any nodes $u \in nodes_i$ and $v \in nodes_j$ in the PN, considering the flow directions in the network.

This constraint addresses SR1.

Direct access to a source for each sector: This constraint enforces that each identified sector must have a *direct access* to at least one water source; *i.e.*, the path from the sector to at least one source must not contain any nodes in other sectors. [7]

$$\begin{aligned} &\forall\, dma_i \in \overline{G},\ \forall\, u \in nodes_i\ \exists\, s \in sources \\ &\text{such that there is a } s-u \text{ path in the PN,} \\ &\text{where } \forall\, v \text{ in the } s-u \text{ path, } \nexists\, dma_j \text{ such that } v \in dma_j. \end{aligned} \tag{3.15}$$

where,

- PN is the partitioned network.
- A $s-u$ path is defined similar to the definition in 3.14.

This constraint addresses SR2.

Connectedness of the partitioned network: This constraint warrants that each *i*DMA should be a connected subgraph and the whole network after partitioning should also be connected to ensure that all the nodes will receive water.

$$\forall\, u,\ v \in nodes: \ \text{ there is a } \ u-v \ \text{ path in the PN.} \tag{3.16}$$

where,

- PN is the partitioned network.

This constraint addresses SR3.

Sector size constraint: This constraint enforces that the size of each identified sector should be within a predefined boundary [79].

$$\forall\, dma \in \overline{G} : DmaMinSize < \text{size}\,(dma) < DmaMaxSize. \tag{3.17}$$

[7] In this thesis, *direct access* to a water source for each sector means that the path from a node or a link in the sector to a source does not contain any nodes in other sectors. Therefore, if the path contains nodes which are not assigned to any sector, it is still considered as a direct access.

where,

- *DmaMinSize* and *DmaMaxSize* are pre-defined minimum and maximum sector sizes [8], and
- size is defined as in 3.3.

This constraint addresses SR4.

Hydraulic constraints

Hydraulic constraints govern the hydraulics of the system and include conservation of mass (to address HR1), conservation of energy (to address HR2), pressure constraints (to address HR3), water velocity limitations (to address HR6), and tanks level limitations (to address HR7).

Conservation of mass: Conservation of mass (is also regarded as continuity of flow) dictates that the fluid mass entering any pipe will be equal to the mass leaving the pipe (since fluid is neither created nor destroyed in hydraulic systems). This constraint yields a set of linear algebraic equations in terms of flows [119]. For each node at each time t, flow continuity should be satisfied, *i.e.*, [119]:

$$-\sum Q_{in}^{t} + \sum Q_{out}^{t} + demand_{C}^{t} = 0. \tag{3.18}$$

where,

- $demand_{C}^{t}$ is the demand at node C at time t,
- Q_{in}^{t} and Q_{out}^{t} are the flows entering and leaving node C at time t, respectively.

This constraint addresses HR1. [9]

[8] As discussed in Chapter 2, boundaries like (300-2500), (500-3000), and (500-5000) customer connections can be found in the literature [32, 45, 79]; however, (500-5000) is the most common one.

[9] In modelling, this constraint is always satisfied by a network solver (*e.g.*, EPANET [103].)

Conservation of energy: Conservation of energy (is also regarded as hydraulic equilibrium equations) imposes that the difference in energy between two points in a network must be the same, irrespective of flow path. For hydraulic analysis, this principle can be represented in terms of *hydraulic head* [10] [119]:

$$z_1 + \frac{P_1}{\gamma} + \frac{V_1^2}{2g} + \sum h_P = z_2 + \frac{P_2}{\gamma} + \frac{V_2^2}{2g} + \sum h_L + \sum h_M. \tag{3.19}$$

where,

- Z_1 and Z_2 are elevations at points 1 and 2, respectively, in (m),
- P_1 and P_2 are pressures at points 1 and 2, respectively, in (N/m^2),
- γ = fluid's (water) specific weight, in (N/m^3),
- V_1 and V_2 are flow velocities at points 1 and 2, respectively, in (m/s),
- g= acceleration due to gravity, in (m/sec^2)
- h_P= pumping head gain, in (m),
- h_L= head loss in pipes, in (m), and
- h_M= head loss due to minor losses, in (m).

This constraint addresses HR2. [9]

Pressure constraints: The partitioned network should guarantee the minimum and maximum admissible pressure for each node in any consumption scenario at all times; *i.e.*, :

$$P_s^{min} \leqslant P_{i,s}^t \leqslant P_s^{max} \quad \forall i \in nodes, t \in T. \tag{3.20}$$

where,

[10]The total energy associated with a fluid per unit weight of the fluid is called *head* (H). The kinetic energy is called velocity head ($\frac{V^2}{2g}$), the potential energy is called elevation head (Z), and the internal pressure energy is called pressure head ($\frac{P}{\gamma}$). While typical units for energy are foot-pounds (Joules), the units of total head are feet (meters); *i.e.*, $H = Z + \frac{P}{\gamma} + \frac{V^2}{2g}$, where H = total head, in (m); γ = fluid's (water) specific weight, in (N/m^3); V is flow velocity, in (m/s); and g= acceleration due to gravity, in (m/sec^2) [119].

- P_s^{min} and P_s^{max} are the maximum and minimum pressure requirements, respectively, for each scenario s,
- $P_{i,s}^t$ is the service pressure in node i at time t in scenario s,
- *nodes* is the set of all nodes, and
- T is the operational time period (usually 24 hours, with time step of 1 hour).

This constraint addresses HR3.

Water velocity constraint: The minimum and maximum velocity of water in pipes in the partitioned network should be within a certain limit [56]; *i.e.*, :

$$V_{min} \leqslant V_{p,s}^t \leqslant V_{max} \quad \forall p \in links, t \in T. \tag{3.21}$$

where,

- $V_{p,s}^t$ is the water velocity in pipe p at time t in scenario s in (m/s), and
- V_{min} and V_{max} are the minimum and maximum admissible water velocity in (m/s); typically, 1 and 3 (m/s), consecutively [119].

This constraint addresses HR6; however, it is not a hard constraint; *i.e.*, minor violations could be accepted but not preferred.

Tanks level constraints: Tanks are special types of source nodes. They neither produce water nor consume it; but they can store water and when it is needed, they can feed the network. The level of water in tanks should be within a boundary. The maximum allowable water level is typically the top of the tank, and the minimum allowable water level is typically above the bottom of the tank to provide some residual storage for potential fire defeat events. For each operational time interval t and tank i the constraints can be stated as [86]:

$$L_i^{min} \leqslant L_{i,s}^t \leqslant L_i^{max} \quad \forall i \in Tanks, t \in T. \tag{3.22}$$

where $L_{i,s}^t$ is the level of tank i at time t in scenario s (m); and L_i^{min} and L_i^{max} are minimum and maximum admissible levels for tank i, respectively. Additionally, the difference between the water level at the start and the end of simulation period should be limited (usually less than 10 percent of the tank height); *i.e.*, :

$$L_{i,s}^{end} - L_{i,s}^{start} \leqslant MLD_i \quad \forall\, i \in Tanks,\, t \in T. \tag{3.23}$$

where,

- $L_{i,s}^{start}$ and $L_{i,s}^{end}$ are the levels of the tank i at the beginning and the end of the simulation in the scenario s in (m), and
- MLD_i is the maximum level difference for tank i in (m).

This constraint addresses HR7. While 3.22 is a hard constraint (*i.e.*, no violations is accepted), 3.23 is not a hard constraint (*i.e.*, minor violations could be accepted but not preferred).

Explicit bound constraints

As discussed earlier in this chapter, status of links are the only decision variables in this problem, which are either 1 (open) or 0 (closed); *i.e.*, :

$$\begin{gathered} X = (x_1, x_2, \ldots, x_m) \\ x_i \in \{0,1\} \text{ is the status of link } i. \end{gathered} \tag{3.24}$$

where,

- X is a vector including the decision variables.

3.4 Summary

This chapter provided details about water network sectorisation (WNS) problem by introducing metrics that quantify WNS requirements discussed in Chapter 2. It also provided a formulation of the WNS problem as a constrained many-objective optimisation problem. The next chapter will discuss how the explained problem is solved in this thesis.

Chapter 4

Design and implementation

Water network sectorisation (WNS) is a multifaceted problem embracing structural, hydraulic, and economic aspects (*cf.*, Chapter 3). The structural aspects of the problem may be handled using different graph theory algorithms for graph partitioning, clustering, and community detection among others. However, as the goal is to partition a WDN into isolated sectors (to address SR1) all having direct access to a water source (to address SR2), the existing graph-theory techniques are not sufficient (*cf.*, Section 2.1.1); therefore, there is a need to develop an algorithm that takes the structural requirements of the problem into account. Additionally, besides the structural requirements, WNS has hydraulic and economic requirements (*cf.*, Chapter 3), therefore, it is a multi-objective (many-objective) optimisation problem.

This chapter first discusses design decisions related to different aspects of the WNS problem to motivate the choice for the novel method proposed in this thesis (*cf.*, Section 4.1). Then, it discusses the challenge that the objectives and constraints of the problem are not explicitly related to the decision variables, and how this thesis addresses this challenge (*cf.*, Section 4.1.1). Finally, the section thereafter (*cf.*, Section 4.2) explains how the proposed solution (*i.e.*, WDN-PARTITION) integrates the design decisions and addresses the requirements and challenges of the problem.

4.1 Design decisions

Optimisation is 'the act of obtaining the best result under given circumstances' [101]; in other words, it is 'the process of determining the best design' [91]. There are four different approaches to deal with multi-objective optimisation problems:

1) **Mathematical approaches**, which consist of maximising or minimising some real functions by systematically choosing input values from the domain of decision variables and calculating the value of the functions [108]. However, in the WNS problem, objectives and constraints are not explicit functions of the decision variables (*cf.*, Chapter 3). Therefore, it is impossible to use mathematical methods directly to solve the WNS problem.

2) **Brute force search**, also known as *exhaustive search* and/or *generate and test*, is a very general problem solving technique that consists of systematically enumerating all possible candidates for the solution and checking whether each candidate solution satisfies the problem statement [89]. The issue with this method is the very large search space of the WNS problem, which is 2^n, for n being the number of links in the network. As an example, a network used as a case-study in Chapter 5, has 14,831 links. It is practically infeasible to test 2^{14831} possible solutions, assess their structural feasibility, and examine their hydraulic viability by simulation [1]. Another problem with this strategy is *scalability*: if the number of links in a network increases by 1, the size of the search space will double.

3) **Meta-heuristic techniques**, which are problem-independent search techniques [9]. As such, they do not take advantage of the specificities of the problem, and therefore, can be used as black boxes. One possible classification of meta-heuristic techniques consider whether the approach uses a single solution or it involves a population-based search [9, 110].

[1]To have an intuition of how large the number is one may consider that 2^{8192} has 2467 digits and 2^{16384} has 4933 digits.

(a) *Single solution approaches* focus on modifying and improving a single candidate solution. Simulated annealing, Tabu search, iterated local search, guided local search, and variable neighbourhood search are examples of this category.

(b) *Population-based approaches* maintain and improve multiple candidate solutions, and often use population characteristics to guide the search. Evolutionary computation and genetic algorithms are examples of this category.

(c) *Swarm intelligence approaches* take advantage of the collective behaviour of decentralised, self-organised agents in a population or swarm. Multi-agent systems, ant colony optimisation, particle swarm optimisation, social cognitive optimisation, and artificial bee colony algorithms are examples of this category.

All the three approaches of meta-heuristic techniques have been examined in different stages of this study (*i.e.*, Tabu search [60], genetic algorithms (NSGA-II) [63], and multi-agent systems [59]). However, the solutions generated by these techniques usually do not satisfy the structural requirements (specifically, SR1 and SR2), as they are typically general search techniques and do not consider the specificities of the problem.

4) **Heuristic techniques**, which are problem-specific, adapted to the problem at hand, and try to take full advantage of the particularities of the problem [92]. This approach has been used in this thesis.

The method proposed in this thesis, hereafter called WDN-PARTITION, is a heuristic technique that satisfies water security related requirements and generates a collection of structurally feasible solutions (*i.e.*, solutions that address structural requirements of WNS, namely SR1, SR2, SR3, SR4, and SR8, using a novel structural graph partitioning method. Then, the near-optimal solutions are identified using a many-objective optimisation method considering the corresponding constraints and objectives (*cf.*, Section 3.3).

4.1.1 Design challenge

After applying the structural graph partitioning method, the hydraulic and economic objectives and constraints should be considered. The challenge is that the objective functions are not explicitly defined based on the decision variables. As discussed in Section 3.2, the decision variables are the links status with values either 0 or 1. The parameters of the objective functions and constraints (*cf.*, Section 3.3) are affected by the changes in the decision variables, however, their relationships are not trivial to formulate (*i.e.*, opening or closing links in the network affects nodal pressure, link flows, etc., but the exact effect on, for example, each node's pressure or each link's flow cannot be easily formulated). A disaggregated methodology is used to solve this challenge [61]. Although the decision variables do not appear in the objective functions and constraints, the results of changing them can be obtained indirectly using a network solver, *i.e.*, a water distribution network hydraulic simulator, *e.g.*, EPANET [103].

Each solution will be sent from the graph partitioning algorithm to the network solver to satisfy the hydraulic constraints. Once the network solver has solved the hydraulic equations and satisfied the hydraulic constraints, it will give the values for the parameters used in the hydraulic objectives and constraints (flow, head, pressure, etc.). Then, the values of the objectives and constraints can be calculated and checked in the optimisation procedure.

Another issue is the optimisation method. WNS is a many-objective [2] optimisation problem, so it is usually not possible to have a single solution as the best one, as some objectives are conflicting. The candidate solutions should be compared against each other in terms of different objectives, *i.e.*, formulae 3.1 through 3.13. There are some multi-objective optimisation algorithms that perform well in dealing with two or three objectives [19, 21, 29, 55, 70, 126]; however, they cannot adequately handle more than three objectives [67]. The number of objective functions in WNS is 13 (*cf.*, Section 3.3), so these methods are not suitable to handle this problem. Recently some algorithms have been proposed to deal with many (more than three) objectives [67, 75, 124]. However, because of the structural

[2] More than three objectives [48].

constraints of WNS, these methods cannot generate feasible solutions and/or modify the solutions in such a way that the new solutions are also feasible. This thesis exploits the *domination* [3] concept [29] to compare different solutions with more than three objectives without compromising the search procedure, as a collection of feasible solutions is identified beforehand.

To demonstrate the different steps of the method, the WDN for the city of Novato, California is used. This WDN is the most complex example included in EPANET 2 [103] which covers an area of about 150 km^2. It is a dual-source network, composed of 92 junctions, 2 reservoirs, 3 tanks, which are interconnected thorough 117 pipes and 2 pumps. This network is referred to as the Novato network, hereafter. Figure 4.1 illustrates this network as it can be seen in EPANET.

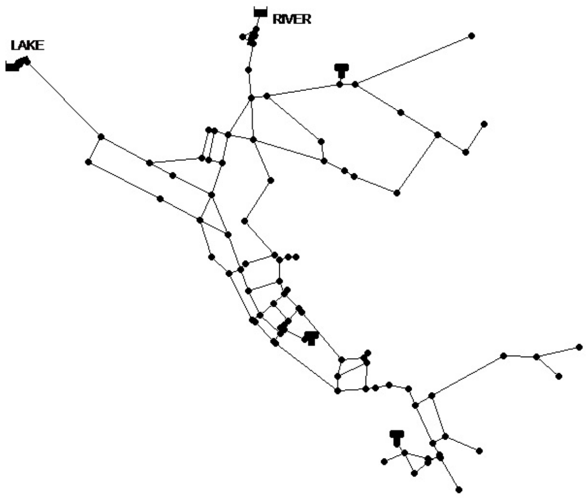


Fig. 4.1 The Novato network as it can be seen in EPANET [103].

[3] In multi-objective optimisation, solution A is said to *dominate* solution B if and only if A is no worse than B in terms of all the objective functions, and A is exactly better than B in terms of at least one objective function. Typically, in the optimisation procedure all solutions are compared against each other and if one solution dominates another one, the dominated solution is removed from the candidate solutions set. Therefore, all the non-dominated solutions will remain, that create the *Pareto front* [20, 29].

4.2 Proposed solution

Based on valuable insights from existing solutions [47, 104] (*cf.*, Section 2.2.1), this thesis proposes a novel heuristic WDN partitioning technique, called WDN-PARTITION, to partition a water distribution network into isolated sectors; to tackle the water network sectorisation (WNS) problem. As discussed, WDN-PARTITION first satisfies water security related constraints and generates a collection of feasible solutions; *i.e.*, solutions that address structural requirements of WNS problem, namely sector isolation (SR1), direct access to at least one source for each sector (SR2), connectedness of the partitioned network (SR3), sector size limitations (SR4), and the requirement that sectors should cross as few mains as possible (SR8). Then, it identifies the near-optimal solutions using a many-objective optimisation method considering the constraints and values for the objective functions (*i.e.*, the metrics for evaluation).

In a nutshell, WDN-PARTITION first identifies pipes that transport water from sources to major areas of the network (*major flow paths*), then identifies the groups of nodes that are connected to the major flow paths but are isolated from the rest of the network (*islands*), then checks the size of the identified islands, and partitions them if needed (*i.e.*, their size is larger than a pre-defined boundary). Finally the best solutions are selected using a many-objective optimisation procedure. Figure 4.2 illustrates a flowchart of WDN-PARTITION, which consists of the following steps.

Algorithm 1 shows the overall procedure of this step.

Step 1: Initialisation. First of all, the network data is read from an EPANET [103] input file, and a model (*Model*, which is a structure that contains all items and parameter of the network) of the network is created, and some parameters are being set. An adjacency matrix $Model.A$ of the network is created using the network data. The $Model.A(i, j)$ and $Model.A(j,i)$ entries in the matrix are set to the diameter of a link l if there is such a link between two nodes i and j in the network, infinity if there is a pump or valve between i and j, and 0 if there is no link between i and j. Some other information are extracted out of the input file and set in the model, such as reservoir nodes which are set to $Model.Sources$. Additionally, flow directions are identified as a result of an extended simulation, and set for each link. If flow path is always one way, then the link is considered as a directed link; otherwise as an undirected link (*i.e.*, if flow direction is always from i to j, then $i-j$ link is

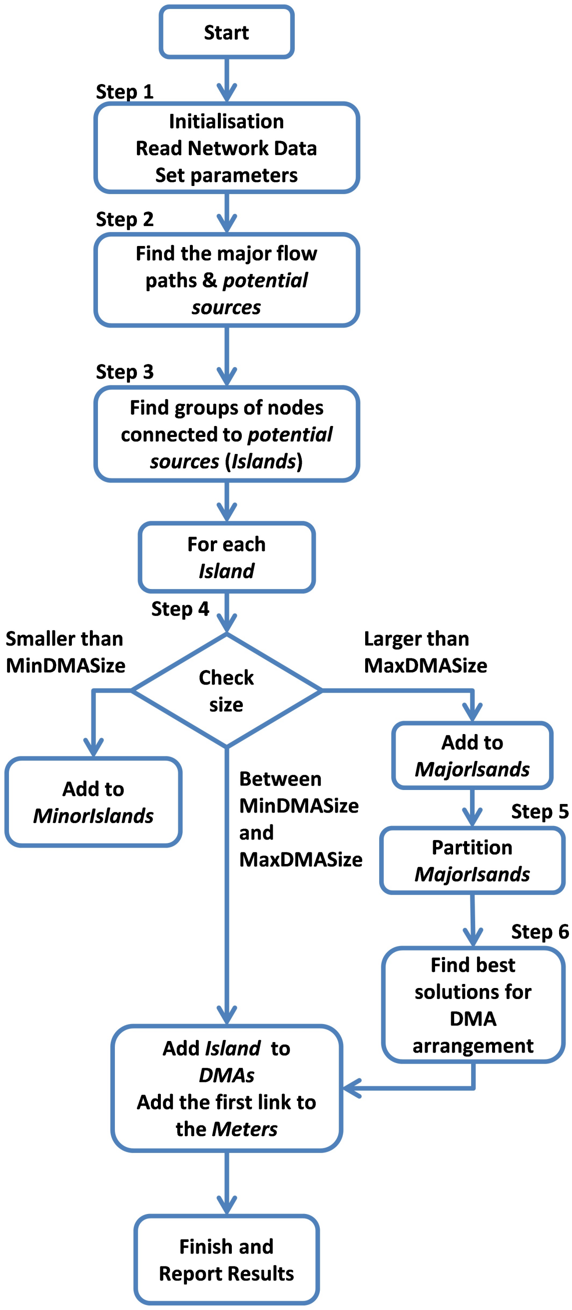


Fig. 4.2 Flowchart illustrating different steps of WDN-PARTITION.

Algorithm 1: WDN-PARTITION Partitions a WDN into isolated sectors.

Input:
- *inp*: network data

Output:
- *DMAs*: a list of isolated sectors and their nodes
- *Valves*: a list of links that should be closed
- *Meters*: a list of links that should be equipped with meters

```
Model ← CreateModel(inp) ;                                  // Initialization
Model.ptnSrcs ← findAllPotSrcs(Model) ;                     // Algorithm 2
Model.Islands ← IdentifyIslandss(Model) ;                   // Algorithm 3
Model ← CheckIslandsSize(Model) ;                           // Algorithm 5
Model ← PartitionMajorIslands(Model) ;                      // Algorithm 6
Model ← FindBestSols(Model) ;           // Optimisation (Algorithm 11)
return Model.DMAs, Model.Valves, Model.Meters;
```

considered as a directed link from i to j; otherwise it is considered as an undirected link).

The following parameters are asked from the user during the initialization phase:

- *Model.MainsSizeThreshold* which is the minimum threshold size for the main pipes,
- *Model.DmaMinSize* and *Model.DmaMaxSize* which are the minimum and maximum size [4] of a sector,
- *Model.MaxIter*, which is the maximum number of iterations in the loop in Step 5.1.2,
- A list of required criteria among the objective functions, which are considered as being important for the specific network at hand, and their priorities (if more than one are chosen).

Then, two parameters are set based on the network data:

- *Model.LinkCount* which is the number of links in the network, and
- A solution template (called *Model.position*), which is a bit vector of length *Model.LinkCount* and is used as a template for possible solutions. Each bit in the position vector uniquely represents a link in the network; 0 represents a closed link

[4] Size is defined in formula 3.3, *cf.*, Section 3.3, page 42.

and 1 represents an open link. The $Model.position$ vector is initialised by 1 in all places.

Finally, two sets are defined during initialization: $Model.Meters$, which is the set of links that should be equipped with meters (the links that connect the sectors to large water mains), and $Model.Valves$, which is the set of links that should be closed off (*i.e.*, the neighbouring links in the boundaries of different sectors identified after partitioning).

Step 2: Finding major flow paths and potential sources. Once the network data is read and the corresponding adjacency matrix is created, it is possible to explore the network. First, WDN-PARTITION finds the large water mains (*major flow paths*) from the main sources (reservoirs) and considers nodes in these paths as new *potential sources*. It starts from the main sources in the network and using a modified depth-first search (DFS) [112] algorithm, it adds nodes to a list called $Model.ptnSrcs$. In this process, the nodes are identified in the direction of flow, while their connecting links are larger than a specified threshold ($Model.MainsSizeThreshold$). The sectors then can be identified starting from these *potential sources* to guarantee direct access to a source (SR2), as these nodes are directly connected to a source, and will not be assigned to any sector. Figure 4.3 shows the result of applying Step 2 on the Novato network. Algorithm 2 shows the overall procedure of this step.

Algorithm 2: FINDALLPOTSRCS Finds all potential sources in a graph.

Input:
- *Model*: a data structure containing all the information about the network

Output:
- *ptnSrcs*: a list of potential sources

```
foreach source ∈ Model.Sources do
    Run a DFS on Model.A and add all the nodes with links larger than
    Model.MainsSizeThreshold to ptnSrcs;
return ptnSrcs;
```

Step 3: Identifying Islands. In this step, the groups of nodes which are connected to each node in $Model.ptnSrcs$ are identified and added to a list called $Model.Islands$. To this end,

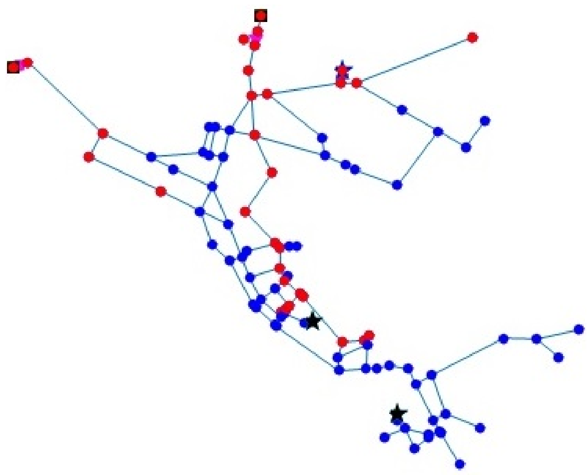


Fig. 4.3 Identifying large water mains and the nodes on them (*Model.ptnSrcs*) in the Novato network. The identified nodes are highlighted. These nodes are used in the later steps to identify *islands*; *i.e.*, groups of isolated nodes that are connected to a source while their path to the source does not contain nodes in other groups.

each node in *Model.ptnSrcs* is selected and a breadth-first search (BFS) [96] algorithm is used to find all the nodes that are reachable from it. Algorithm 3 shows the overall procedure of this step.

For the first-level neighbours (immediate neighbours), the left-hand side nodes are separated from the right-hand side nodes (or the upper side nodes are separated from the lower side nodes), considering the large water mains as a reference line (lines 2 to 7 in Algorithm 3). This separation guarantees that the identified groups of nodes (*Islands*) and so the identified sectors in the next steps do not cross the large water mains (to address the SR8 requirement). This is one of the advantages of WDN-PARTITION to the other approaches [36, 45, 47, 104]. Figure 4.4 shows the result of applying this step on the Novato network.

After identifying the groups of nodes directly connected to large water mains (*Islands*), the method may end up with some equal or overlapping islands. This may happen for the islands that are connected to the large water mains by more than one link. Redundant or overlapping islands are then merged. Algorithm 4 shows the overall procedure of this step.

Step 4: Checking the size of islands. In this step, the sector size constraint is examined

Algorithm 3: IDENTIFYISLANDS Finds all potential independent groups of nodes (*Islands*) in the network.

Input:
- *Model*: a data structure containing all the information about the network

Output:
- *Islands*: A list of nodes belonging to independent groups

foreach $ptnSrc \in Model.ptnSrcs$ **do**
  $firstNeighbours \leftarrow$ Find the first neighbours of *ptnSrc* which are not in *Model.ptnSrcs*;
  $SortedPotSrcs \leftarrow$ Sort *Model.ptnSrcs* based on their distance (in terms of number of links) from the reservoir which is connected to *ptnSrc*;
  $A \leftarrow$ the node exactly before *ptnSrc* in *SortedPotSrcs*;
  $B \leftarrow$ the node exactly after *ptnSrc* in *SortedPotSrcs* (if there is no node after *ptnSrc*, then $B \leftarrow ptnSrc$);
  foreach *C in firstNeighbors* **do**
    /* *x* and *y* are coordinates of nodes */
    if $(B.x - A.x) * (C.y - A.y) - (B.y - A.y) * (C.x - A.x) > 0$ **then**
      /* *C* is in the left-hand (upper) side of *ptnSrc* */
      Start from *C*, run a BFS on *Model.A* in the direction of flow and add all the identified nodes to *Island*;
      Add *Island* to *Islands*;
    else
      /* *C* is in the right-hand (lower) side of *ptnSrc* */
      Start from *C*, run a BFS on *Model.A* in the direction of flow and add all the identified nodes to *Island*;
      Add *Island* to *Islands*;
$Islands \leftarrow$ MergeOverlapingIslands(*Islands*);
return *Islands*;

Algorithm 4: MERGEOVERLAPINGISLANDS Merges overlapping islands.

Input: *Model.Islands*: A list of all islands
Output: *Model.Islands*: A list of islands, with redundant or overlapping ones merged

for $island_i \in Model.Islands$ **do**
  for $island_j \in Model.Islands$ **do**
    if $island_i \cap island_j \neq \varnothing$ **then**
      Add all nodes in $island_j$ to $island_i$;
      Remove $island_j$ from *Model.Islands*;
return *Model.Islands*;

to address the SR4 requirement. For each *island* in the *Model.Islands* list, the size of the island is assessed, based on the definition of size in formula 3.3 (*cf.*, Section 3.3). There are three possible situations:

- If the *island* size is in the range of the allowed size of a sector (between *Model.DmaMinSize* and *Model.DmaMaxSize*), the *island* will be considered as a sector, and the first link from the corresponding potential source(s) will be added to the *Model.Meters* set.
- If the *island* size is less than the minimum allowable size of a sector (*Model.DmaMinSize*), it is considered as a *MinorIsland* and is added to *Model.MinorIslands*. No meter is defined for *MinorIslands* as the cost is not justified [47, 104]. However, these islands cannot cause any water security threat for the rest of the network as there is no flow path from these islands to the other parts of the network.
- If the *island* size is greater than the maximum allowable size of a sector (*Model.DmaMaxSize*), it is considered as a *MajorIsland*, is added to *Model.MajorIslands*, and must be partitioned into a set of isolated sectors all having direct access to the large water mains, and consequently, to a water source. *Model.MajorIslands* will be partitioned in Step 5.

Algorithm 5 shows the overall procedure of this step.

It can be seen in Figure 4.5 that two *MinorIslands* and one *MajorIsland* are identified in the Novato network. No proper size island (*i.e.*, *i*DMA) is identified in this network in this step.

Step 5: Partitioning *MajorIslands*. In this step, which is the central part of WDN-PARTITION, *MajorIslands* are partitioned into smaller sub-networks of proper sizes, each of them with direct access to at least one water source. For each *MajorIsland* in *Model.MajorIslands*, partitioning is done by growing subgraphs, starting from the nodes in the *MajorIsland* that are directly connected to the large water mains (*i.e.*, nodes

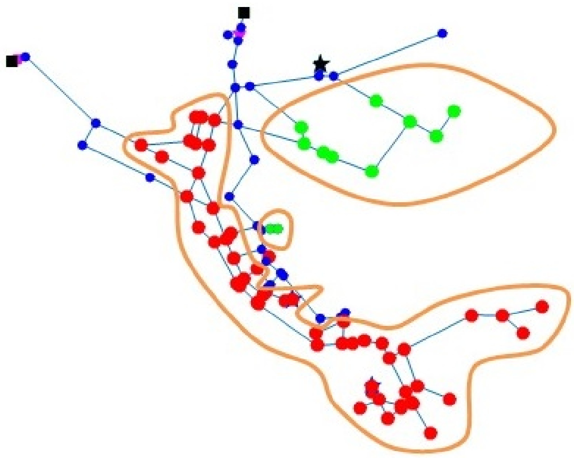


Fig. 4.4 Identifying *Model.Islands* in the Novato network. Nodes that are highlighted are *Islands*.

Algorithm 5: CHECKISLANDSSIZE Checks the sizes of *Model.Islands* and identifies whether they should be partitioned or not.

Input:
- *Model*: a data structure containing all the information about the network

Output:
- *Model*: updated *Model* in which *MinorIslands*, *InitialDMAs*, and *MajorIsands* are set

```
foreach Island ∈ Model.Islands do
    if Model.DmaMinSize ≤ size(Island) ≤ Model.DmaMaxSize then
        Add Island to Model.InitialDMAs;
    if size(Island) < Model.DmaMinSize then
        Add Island to Model.MinorIslands;
    else
        Add Island to Model.MajorIslands;
return Model;
```

---

in both $MajorIsland$ and $Model.ptnSrcs$). Therefore, the resulting subgraphs, and so the resulting sectors, are guaranteed to be directly connected to the large water mains, and so to the main water sources. If the sizes of all the resulting subgraphs are within the proper boundary (larger than $Model.DmaMinSize$ and smaller than $Model.DmaMaxSize$), the arrangement is considered as a feasible solution for the $Model.MajorIsland$. Algorithm 6 shows the overall procedure of this step. Figure 4.6 illustrates the flowchart of this process. As it can be seen in the flowchart, the following sub-steps are taken for each $MajorIsland$:

Step 5.1: Partition a $MajorIsland$ into K groups: Lines 1 to 15 in Algorithm 6 show the process of this step. First, the method examines to see if there are enough nodes to be used as seeds to grow graphs from. Let us consider $ptnSrcs4MI = PotentialSources \cap MajorIsland$, $K_{max} = \lfloor |MajorIsland| / DmaMinSize \rfloor$ [5], and $K_{min} = \lceil |MajorIsland| / DmaMaxSize \rceil$. If the number of nodes in $PtnSrcs4MI$ is less than K_{max}, the number of seed nodes is not enough, so new nodes need to be found and used as seeds. In such a case, Step 5.1.1 should be taken, otherwise Step 5.1.2.

Step 5.1.1: Find new potential sources. In this stage, the method finds new nodes in the $Model.MajorIsland$ that are connected to the large water mains with large pipes (which are definitely smaller than the $Model.MainsSizeThreshold$, as all the nodes that are connected to large water mains with links of size $Model.MainsSizeThreshold$ or larger were previously identified and added to $Model.ptnSrcs$). To this end, for each node in $PtnSrcs4MI$ a DFS algorithm is used to identify the neighbours in $Model.MajorIsland$ which are connected with links of size $Model.MainsSizeThreshold - ReductSteps$. The identified nodes are added to the $PtnSrcs4MI$ set. The method starts by $ReductSteps = 1$ and continues increasing it until at least K_{max} nodes are in $PtnSrcs4MI$. Algorithm 7 shows the overall procedure of this step.

Step 5.1.2: Select seeds and grow graphs. Lines 8 to 14 in Algorithm 6 show the process of this step. Once there are enough seed nodes, K is set to be the number of partitions; starting from $K = K_{min}$ and continuing to K_{max} (for example if size of

[5] Floor of size($MajorIsland$) divided by the predefined constant $DmaMinSize$.

Algorithm 6: PARTITIONMAJORISLANDS Partitions major islands into isolated components.

Input:
- *Model*: a data structure containing all the information about the network

Output:
- *Model*: updated *Model* in which *MajorIsands* are partitioned into new *DMAs*

$MI \leftarrow 1$;
foreach $MajorIsland$ *in* $Model.MajorIslands$ **do**
  $PtnSrcs4MI \leftarrow Model.PotentialSources \cap MajorIsland$;
  $K_{min} \leftarrow \lceil |MajorIsland| / Model.DmaMaxSize \rceil$;
  $K_{max} \leftarrow \lfloor |MajorIsland| / Model.DmaMinSize \rfloor$;
  if $|PtnSrcs4MI| < K_{max}$ **then**
    $ptnSrcs4MI \leftarrow$ FindNewPotSrcs$(Model, MajorIsland, K_{max})$
  for $K = K_{min}$ **to** K_{max} **do**
    $grps \leftarrow []$;
    for $i = 1$ **to** $Model.MaxIter$ **do**
      $srcs \leftarrow$ Choose K nodes from $ptnSrcs4MI$ on random (with replacement);
      $grps[i] \leftarrow$ GrowGraphs($Model, MajorIsland, srcs$);
    if $\forall i, Model.DmaMinSize \leq |grps[i]| \leq Model.DmaMaxSize$ **then**
    `// Feasible solution`
      $[Positions[MI], PositionDMA[MI]] \leftarrow$ DefineDMAs($Model, MajorIsland, BestScrs, grp$);
  $MI \leftarrow MI + 1$
$TotNumSols \leftarrow \prod_{i=1}^{MI} |Positions[i]|$;
for $i = 1$ **to** MI **do**
  $indx \leftarrow 0; n \leftarrow$ number of Positions for $MajaorIsland[i]$;
  for $j = 1$ **to** n **do**
    for $k = 1$ **to** $TotNumSols/n$ **do**
      $Sols[indx + k].Position \leftarrow Positions[indx + k] \& Positions[j]$;
      $Sols[indx + k].DMAs \leftarrow$ CombineDMAs $(PositionDMA[indx + k], PositionDMA[j])$;
    $indx \leftarrow indx + n$;
for $i = 1$ **to** $|Sols|$ **do**
  $Model.OverallSols[i].Position \leftarrow Sols[i].Position$;
  $Model.OverallSols[i].DMAs \leftarrow$ CombineDMAs $(Sols[i].DMAs, InitialDMAs)$;
return $Model$;

Algorithm 7: FINDNEWPOTSRCS Finds new potential sources for an *Island*.

Input:
- *Model*: a data structure containing all the information about the network
- *MajorIsland*: the group of nodes for which there is a need to find more potential sources
- *S*: the number of potential sources needed for the given *MajorIsland*

Output:
- *NewPotSrcs*: a list of nodes which can be considered as potential sources for the *MajorIsland*

```
SizeThreshold ← Model.MainsSizeThreshold;
step ← a positive integer number e.g., 1;
NewPotSrcs ← {};
while |NewPotSrcs| < S do
    SizeThreshold ← SizeThreshold − step;
    foreach potSrc ∈ Model.potSrcs ∩ MajorIsland do
        Run a DFS on MajorIsland and add all the nodes with links larger than
        SizeThreshold to NewPotSrcs;
return NewPotSrcs;
```

Model.MajorIsland is 10,000 and the proper sector size is between 500 and 5,000, any number between 2 and 20 partitions is acceptable; therefore, *K* starts from 2 and continues to increase up to 20). Then, in a loop of *Model.MaxIter* times (*e.g.*, 100), *K* nodes are selected from *PtnSrcs*4*MI* (random selection by replacement) to serve as *seeds*.

Then using *K* parallel BFS algorithms, the method grows *K* graphs from the selected seeds. Algorithm 8 shows the overall procedure of this step. The graph growing algorithm works as follows: starting from each *K* seed node in parallel, all the neighbouring nodes that are in *Model.MajorIsland* are added to a group corresponding to the seed node and will be removed from *Model.MajorIsland*. This process continues until all the nodes in *Model.MajorIsland* are visited and removed from it and are assigned to different groups.

The resulting arrangement of nodes into identified groups could be a feasible solution for partitioning this *Model.MajorIsland* if the size of all the resulting groups are within the limits of a proper sector size (larger than *Model.DmaMinSize* and smaller than *Model.DmaMaxSize*). If this condition holds for all the identified groups (line 13 in Algorithm 6), the resulting subgraphs are defined as *i*DMAs using the DefineDMAs

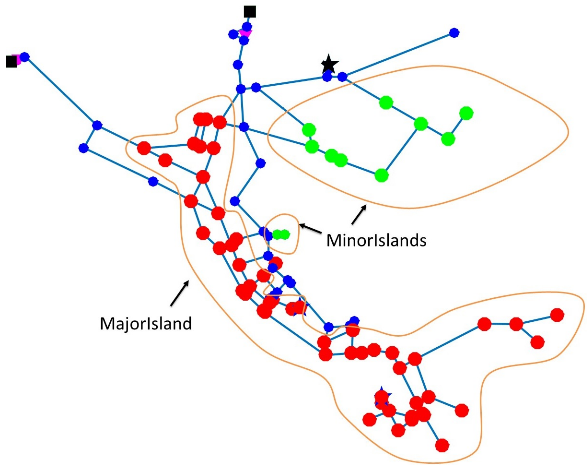


Fig. 4.5 Checking the size of $Islands$. The left-side group is a $MajorIsland$, while the tow smaller groups are $MinorIslnads$. No proper size island (*i.e.*, *i*DMA) is identified in this network at this step.

Algorithm 8: GROWGRAPHS Grows graphs in a major island to identify smaller components.

Input:

- *Model*: a data structure containing all the information about the network
- $MajorIsland$: the group of nodes that should be partitioned into smaller components
- *Srcs*: selected nodes to serve as potential sources for the given $MajorIsland$

Output:

- *grps*: a list containing groups of nodes created after partitioning the given $MajorIsland$

$grps \leftarrow []$;
forall the $src \in scrs$ **do in parallel**
    Grow a graph in $MajorIsland$ using BFS in the direction of flow and add the neighbours to $grps[src]$ and remove it from $MajorIsland$ until $MajorIsland$ empties;
return $grps$;

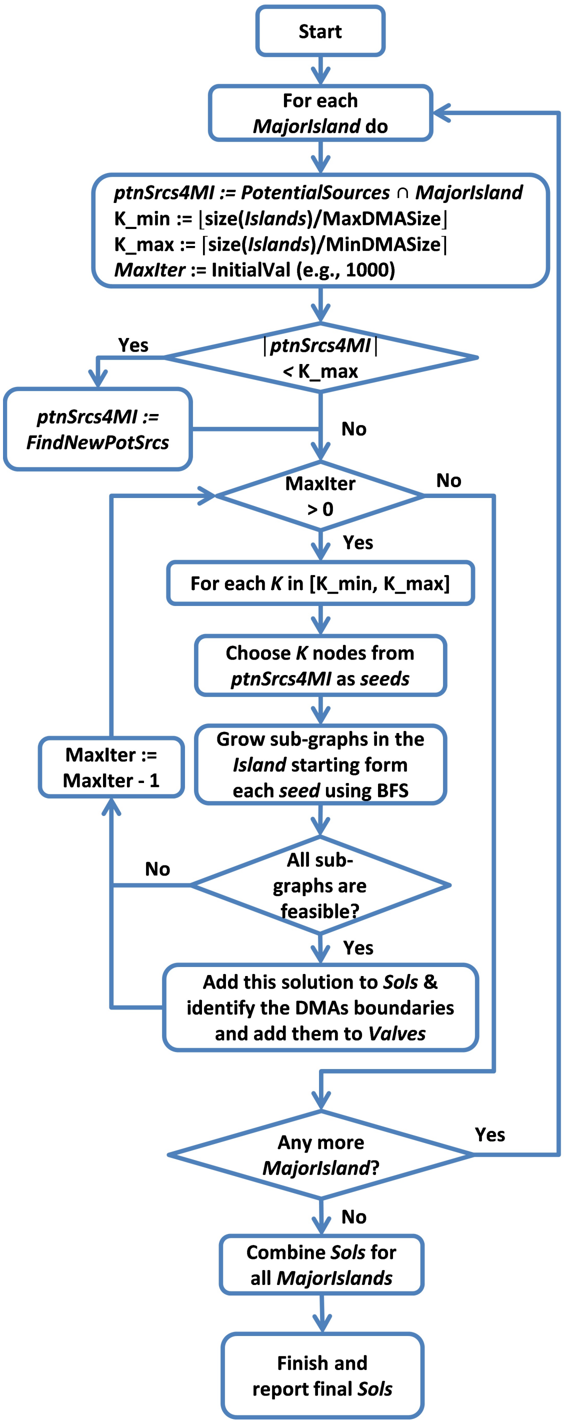


Fig. 4.6 Flowchart showing different steps of Partitioning *MajorIslands*.

method, which adds the identified subgraphs to the *i*DMAs corresponding to the solution; the first links from the selected *seeds* are added to $Model.Meters$ set; the neighbouring links [6] are identified and added to the $Model.Valves$ set, and the corresponding bits in the $Model.Position$ vector of the solution (the bit string representing the status of links in a solution) are set to 0. This process is repeated $Model.MaxIter$ times for each K, and the feasible solutions are identified [7]. Algorithm 9 shows the overall procedure of this step.

Algorithm 9: DEFINEDMAS Defines DMAs, identifies the positions of boundary links *i.e.*, meters and valves.

Input:
- $Model$: a data structure containing all the information about the network
- $MajorIsland$: the group of nodes that should be partitioned into smaller sectors
- $Srcs$: selected nodes to serve as potential sources for the given $MajorIsland$
- $grps$: a list of subgraphs created after graph growing to partition the given $MajorIsland$

Output:
- Sol: a vector of size $Model.LinkCount$ representing the status of links in the partitioned network
- $SolDMAs$: a list of DMAs corresponding to the given Sol

```
Sol ← Model.Position;
Find all links with one end node in one group and another one in another group and
set the corresponding bit in Sol to 0;
i ← 1;
foreach grp in grps do
    SolDMAs[i].Nodes ← grp;
    SolDMAs[i].Meters ← links between grp and Srcs;
    i ← i + 1;
return Sol, SolDMAs;
```

This is the end of Step 5.1 (Partition a $Model.MajorIsland$ into K groups). After that, the resulting solutions of partitioning different $MajorIslands$ should be combined to create overall solutions, which cover all the network. Algorithm 10 shows the overall procedure of this step.

Step 5.2: Combine solutions for different $MajorIslands$**.** As discussed, if there is

[6] The neighbouring links are the links with exactly one end in a group.

[7] If no feasible solution has been found, $Model.MaxIter$ should be increased (*e.g.*, to 500, 1000, etc.) and the algorithm should be run again. This issue is discussed in details in Section 5.5, page 99.

Algorithm 10: COMBINEDMAS Combines DMAs identified in different steps.

Input:
- dma_1: a list of identified DMAs created after partitioning one major island
- dma_2: a list of identified DMAs created after partitioning another major island

Output:
- *dmas*: a list containing of identified DMAs after combining the DMAs identified after portioning two major islands

```
dmas ← {};
if dma_1 is empty then
    dmas ← dma_2;
else if dma_2 is empty then
    dmas ← dma_1;
else
    indx ← 0;
    foreach dma in dma_1 do
        indx ← indx + 1;
        dmas[indx] ← dma;
    foreach dma in dma_2 do
        indx ← indx + 1;
        dmas[indx] ← dma;
Remove identical items from dmas;
return dmas;
```

more than one *Model.MajorIsland* in the network, all the resulting components of partitioning different *MajorIslands* should be reflected in the overall solution that covers the entire network. Therefore, the solutions related to different *MajorIslands* should be combined into overall solutions. This is done by performing bitwise *and* operations on *positions* of partial solutions and combining the corresponding *i*DMAs. Note that the solutions related to each *Model.MajorIsland* deal with a part of the network, so they are called partial solutions. Each partial solution corresponds to a specific set of links and represents the links that should be closed (the positions with 0) or remain open (the positions with 1). By doing bitwise *and* between all partial solutions, the overall positions of the links that should be closed are identified (note that the result of a bitwise *and* of 0 and either 0 or 1 will be 0; therefore, if a link is identified to be closed in one partial solution, it will be identified to be closed in the overall solution as well). As there might be multiple solutions for partitioning each *Model.MajorIsland*, all possible combinations of all solutions for different *MajorIslands* should be considered. Finally, the identified *i*DMAs in this step should be combined with the identified *i*DMAs from Step 4 to make the final candidate solutions. Lines 16 to 27 in Algorithm 6 show the process of this step.

After this stage, if there is at least one *Model.MajorIsland* in the network, it is possible to end up having a couple of different solutions, as there might be more than one solution for partitioning each *Model.MajorIsland*. Therefore, the best solution(s) should be identified. This will be done in the next step.

Step 6: Selecting the best solution(s). In this step, the solutions that do not have any advantages over the other solutions, with respect to their values for the selected objective functions [8], are removed. As explained in Chapter 3, WDN partitioning is a many-objective optimisation problem, so it might be impossible to choose a single solution as the best one, as some of the objectives are conflicting. Therefore, a set of Pareto optimal solutions or a representative sample of it should be identified, which make the Pareto front [55]. Pareto optimal solutions (in short, Pareto solutions) are the *non-dominated* ones. To

[8] In Step 1.

compare candidate solutions, the *dominance* concept [55] is used, in which a relationship is established between solutions instead of giving a scalar value to each solution. Solution A is said to dominate solution B if and only if A is no worse than B in terms of all the objective functions, and A is exactly better than B in terms of at least one objective function [55] (Algorithm 12 shows the logic of this comparison). All solutions are compared against each other, and if one solution dominates another one, the dominated solution will be removed from the candidate solutions set. Therefore, all the non-dominated solutions will remain. Figure 4.7 illustrates the flowchart of this step.

Algorithm 11: FINDBESTSOLS Performs optimisations and identifies the best solutions.

Input:
- *Model.OverallSols*

Output:
- *Model.BestSols*

```
/* Calculate costs                                              */
n ← number of solutions in Model.OverallSols;
for i = 1 to n do
    Model.OverallSols[i].Costs ← CalcCosts(Model.OverallSols[i])
Model.BestSols ← Model.OverallSols;
/* Removing dominated solutions                                 */
for j = n to 2 do
    for i = 1 to j-1 do
        A ← Model.BestSols[i];
        B ← Model.BestSols[j];
        if Dominates(A,B) then
            Remove Model.BestSols[j];
return Model.BestSols;
```

Each candidate solution's costs is calculated using the formulae 3.1 through 3.13. In order to obtain the values for the parameters involved in the objective functions 3.8 through 3.13 (*e.g.*, flow, pressure, hydraulic head, tank level, water velocity in pipes, and water age), an extended period simulation should be run for the candidate solutions. To this end, the resulting networks (the network in which the status of links are set based on the $Model.OverallSols[i].Position$ vector, for each i) are being sent to the hydraulic solver

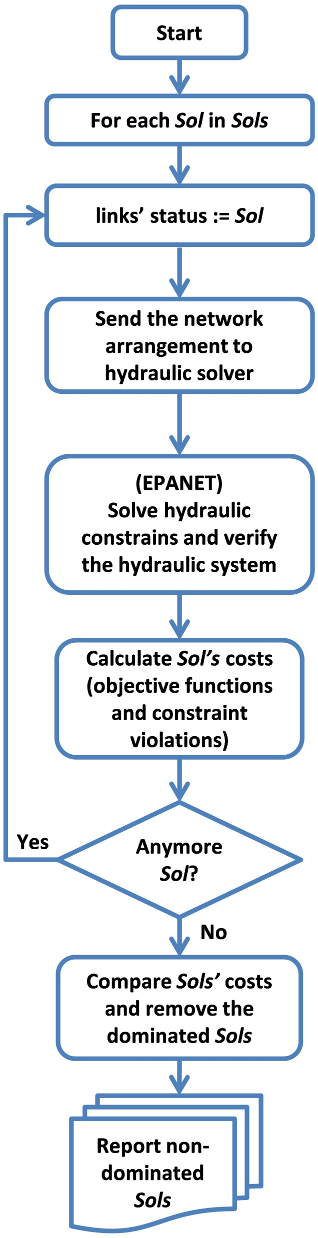


Fig. 4.7 Flowchart illustrating how the best candidate solutions are selected.

Algorithm 12: DOMINATES Identifies weather a solution dominates another one.

Input:

- Sol_A: a solution
- Sol_B: another solution

Output:

- $flag$: a boolean value which is $true$ if Sol_A dominates Sol_B, and $false$ otherwise

```
if all(Sol_A.Costs <= Sol_B.Costs) && any(Sol_A.Costs < Sol_B.Costs) then
    flag ← true;
else
    flag ← false;
return flag;
```

(*i.e.*, EPANET [103]) to solve the network hydraulic equations (*i.e.*, formulae 3.18 and 3.19), and achieve the network hydraulic information, *e.g.*, links' flow, nodal pressure, nodal hydraulic heads, tank levels, water velocity in pipes, nodal water age, etc. Once the required information is returned from the hydraulic solver, the objective functions can be calculated, and the constraint violations can be examined for the candidate solutions.

After removing the dominated solutions, the final results including the identified sectors ($Model.DMAs$), the set of links that should be equipped with meters ($Model.Meters$), the set of links that should be closed off ($Model.Valves$), and the related costs of each non-dominated solution will be reported in a lexicographical order [9], regarding the chosen objectives and their priorities. The domain experts should choose the best solution among the results based on their experience and other considerations.

Figure 4.8 illustrates one arrangement of the identified sectors in the Novato Network.

4.3 Implementation

WDN-PARTITION was implemented in MATLAB 2014b, and was integrated with EPANET 2. The software tool was implemented in a modular and interactive approach.

[9] 'In the lexicographic method, the objectives are ranked in order of importance by the designer. The optimum solution is then found by minimising the objective functions starting with the most important and proceeding according to the order of importance of the objectives' [101].

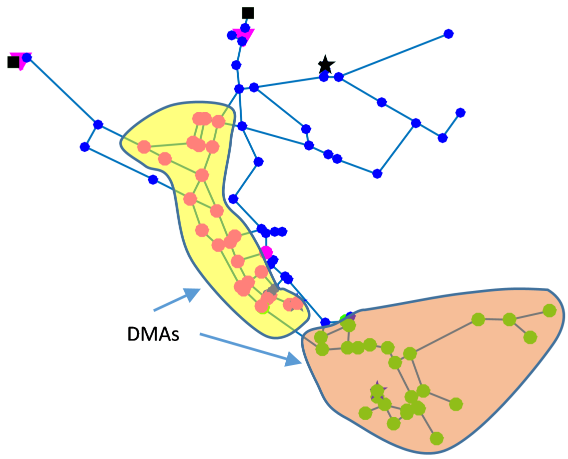


Fig. 4.8 Identified sectors in the Novato Network. The network is partitioned into 2 isolated sectors.

Different modules were implemented for the different steps of the method (*cf.*, 4.2). In each run, the user can choose the network file, view the topology of the network, choose the required objectives and their priorities, and set the initialisation parameters, *i.e.*, the minimum threshold size for the main pipes, the minimum and maximum sector size, and the maximum number of iterations in the loop. The tool starts partitioning the network and generating output reports about the details of the process. After finding structurally feasible solutions, the candidate solutions will be sent to the hydraulic solver (*i.e.*, EPANET 2). Hydraulic simulations can be time consuming for large networks. Once the hydraulic simulations are finished, the method will have all the required data to calculate different objective functions for the candidate solutions. Finally, the non-dominated solutions will be selected and their details will be reported. Additionally, the layout of the identified sectors will be presented as one of the outputs of the execution.

The experimental settings presented in Table 4.1.

Processor	**Cores**	**Memory**	**OS**	**Development Environment**
Intel Core i5	2	8 gb	Windows 8.1 64 bit	MATLAB 2014b 64 bit

Table 4.1 Experimental settings.

The execution time for 7 benchmark networks (*cf.*, Chapter 5) are given in Table 4.2.

As it can be seen in Table 4.2, the WDN-PARTITION execution time can be long for

Network Name	Nodes	Links	Simulation Duration	Hydraulic Time Step	Execution Time
Pescara Network	68	99	48 h	1 h	38 s
Nevato Network	97	119	24 h	1 h	69 s
BWSN Network 1	129	178	96 h	0.5 h	153 s
Modena Network	268	317	48 h	1 h	315 s
Balerma Network	443	454	48 h	1 h	217 s
Exeter Network	1891	3032	4h h	1 h	1.13 h
BWSN Network 2	12527	14831	48 h	1 h	~15 h

Table 4.2 Execution times for seven benchmark networks.

large networks (up to 15 hours for the largest one, *i.e.*, BWSN Network 2). However, it is not an issue for sectorisation projects, as they are not real-time applications. The analysis, design, and implementation of such projects are costly and time consuming. Typically, sectorisation is not a day-to-day task; *i.e.*, once it is completed, there is no need to repeat it (soon). Additionally, the experimental infrastructure was a personal laptop. High performance computing can be utilised if faster execution is needed.

4.4 Summary

This chapter explained how the WNS problem is solved in this thesis. It first discussed the design challenges and decisions. Then, the proposed method is discussed in details. Finally, a short section explained the implementation of WDN-PARTITION and the experimental setup. The next chapter will discuss how well WDN-PARTITION achieves its objectives in partitioning a WDN into a collection of *i*DMAs with minimal negative impact on the hydraulic requirements.

Chapter 5

Evaluation

This chapter evaluates how well WDN-PARTITION works for partitioning a WDN into isolated sectors all having direct access to a water source, with minimum negative impact on hydraulic requirements. It also compares the proposed method with the existing approaches, *i.e.*, Murray et al. [80], Diao et al. [38], and Ferrari et al. [47, 104].

The chapter is organised in five sections: the first section (*cf.*, Section 5.1), describes the general settings for the experiments. The second section (*cf.*, Section 5.2), explains the metrics that are used for evaluation of the results. The third section (*cf.*, Section 5.3), explains the networks that were used in this thesis as case studies to evaluate WDN-PARTITION, and the results of partitioning them into isolated sectors. Then, in the fourth section (*cf.*, Section 5.4), the proposed approach is compared against the existing approaches using the largest available network in the literature. Finally, in the fifth section (*cf.*, Section 5.5) the potential threats that may affect the validity of the study and the results are discussed.

5.1 Experimental setup

The evaluation of the proposed method is based on simulations, because they allow for controllable, repeatable, and scalable experiments. Additionally, evaluation of a partitioning scheme in the real world is not practical, as water distribution networks are critical urban infrastructures, and are not suitable for test and learn practices.

The hydraulic network simulator and solver EPANET 2 [103] was used to simulate the behaviour of each water distribution network under different network arrangements. Computational aspects of the experimental settings were described in Table 4.1 (page 78). The hydraulic aspects of the experimental setting are as follows: the default values of the network files are used in the experiments for all settings, except for the ‘Quality Parameter’ which is set to ‘Age’ to enable water age analysis, ‘Total Duration’ which is set to 48:00 hours (unless mentioned explicitly), and ‘Hydraulic Time Step’ which is set to 1:00 hour (unless stated otherwise).

WDN-PARTITION was designed and implemented in such a way that a network file and a configurable list of requirements and their priorities (from the list of objective functions stated in Chapter 3) are input. Therefore, for each network or each run, the user can choose what criteria are important and in what order.

5.2 Metrics for evaluation

The performance metrics used for evaluation were explained in detail in Chapter 3 (*cf.*, 3.3). They are derived from the WNS requirements (*cf.*, Chapter 2, Section 2.1), and are summarised in Table 5.1.

Metric	**Description**	**Related WNS requirement**	**Source (Equation)**
CS	cut size	SR6	3.1
CW	cut weight	SR7	3.2
SSI	sector size imbalance	SR5	3.3
AExp	average number of customer connections per sector	SR4	3.4
MExp	maximum number of customer connections in a sector	SR4	3.5
ALnkExp	average pipe length per sector	SR4	3.6
MLnkExp	maximum pipe length in a sector	SR4	3.7
PV	pressure violations	HR4	3.8
DP	dissipated power	HR9	3.9
ED	elevation differences within sectors	HR5	3.10
ANR	average network resilience	HR8	3.11
WA	average water age during the last 24 hours	HR10	3.12

Table 5.1 Performance metrics used for evaluation.

It should be noted that sector isolation (SR1), direct access (SR2), connectedness (SR3), conservation of mass (HR1), conservation of energy (HR2), and limited pressure (HR3) are considered as hard constraints in WDN-PARTITION, therefore, the method first satisfies them and then optimises for the objectives. Therefore, they are not considered as metrics for evaluation (but they are all satisfied in the method).

In running WDN-PARTITION on the case study networks, cost and background leakage are not chosen as criteria, as there is no exact data to calculate them for these networks. Therefore, no evaluation is presented for these two design objectives as well. As discussed in Section 3.3, page 47, to calculate background leakage, there should be data about the model parameters, *i.e.*, α_l and β_l, for different sectors, which is not available for the case study networks. For cost, the exact cost of partitioning, including the costs of cutting pipes and installing valves and meters and operating them, depends on the exact situations of the network at hand, which are not also available for the current cases. CS and CW can be regarded as surrogate measures for cost [104].

Additionally, the assessments of the majority of the metrics are not presented for the majority of the case studies, as there were not any baseline to compare the results. For case studies 1 to 6 only the assessment of CS, ANR, and WA are presented. For case study 7 (*i.e.*, the BWSN Network 2), the assessments of all the metrics in Table 5.1 are presented.

5.3 Case studies

WDN-PARTITION has been applied to a series of publicly available water distribution networks, including the 3 networks that were shipped with EPANET as examples [103], the 12 networks that were used as benchmarks for multi-objective design optimisation of water distribution networks [120] [1], and the 2 networks that were used as benchmarks for battle of water network sensor placement [87] [2].

It should be noted that partitioning is useful for large networks, and a network with 50 or even 500 nodes cannot adequately reflect the benefits of partitioning. Of the networks

[1]Site address (last visit: 1 Aug. 2015):
http://emps.exeter.ac.uk/engineering/research/cws/resources/benchmarks/design-resiliance-pareto-fronts/.
It should be noted that the data files for these networks are based on the best-known Pareto fronts for two objective (network resilience and the total cost) design optimisation of these networks.

[2]Site address (last visit: 1 Aug. 2015):
http://emps.exeter.ac.uk/media/universityofexeter/emps/research/cws/downloads/Ostfeld_et_al._BWSN_files.zip

that were used as case studies, two are large networks, namely, the Exeter Network and the BWSN Network 2. However, as there are not many large networks available in the literature, five small to medium sized networks are chosen as case studies as well, namely, the Pescara Network, the Novato Network, the BWSN Network 1, the Modena Network, and the Balerma Network. The sizes of all these networks are 50 or more nodes. [3] For the smaller networks, it was necessary to set the sector size boundaries to small numbers (in terms of the number of nodes) to enable their partitioning. These numbers will be stated explicitly for each case.

Table 5.2 illustrates the networks used as case studies in this thesis and gives a summary of the results of applying WDN-PARTITION on them.

	Nodes	Pipes	Res	Pareto Sols	DMAs	CS	Meters	ANR	WA
Pescara Network	68	99	3	3	3	2 - 3	4-5	0.95 - .0.98	0.25 - 0.44
Nevato Network	92	117	2	4	2	1 - 3	2	0.89 - 0.84	7.08 - 9.52
BWSN Network 1	126	168	1	8	3	5 - 10	7 - 8	0.87 - 0.83	34.12 - 42.56
Modena Network	268	317	4	51	4 - 9	11 - 24	7-21	0.88 - 0.92	1.04 - 1.38
Balerma Network	443	454	4	11	5 - 8	1 – 3	10 - 15	089 – 0.93	12.54 – 14.73
Exeter Network	1891	3032	2	47	14 - 29	2 - 19	70 - 90	0.51 – 0.55	15.05 - 15.69
BWSN Network 2	12523	14822	2	78	28 - 48	66 - 161	56 - 78	0.83 - 0.72	31.01 - 33.14

*__Nodes:__ number of nodes in the network; **Pipes:** number of pipes in the network; **Res:** number of reservoirs in the network; **Pareto Sols:** number of non-dominated (Pareto) solutions; **DMAs:** number of identified sectors; **CS:** cut size; **Meters:** number of meters that should be installed; **ANR:** average network resilience; **WA:** average water age during the last 24 hours.

Table 5.2 Networks used as case studies in this thesis and a summary of the results of applying WDN-PARTITION on them.

These networks are studied in the order of their size (in terms of the number of nodes) in the rest of this chapter. For the case studies 1 to 6, only the sectorisation arrangements are illustrated and brief summaries of the results are given. The largest network, *i.e.*, BWSN Network 2, is evaluated in more detail in case study 7, and is used as a benchmark for comparing WDN-PARTITION against the existing approaches (*i.e.*, Murray et al. [80], Diao et al. [38], and Ferrari et al. [47, 104]).

[3] The other networks in [120], *i.e.*, the Two-Reservoir Network, the Two-Loop Network, the BakRyan Network, the New York Tunnel Network, the Hanoi Network, the GoYang Network, and the Fossolo Network, are too small (less than 50 nodes), so they are not discussed in this chapter.

5.3.1 Case study 1: The Pescara Network

The Pescara Network [12] is a real world WDN of a city in Italy. It includes 68 demand nodes, 3 reservoirs, and 99 pipes. Figure 5.1 illustrates one arrangement of the identified sectors in the Pescara Network as a result of applying WDN-PARTITION on the network. The minimum and maximum size of a sector were set to 10 and 30 nodes respectively to enable its partitioning. This network was partitioned into three isolated sectors. There were 3 non-dominated solutions in the Pareto front, with 4 to 5 meters each, and 2 to 3 pipes that should be closed off to isolate the sectors. The average resilience index is between 0.95 to 0.98, and average water age during the last 24 hours of simulations is between 0.25 and 0.44 hours for different solutions.

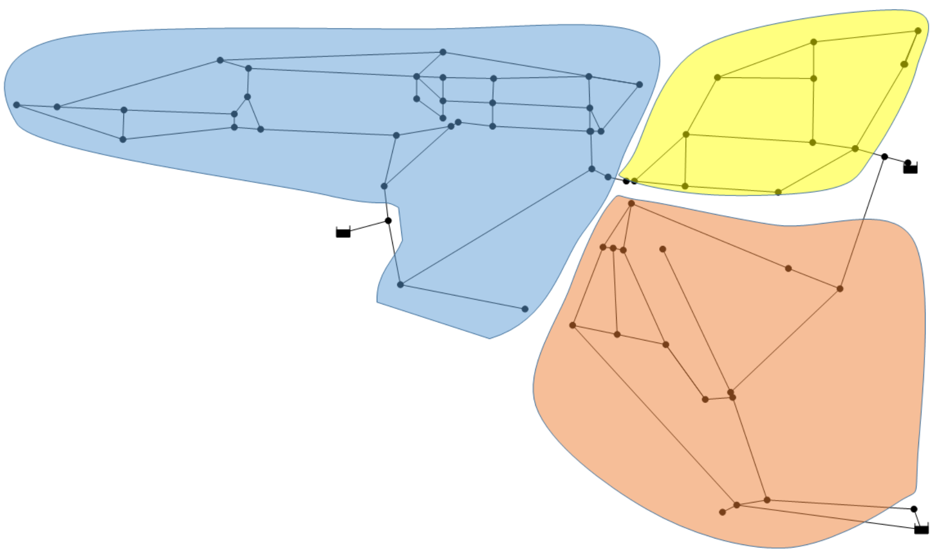


Fig. 5.1 Identified sectors in the Pescara Network. The network is partitioned into 3 isolated sectors.

5.3.2 Case study 2: The Novato Network

The Novato Network [103] was introduced and studied in Chapter 4 to illustrate the different stages of the proposed approach. As discussed, the Novato Network has 92 nodes and 117 pipes, and is supplied from 2 reservoirs. The minimum and maximum size of a sector were set to 20 and 50 nodes respectively to enable its partitioning. WDN-PARTITION partitions this network into two isolated sectors. Figure 5.2 illustrates one arrangement of the identified sectors in the Novato Network as a result of applying WDN-PARTITION on the network. There were 4 non-dominated solutions in the Pareto front, with 2 meters each, and between 1 to 3 pipes that should be closed off to isolate the sectors. Average resilience index is between 0.89 to 0.84, and average water age during the last 24 hours of simulations is between 7.08 and 9.52 hours for different solutions.

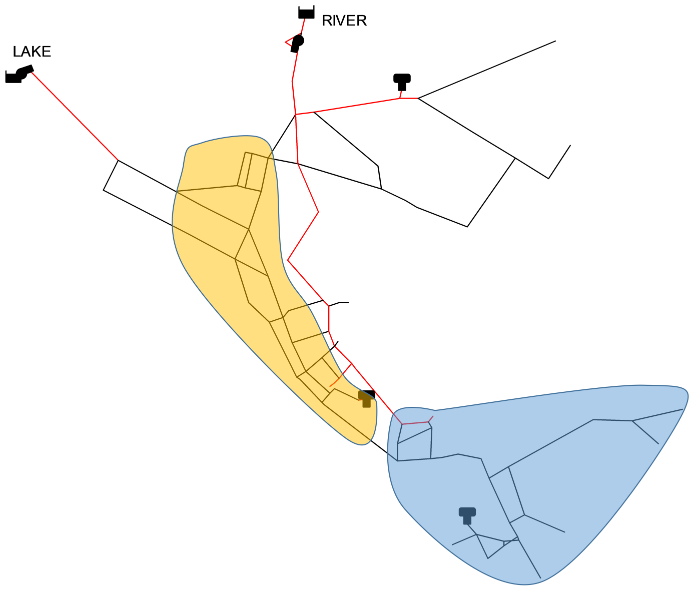


Fig. 5.2 Identified sectors in the Novato Network. The network is partitioned into 2 isolated sectors.

5.3.3 Case study 3: The BWSN Network 1

BWSN Network 1 [87] has 126 nodes and 168 pipes, and is supplied from 1 reservoir. The network is a real water distribution system, but it has been modified to preserve its anonymity; however, these modifications do not affect the connectivity and the hydraulic behaviour of the network [87]. Figure 5.3 illustrates one arrangement of the identified sectors in the BWSN Network 1 as a result of applying WDN-PARTITION on the network. The minimum and maximum size of a sector were set to 20 and 40 nodes respectively to enable its partitioning. WDN-PARTITION partitions this network into three isolated sectors. There were 8 non-dominated solutions in the Pareto front, with 7 to 8 meters, and between 5 to 10 pipes that should be closed off to isolate the sectors. The average resilience index is between 0.87 to 0.83, and average water age during the last 24 hours of simulations is between 34.12 and 42.56 hours for different solutions.

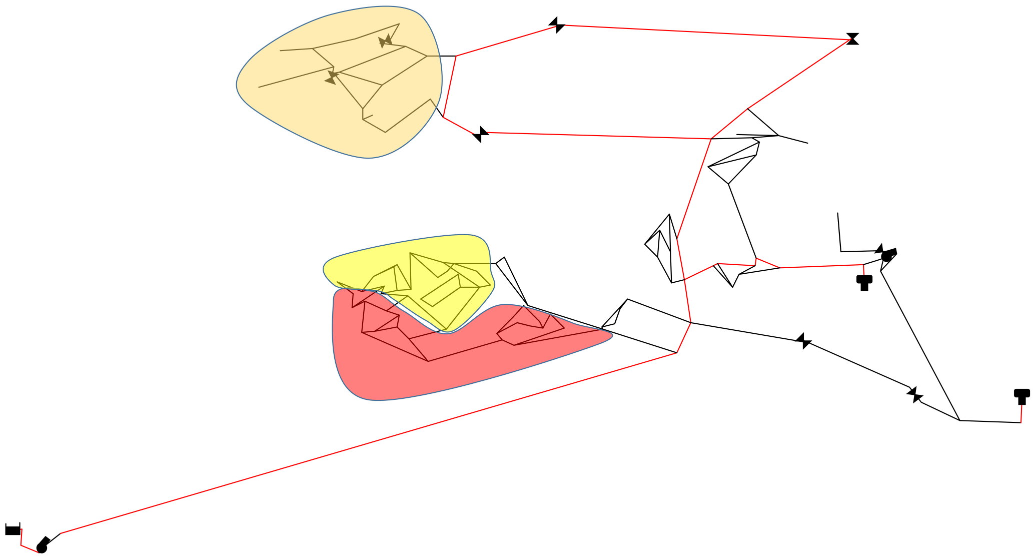


Fig. 5.3 Identified sectors in the BWSN Network 1. The network is partitioned into 3 isolated sectors.

5.3.4 Case study 4: The Modena Network

The Modena Network [12] includes 268 demand nodes, 4 reservoirs, and 317 pipes, and is a real world WDN of a city in Italy. Figure 5.4 illustrates one arrangement of the identified sectors in the Modena Network as a result of applying WDN-PARTITION on the network. The minimum and maximum size of a sector were set to 10 and 60 nodes respectively to enable its partitioning. There were 51 non-dominated solutions in the Pareto front, that partition the network into 4 to 9 isolated sectors, with 7 to 21 meters each, and 11 to 24 pipes that should be closed off to isolate the sectors. The average resilience index is between 0.88 to 0.92, and average water age during the last 24 hours of simulations is between 1.04 and 1.38 hours for different solutions.

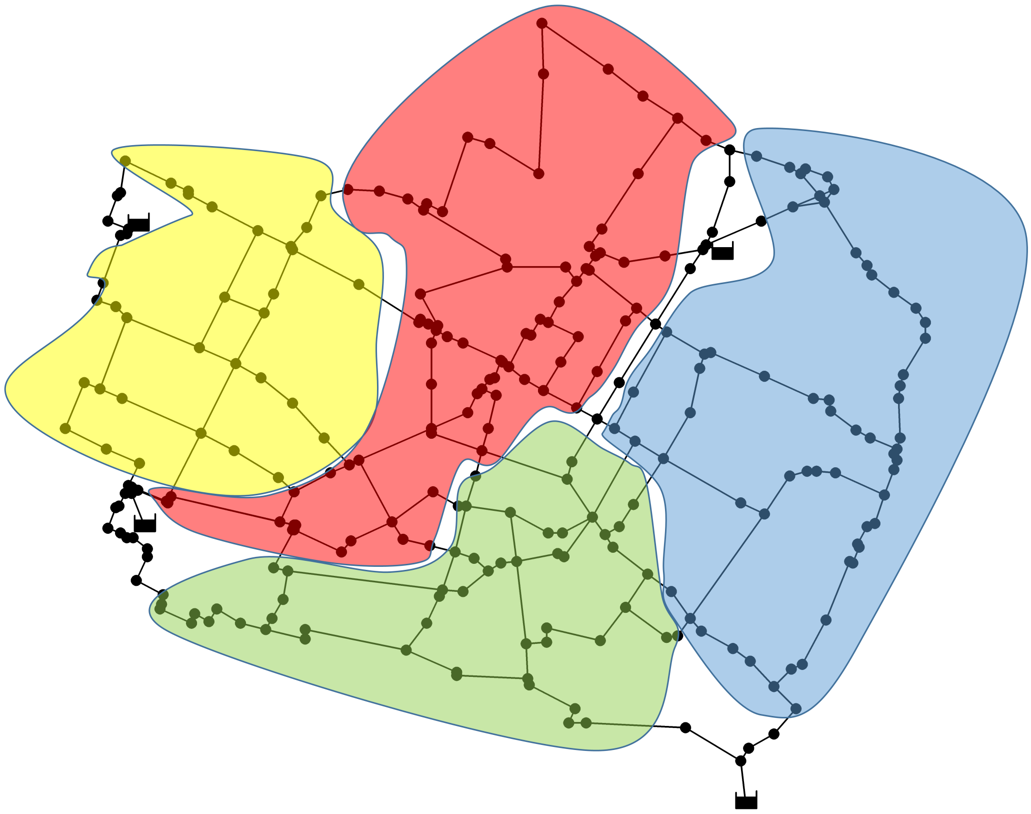


Fig. 5.4 Identified sectors in the Modena Network. This arrangement partitions the network into 4 isolated sectors.

5.3.5 Case study 5: The Balerma Network

The Balerma Irrigation Network [102] includes 443 demand nodes, 4 reservoirs, and 454 pipes, and is a real world WDN in Italy. Figure 5.5 illustrate one arrangement of the identified sectors in the Balerma Network as a result of applying WDN-PARTITION on the network. The minimum and maximum size of a sector were set to 30 and 100 nodes respectively to enable its partitioning. There were 11 non-dominated solutions in the Pareto front, that partition the network into 5 to 8 isolated sectors, with 10 to 15 meters each, and 1 to 3 pipes that should be closed off to isolate the sectors. The average resilience index is between 0.89 to 0.93, and average water age during the last 24 hours of simulations is between 12.54 and 14.73 hours for different solutions.

5.3.6 Case study 6: The Exeter Network

The Exeter Network [44] includes 1891 demand nodes, 2 reservoirs, and 3032 pipes. Figure 5.6 illustrate one arrangement of the identified sectors in the Exeter Network as a result of applying WDN-PARTITION on the network. The minimum and maximum size of a sector were set to 20 and 200 nodes respectively to enable its partitioning. There were 47 non-dominated solutions in the Pareto front, that partition the network into 14 to 29 isolated sectors, with 70 to 90 meters each, and 2 to 19 pipes that should be closed off to isolate the sectors. The average resilience index is between 0.51 to 0.55, and average water age during the last 24 hours of simulations is between 15.05 and 15.69 hours for different solutions.

5.3.7 Case study 7: The BWSN Network 2

The BWSN Network 2 [87] comprises 12523 nodes, 2 constant head sources, 2 tanks, 14822 pipes, 4 pumps, 5 valves, and is subject to 5 variable demand patterns. It serves about 150000 people and is a good example of a large water distribution system. The system was simulated for a total period of 48 hours. This WDN was used as a test bed for a number of modelling exercises, including the Battle of the Water Sensor Networks competition [87]. The network is a real water distribution system, but it has been modified to preserve its

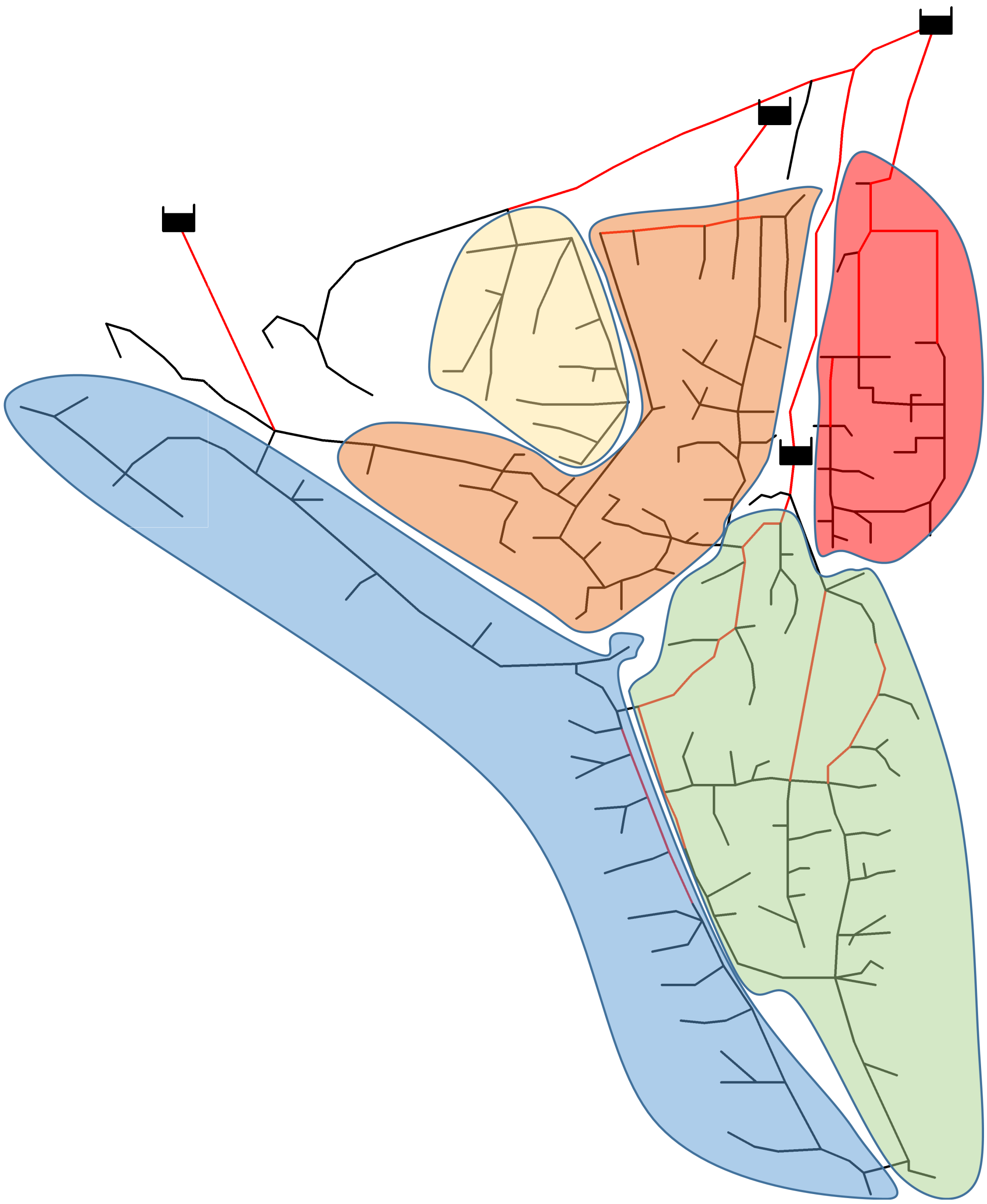


Fig. 5.5 Identified sectors in the Balerma Network. This arrangement partitions the network into 5 isolated sectors.

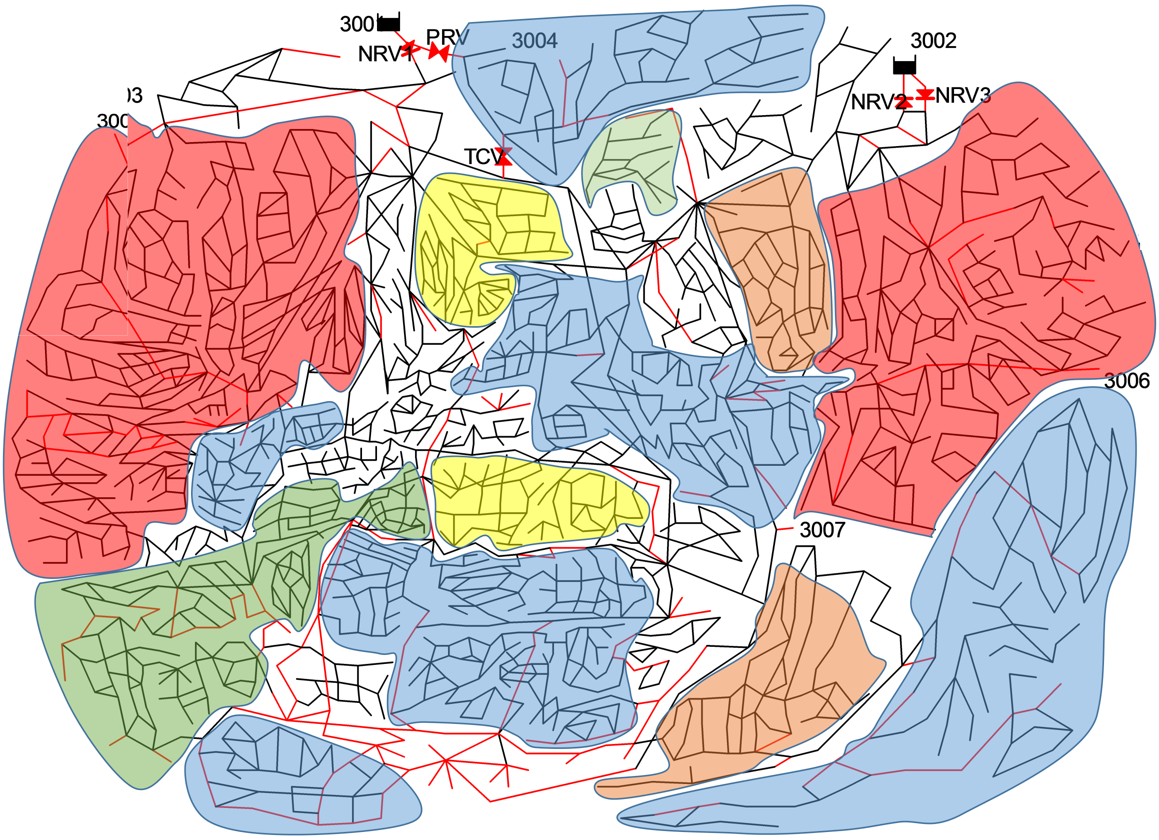


Fig. 5.6 Identified sectors in the Exeter Network. This arrangement partitions the network into 14 isolated sectors.

anonymity; however, these modifications do not affect the connectivity and the hydraulic behaviour of the network [87].

The BWSN Network 2 was used as benchmark in Murray et al. [80], Diao et al. [38], and Ferrari et al. [47, 104]; therefore, it will be studied in more detail in the rest of this chapter to support the evaluation of WDN-PARTITION.

Figure 5.7 illustrates the skeletonised layout of the BWSN Network 2 as it can be seen in the EPANET software, and Table 5.3 summarises the network characteristics and the hydraulic parameters which are used in the experiments for the BWSN Network 2 (these parameters are the default ones set in the network file).

As there is no exact data for the number of customer connections per node in the BWSN Network 2, the average customer connection per node was used instead. The total number of customer connections is 77916, so to have a boundary of 500-5000 customer connections per sector, 80-800 nodes per sector on average were considered [4]. Additionally, the threshold size for the large water mains is set to 14 inches (350 mm) for this network in the initialisation phase; *i.e.*, $Model.MainsSizeThreshold = 14$ (inches). The maximum number of iterations in Step 5.1.2 (*i.e.*, Select seeds and grow graphs, *cf.*, Section 4.2), is set to 100 in the initialisation phase (*i.e.*, $Model.MaxIter = 100$). As there is no detailed information for each individual node, $H_{s,j}^{min}$ in Equation 3.8 (which is the minimum admissible head at node j in scenario s) and H_i^* in Equation 3.11 (which is the minimum required head at node i) were set to 28 (m) for all the demand nodes [5], in all scenarios (*cf.*, Section 3.3). Finally, $C_{s,j}^{pen}$, which is the penalty cost of pressure deficiency for node j in scenario s, is assumed to be 1 in all cases, in the absence of exact penalty costs for different nodes in this network.

Table 5.4 summarises the algorithm parameters which are used in the experiments for the BWSN Network 2.

---

[4] $12523/77916 * 500 \simeq 80$, and $12523/77916 * 5000 \simeq 803$.

[5] 28 m (~30 psi) is the default minimum head usually considered in design of water distribution networks. The same number is used in the baseline approaches, *i.e.*, [47, 80].

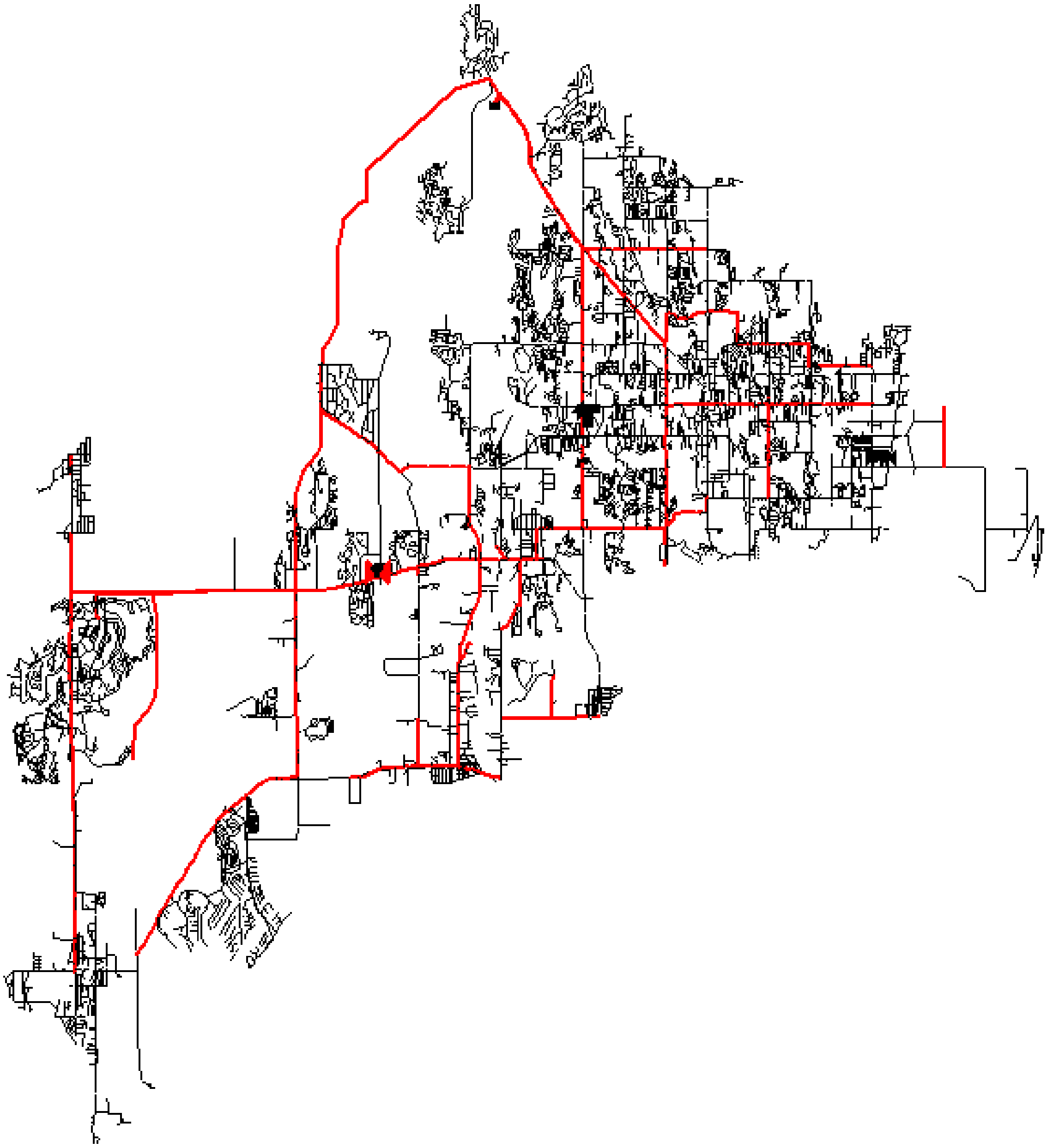


Fig. 5.7 The skeletonised layout of the BWSN Network 2 as it can be seen in the EPANET software. The large water mains (pipes with diameters larger than or equal to 14 inches (350 mm) are highlighted.

Parameter	Value
Population served	$\sim$ 150000
Customer connections	77916
Number of junctions	12523
Number of reservoirs	2
Number of tanks	2
Number of pipes	14822
Number of pumps	4
Number of valves	5
Quality parameter	Age
Total duration	48:00
Hydraulic time step	1:00
Quality time step	0:05
Pattern time step	1:00
Pattern start time	0:00
Reporting time step	1:00
Report start time	0:00
Clock start time	12 am
Pump efficiency	75%
Flow units	GPM
Headloss formula	Hazen-Williams
Specific gravity	1
Relative viscosity	1
Maximum trials	40
Accuracy	0.001
If unbalanced	Stop
Default pattern	Pattern-0
Demand multiplier	1.0

Table 5.3 The BWSN Network 2 network information and the hydraulic parameters used in the experiments, which are the default parameters set in the network file.

Results

WDN-PARTITION identifies three $MajorIslands$. After partitioning them and combining the solutions, the method recognises 81 structurally feasible candidate solutions. After removing the dominated ones, the method ended up with 78 non-dominated solutions [6], *i.e.*, Pareto solutions. The Pareto solutions have between 28 and 48 sectors of sizes between 1515 and 2425 customer connections per sector. The total number of closed pipes vary

[6] The reason why the majority of the solutions remained non-dominated is the large number of objectives.

Parameter	**Value**
Sector size	80-800 (nodes per sector)
$Model.MainsSizeThreshold$	14 inches (350 mm)
$H_{s,j}^{min}$ and H_i^*	28 (m)
$C_{s,j}^{pen}$	1
$Model.MaxIter$	100

Table 5.4 Initialisation parameters to run the algorithm on BWSN Network 2.

from 66 to 161, and the total number of meters vary from 56 to 78 in different solutions.

The values of objective functions for 10 different Pareto solutions are reported in Table 5.5. These 10 solutions are selected from the set of 78 non-dominated solutions after sorting them based on the least pressure violations (*i.e.*, PV), most average resilience index (*i.e.*, ANR), and least average water age during the last 24 hours of simulation (*i.e.*, WA), as they were chosen as high priority requirements [7].

	DMAs	**CS**	**CW**	**SSI**	**AExp**	**MExp**	**ALnkExp**	**MLnkExp**	**PV**	**DP**	**ED**	**ANR**	**WA**
Sol-01	28	66	494	0.879	2423	4844	1.55E+5	3.25E+5	0.00E+0	3.83E+7	330.14	0.83	31.01
Sol-02	30	79	624	0.860	2263	4206	1.56E+5	3.23E+5	0.00E+0	3.82E+7	329.49	0.83	31.05
Sol-03	30	78	618	0.860	2262	4206	1.38E+5	3.34E+5	0.00E+0	3.88E+7	362.53	0.83	32.08
Sol-04	30	87	662	0.880	2261	4906	1.66E+5	3.27E+5	2.11E+0	1.10E+8	317.74	0.82	32.12
Sol-05	30	87	662	0.880	2262	4906	1.66E+5	3.26E+5	2.13E+0	1.10E+8	317.66	0.82	32.42
Sol-06	30	90	629	0.882	2262	4969	1.42E+5	2.93E+5	5.32E+0	8.53E+7	356.80	0.81	32.41
Sol-07	30	90	639	0.882	2263	4969	1.42E+5	2.93E+5	5.44E+0	8.44E+7	356.56	0.81	32.45
Sol-08	30	84	588	0.875	2263	4719	1.56E+5	3.33E+5	9.92E+0	8.33E+7	329.31	0.81	32.45
Sol-09	30	82	562	0.879	2262	4850	1.56E+5	3.86E+5	1.11E+1	8.65E+7	330.67	0.80	32.46
Sol-10	31	94	673	0.859	2190	4175	1.34E+5	3.02E+5	1.70E+1	8.28E+7	365.31	0.80	32.43

***DMAs**: number of identified sectors; **CS**: cut size; **CW**: cut weight; **SSI:** sector size imbalance; **AExp**: average number of customer connections per sector; **MExp**: maximum number of customer connections in a sector; **ALnkExp**: average pipe length per sector; **MLnkExp**: maximum pipe length in a sector; **PV**: pressure violations; **DP**: dissipated power; **ED**: sum of the standard deviations of elevation differences within sectors; **ANR**: mean network resilience; **WA**: average water age during the last 24 hours of the simulation.

Table 5.5 Top 10 non-dominated solutions for partitioning the BWSN Network 2, sorted lexicographically based on least pressure violations, most average network resilience, and least average water age during the last 24 hours of simulation.

Sol-01 has the best objective functions in terms of cut size, cut weight, pressure violations, resilience index, and water age, but is not doing as well in terms of other objective functions; *i.e.*, Sol-02 outperforms Sol-01 in terms of sector size imbalance, average and maximum number of customer connections per sector. For other objective

[7] As discussed, the criteria and their priorities can be set for each network and each run.

functions the results are almost the same. Sol-06 and Sol-07 have the same cut size, but their cut weight is different, which mean they have identified different neighbouring links to be cut to isolate different sectors. As it goes downwards, the identified solutions are subordinate in terms of the main three selected objective functions, but as they are better in terms of other objective functions (especially, sector size imbalance, average and maximum number of customer connection per sector) they are in the Pareto front of non-dominated solutions.

The arrangement of one of the Pareto solutions (Sol-01) is illustrated in Figure 5.8, which identifies 28 sectors of sizes between 588 to 4844 nodes (2423 on average). This solution demonstrates a 1% decrease in resilience index and a less than 1% increase in water age, comparing to the original network.

Figure 5.9 illustrates the Pareto front of 10 lexicographically ordered non-dominated solutions and how they behave against each other in terms of pressure violations, network resilience, and water age objective functions.

The results demonstrate that WDN-PARTITION generally achieves its design objectives to partition a water network into isolated sectors. The sizes of all the identified sectors are within the minimum and maximum acceptable sector size. Additionally, all the identified sectors are isolated from each other and all have direct access to a water source.

5.4 Comparison with the existing approaches

In this section, WDN-PARTITION is compared with the existing solutions using the BWSN Network 2 as the benchmark.

5.4.1 Baselines

Three methods used BWSN Network 2 as their benchmark:

- Murray et al. [80], which manually partitions BWSN Network 2 into 43 sectors using engineering principles with trial and error. This work is a good representative

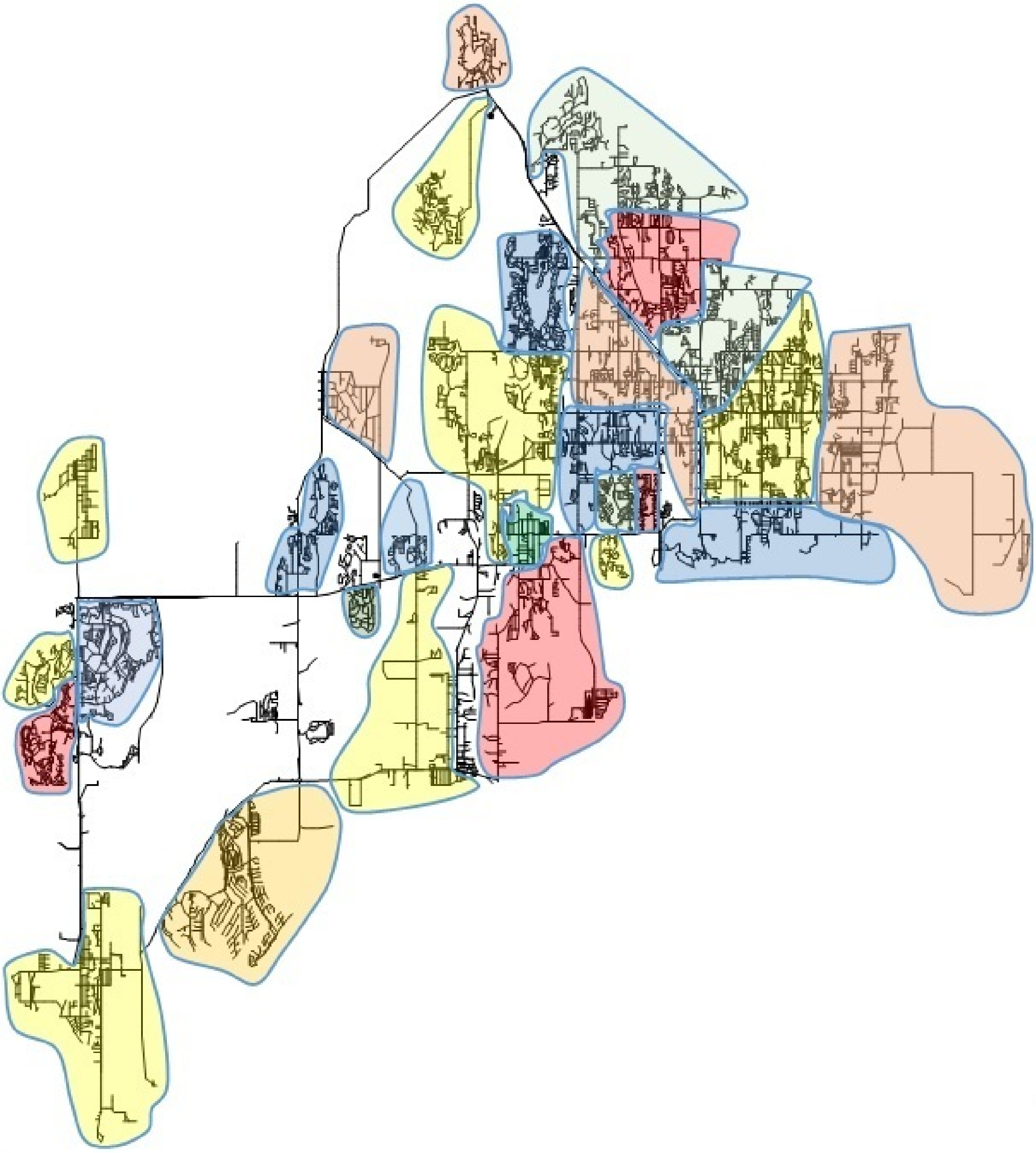
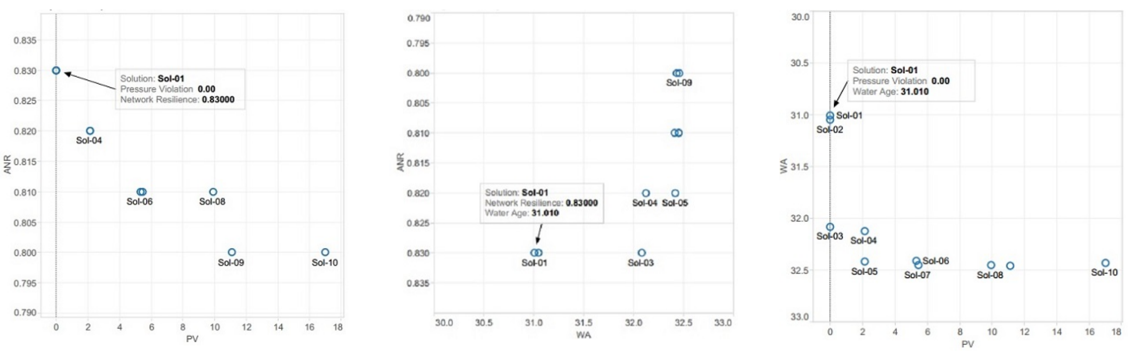


Fig. 5.8 Identified sectors in BWSN Network 2. The arrangement illustrated in this figure is one of the Pareto solutions (Sol-01 in Table 5.5), which identifies 28 sectors, of sizes between 588 to 4844 customer connections (2423 on average).



(a) Pressure violation *vs.* average network resilience.

(b) Average water age during the last 24 hours *vs.* average network resilience.

(c) Pressure violations *vs.* average water age during the last 24 hours.

Fig. 5.9 Pareto Front of 10 lexicographically ordered non-dominated solutions for partitioning the BWSN Network 2. These figures show how different solutions behave against each other in terms of 3 objective functions which had been specified as important criteria in the initialisation phase.

of applying sound engineering by domain experts to partition a WDN into isolated sectors.

- Diao et al. [38], which applies community detection for the automated identification of sector boundaries in BWSN Network 2. This work is a good representative of the approaches that partition a WDN into smaller sectors; however, water security was not a main concern in the design of the method.

- Ferrari et al. [47, 104], which applies graph theory to partition BWSN Network 2 into isolated sectors. This work is an influential work for this thesis, and automatically partitions a WDN into isolated sectors.

5.4.2 Discussion

Table 5.6 illustrates the results of partitioning BWSN Network 2 using WDN-PARTITION in comparison with the three baseline approaches (*cf.*, 5.4.1).

Murray et al. [80] manually partition the BWSN Network 2 into 43 sectors using engineering principles with trial and error. They modified the network by adding 11 pipes to have at least two feed lines for each sector to ensure sufficient redundancy within the system. The average sector size is 1996 connections. Three sectors are smaller than 500 and one is slightly larger than 5000. Additionally, one sector is not directly connected to a source. The total number of closed pipes and meters are 163 and 53, respectively. The average water age during the last 24 hours of simulation is 31.51 hours. The other metrics are not reported in this work.

Diao et al. [38] applied community detection for the automated identification of sector boundaries in BWSN Network 2. This approach identifies 41 sectors, from which at least one is not directly connected to a source (this can be seen in the arrangement illustrated in the paper). Two sectors are larger than the maximum allowable size (5000 customer connections), and one sector is smaller than the minimum allowable sector size (500 customer connections). The average sector size is 2044 customer connections, and the average water age during the last 24 hours of simulation is 32.01 hours. The other metrics are not reported in this work.

Ferrari et al. [47, 104] applied graph theory to partition the BWSN Network 2 into isolated sectors. As this method applied a multi-objective optimisation method to find the sub-optimal solutions, a set of non-dominated solutions results, comprising of solutions with 32 to 43 sectors. The reported sub-optimal solution identifies 36 sectors, of which at least 4 sectors are not directly connected to a source (this can be seen in the arrangement illustrated in the paper). One sector is larger than the maximum acceptable size, and 3 sectors are smaller than the minimum acceptable sector size. The average sector size of the reported solution is 1863 customer connections per sector, and the total number of closed pipes is 152 [8]. The average reported resilience index is between 0.81 and 0.82 for different solutions, while the average reported water age during the last 24 hours of simulation is between 31.04 and 31.62 hours for different solutions. The other metrics were not reported in this work.

WDN-PARTITION identifies between 28 and 48 sectors in different Pareto solutions. All the identified sectors are isolated from each other and have a direct access to a source. No sector has less than 500 or more than 5000 customer connections. The average sector size varies from 1515 to 2425 customer connections for different solutions. The total number of

[8] In [47], 152 is reported, while in [104], between 53 and 132 for different Pareto solutions is reported. Communications with the authors for clarification remained unanswered.

closed pipes is between 66 and 161 for different solutions. The average resilience index is between 0.83 and 0.72, while the average water age during the last 24 hours of simulation is between 31.01 and 33.14 hours for different Pareto solutions. The total number of closed pipes vary from 66 to 161, and the total number of meters vary from 56 to 78 in different Pareto solutions.

All of the identified sectors WDN-PARTITION have sizes within the predefined boundary (*i.e.*, 500-5000 customer connection per sector) and all are isolated from each other to maximise water security, which is the main focus of this research. However, this is obtained at the cost of more cut size, comparing to [47] [9]. In terms of average number of customer connections per sector, resilience index, and water age, WDN-PARTITION identifies solutions that outperform the baselines [38, 47, 80, 104].

	DMAs	ND	LT5000	ST500	CS	Meters	AExp	ANR	WA	PA
Original Network	-	-	-	-			-	0.84	30.71	-
Murray et al. [24]	43	1	1	3	163	53	1996	0.802	31.51	11
Diao et al. [12]	41	1	2	1	NA	NA	2044	NA	32.01	0
Ferrari et al [14]	32 - 43	4	1	3	53 - 132	NA	1863	0.82 - 0.81	31.04 - 31.62	0
WDN-PARTITION	28 - 48	0	0	0	66 - 161	56 - 78	1415 - 2423	0.83 - 0.72	31.01 - 33.14	0

***DMAs**: number of identified sectors; **ND**: number of sectors not directly connected to source; **LT5000**: number of sectors larger than 5000 customer connections; **ST500**: number of sectors smaller than 500 customer connections; **CS**: cut size; **Meters:** number of meters that should be installed; **AExp:** average number of customer connections per sector; **ANR**: average network resilience; **WA**: average water age during the last 24 hours; **PA**: number of pipes added to the network.

Table 5.6 Comparison of the results of partitioning the BWSN Network 2 for WDN-PARTITION against three baseline approaches.

5.5 Threats to validity

The following aspects of the proposed method and the experimental setup potentially limit the conclusions that can be drawn from the study.

Random component in the algorithm. As explained in Section 4.2, there is a random component in the proposed method in Step 5.1.2 (*i.e.*, Select seeds and grow graphs). In this step, if no feasible solution has been found, *Model.MaxIter* should be increased

[9]Unfortunately, [47] did not report the details of different Pareto solutions it identified; therefore, it is not possible to compare different Pareto solutions in detail.

(*e.g.*, to 500, 1000, etc.) and the algorithm should be run again. If no feasible solution is found even with large values for $Model.MaxIter$, the method is not able to find a feasible partitioning solution for the network at hand (which may imply the need for creation of cascading DMAs [10]); however, this has never happened in the case studies using a number of benchmark networks available in the literature (*cf.*, Section 5.3).

Sub-optimality. WDN-PARTITION is a heuristic method, and similar to every other heuristic method finding the global optimal solution is not guaranteed. However, the results are structurally and hydraulically feasible, and they are near-optimal.

Simulation. Although simulation is an established and widely used evaluation method for water distribution network analysis, its simplification of the real world may lead to results that do not reflect the actual network behaviour in the real world. However, as the simulation setup is the same for all the examined sectorisation schemes, the simulation allows for comparing their performance. Additionally, the hydraulic simulator which is used in this research, *i.e.*, EPANET [103], is a widely used hydraulic simulator and network solver in the literature, which is well trusted in both academia and industry and is used as the basis for several commercial models [119].

5.6 Chapter summary

This chapter discussed how well the proposed water distribution network sectorisation method (*i.e.*, WDN-PARTITION) works for partitioning a WDN into isolated sectors all having direct access to a water source, with minimal negative impact on hydraulic requirements. It first explained the experimental setup and the general settings. Next, the results of applying WDN-PARTITION to partition seven case study networks into isolated sectors were discussed in detail. Then, WDN-PARTITION was compared with three baseline approaches (*i.e.*, Murray et al. [80], Diao et al. [38], and Ferrari et al. [47, 104]), and it was shown that the proposed method is competitive with the baseline approaches, so

[10] 'There are situations where there is a hierarchy of metered zones and cascading zones, where the water from one DMA first passes through another' [79].

it can be considered as a good alternative for the existing solutions. The results show that WDN-PARTITION achieves its design objectives to partition a WDN into isolated sectors, satisfying all structural requirements with less than 1% deterioration in resilience index and water age comparing to the original network. Finally, the threats to the validity of the study were discussed.

Chapter 6

Conclusions

This thesis investigated the water network sectorisation problem and proposed a heuristic method to partition a water distribution network into isolated sectors, satisfying structural requirements with minimum negative impact on hydraulic requirements.

This chapter summarises the thesis and its achievements. It discusses the advantages and limitations of the proposed water distribution network sectorisation method, *i.e.*, WDN-PARTITION, and highlights potential areas for future work.

6.1 Discussion

As discussed in Chapter 2, WNS is a multifaceted problem embracing graph-theoretic, hydraulic, and economic aspects. Existing techniques in graph theory cannot handle structural challenges of WNS (*cf.*, Section 2.1.1); therefore, this thesis proposed a novel algorithm that takes the structural requirements of the problem into consideration. Additionally, besides the structural requirements, WNS has hydraulic and economic requirements, *i.e.*, constraints that should be satisfied and objectives that should be optimised; therefore, it is a constrained multi-objective optimisation problem.

In this thesis, a set of WNS requirements and metrics was first defined to reflect WNS structural, hydraulic, and economic requirements. These requirements then were considered in the sectorisation process, and the outcomes of sectorisation were evaluated using the defined metrics.

WDN-PARTITION addresses water security requirements for water network sectorisation, *i.e.*, sector isolation and direct access to at least one source for each sector, with minimum negative impact on the other structural (connectedness of the network and sector size limitations) and hydraulic (customer demand satisfaction, pressure requirements, network reliability, energy efficiency [1], limited water velocity in pipes, minimum nodal elevation differences within the sectors, and water quality) requirements. Simulation based evaluations showed WDN-PARTITION generally achieves its design objectives with a minimal decrease (less than 1%, *cf.*, Chapter 5) in the performance criteria of the network.

WDN-PARTITION partitions a water network into isolated sectors without adding any new component to the network, and works well for both small and large networks, as the number of sources is not a limiting factor for the number of sectors. The criteria of optimisation and their priorities can be specified for each case, which makes the method a general purpose one.

WDN-PARTITION is a heuristic method, which is specific to water network sectorisation, and takes full advantage of the particularities of the problem. It is acknowledged that the optimisation method applied in this work is rather different from the search-based methods which are often used in engineering [91]. Search-based techniques had been applied in the earlier stages of this research; *i.e.*, Tabu search [60], genetic algorithms (NSGA-II) [63], and multi-agent systems [62]. However, the solutions generated by these techniques usually do not satisfy the structural requirements of the problem, specifically, sector isolation and direct access to (at least) one water source for each sector.

In [63], a large number of solutions were generated and the optimal or near-optimal ones were found (which entailed a lot of time-consuming hydraulic simulations). Then the resulting solutions were examined to see if they were structurally feasible. This process was inefficient in terms of both computation time and its capability to generate feasible solutions

[1] Dissipated power reflects energy efficiency (HR9); *cf.*, Section 2.1.

[2]. In practice, it did not work for large networks, *e.g.*, BWSN Network 2. Therefore, the final idea used in this thesis was to reduce the search space by generating a collection of structurally-feasible solutions first, and then finding the best solutions among the feasible ones, which entails a radically reduced number of hydraulic simulations. This way, the efficiency of the optimisation method increases noticeably and the results are structurally feasible as well.

The proposed water network sectorisation method, *i.e.*, WDN-PARTITION, applied a novel structural graph partitioning technique combined with a many-objective optimisation procedure to find near-optimal sectorisation solutions. The key characteristics of the proposed method are:

- Each sector has direct access to a water source, so the path from the sector to a source does not contain any nodes in other sectors.
- The identified sectors are isolated from each other, so there is no flow exchange between different sectors.
- The sizes of the identified sectors are within a pre-defined boundary.
- The candidate solutions are hydraulically analysed, and the best ones are chosen using a many-objective optimisation procedure.
- The criteria of optimisation and their priorities can be specified for each network and each run.
- It works well for both small and large networks.
- It performs sectorisation without adding any new component to the network.

6.2 Thesis contribution

This research contributed to the body of knowledge by providing:

[2]Even after a long execution time, most of the identified results were structurally infeasible.

Extension to WNS requirements and metrics. To perform sectorisation correctly, a general set of WNS requirements must be applied in the process. Additionally, a general set of metrics is needed to quantify the different aspects of the sectorisation results and to support evaluations. Existing design criteria and metrics do not cover all aspects of WNS (*cf.*, Chapter 2). This study extended the current design criteria [79] and proposed a more general set of *i*DMA design criteria (WNS requirements, *cf.*, Section 2.1), which were derived from the structural and hydraulic requirements of water network sectorisation. In addition to the WNS requirements, this study extended the metrics to quantify the performance of a partitioning scheme, so enhanced the measures to compare different approaches. Specifically, sector size imbalance (SSI), average pipe length per sector (ALnkExp), maximum pipe length in a sector (MLnkExp), and elevation differences within sectors (ED) were defined as metrics in this thesis (*cf.*, Section 3.3). The problem statement (*cf.*, Chapter 3) complemented this contribution by providing a mathematical formulation of the WNS problem as a constrained multi-objective optimisation problem, which can be tackled by other researchers or professionals.

Extension to existing WDN partitioning methods. Current solutions for WDN partitioning do not address all the requirements of water network sectorisation. In particular, water security is not addressed in most of the existing approaches as comprehensive as it is addressed in this thesis. Those approaches that do address water security cannot handle large networks (as the number of sources is a limiting factor for the number of sectors), and/or do not address other WDN partitioning requirements (*e.g.*, pressure requirements during different consumption scenarios, energy efficiency [3], limited water velocity in pipes, minimum nodal elevation differences within the sectors, and the requirement that sectors should cross as few mains as possible). WDN-PARTITION partitions a water network into isolated sectors guaranteeing that each sector is of a good size and is directly connected to a water source, while minimising the negative impact on the hydraulic requirements.

[3] Dissipated power is a metric for energy efficiency of the network (*cf.*, Section 2.1).

Extension to graph partitioning techniques. As discussed in Section 1.2.1 and Chapter 2 and 3, water network sectorisation can be reduced to a graph partitioning problem after removing the hydraulic details and complexities. However, no graph-theoretic graph partitioning method that addresses component/sector [4] isolation and direct access to some specific nodes (*sources*) for each component/sector was found in the course of this research. A heuristic digraph partitioning technique is proposed in this thesis that partitions a digraph with two types of nodes (*i.e.*, *sources* and *consumers*), which guarantees component isolation and direct access to at least one *source* for each component, while minimising the number of cuts, cut weight, and sector size imbalance (*cf.*, Chapter 4). However, this method has not been designed for general graphs/networks and only targeted for specific digraphs with two types of nodes (*i.e.*, *sources* and *consumers*) that have the discussed structural characteristics and requirements.

6.3 Limitations

Despite its advantages, this study has limitations. As discussed in Section 5.5, WDN-PARTITION is a heuristic method, and similar to every heuristic method, finding the global optimal solution is not guaranteed. However, the results are structurally and hydraulically feasible, and the best ones within the scope of the generated feasible solutions are being selected (that is why the term *near-optimal* was used in this thesis).

Additionally, WDN-PARTITION is an algorithmic method and can only follow a predefined set of steps. There might be room to improve the results by hydraulic and/or urban infrastructure planning experts.

In the sectorisation method, only closure of links (pipes and valves) was studied. It might be useful in some situations to add new components to the network; however, it is out of the scope of this thesis. Additionally, dynamic pump and/or valve operations optimisation might prove to be effective in DMA management; nonetheless, they were not studied in this

[4] A *component* is the term used in the graph theory literature [14] for the smaller subgraphs resulted after partitioning a graph. A *sector* is the same notion, which is a smaller sub-network resulted after partitioning a water distribution network.

research. Furthermore, infrastructure conditions (*e.g.*, the layout of roads, streets, bridges, etc.) were also not studied in this thesis.

WDN-PARTITION was designed to partition a water distribution network into isolated sectors; *i.e.*, general networks/graphs were not studied, and generalisation of the proposed method to other types of networks is out of the scope of this thesis. The method assumes a previously designed water distribution network as input, and suggests a set of near-optimal partitioning solutions for the input network as output (that can be ordered in terms of costs, pressure violations, or other criteria). Design and calibration of the networks are also out of the scope of this thesis.

Finally, as discussed in Chapter 5, simulations are used to evaluate the proposed method. Although simulation is an established and widely used evaluation method for water distribution network analysis, its simplification of the real world may lead to results that do not reflect the actual network behaviour. However, as discussed in Section 5.5, the simulation setup is the same for all examined sectorisation schemes, therefore, it allows for controllable, repeatable, and scalable experiments. Additionally, the hydraulic simulator which is used in this research, *i.e.*, EPANET [103], is a widely used hydraulic simulator and network solver in the literature, which is well trusted in both academia and industry and is used as the basis for several commercial models [119]. After all, water distribution networks are critical urban infrastructures; it was not practical to use a real world water distribution network as a test bed to apply sectorisation on it in the course of this research.

6.4 Future work

The proposed method could be improved in some aspects; *e.g.*,

Seed selection optimisation. In the seed selection step (Step 5.1.2), it may be helpful to consider some additional criteria to select better seeds (*e.g.*, the diameter and/or flow of the connecting links, the demand of the seed nodes, and the distance of the seed nodes from each other and/or from the large water mains).

Sector boundaries optimisation. In Step 5.1.2, after graph growing, it may be helpful to further optimise the boundaries of the identified subgraphs considering subgraph size, nodal elevation differences in each subgraph, the number and the diameter of the pipes that should be closed off, distances of the nodes in a group, nodal demands, etc.

Merging MinorIslands. In some cases, it might be possible to merge nearby $MinorIslands$ with each other to create a sector, or merge them with the neighbouring sectors.

Adding new components. As discussed in the previous section, WDN-PARTITION only considers the closure of links (pipes and valves). Adding new components (*e.g.*, pipes and/or pumps) to the network to further optimise the performance of the sectorised network can be further studied.

Dynamic pump and/or valve operations. Dynamic pump and/or valve operations optimisation to further optimise the performance of the sectorised network is a possible area of more studies. Additionally, further studies may need to investigate the application of pressure reducing valves (PRVs) to further optimise the performance of the sectorised network, specially in terms of leakage reduction.

Improving the software tool. The software tool was developed as a proof of concept in the course of this research. It can be improved such that it could be used in an interactive 'decision support' mode; *i.e.*, the user could be able to see the results, change the parameters and settings if required, and see the new results based on the changes.

This thesis examined a new research direction for water distribution network sectorisation to improve its manageability and security. Besides its contributions to the body of knowledge, the thesis may serve as a basis for further investigations in areas such as dynamic sectorisation and water smart grid.

Dynamic sectorisation. Currently, partitioning a water distribution network into a collection of sectors is performed permanently by cutting pipes and installing valves. This has the potential to result in water outage in a large area if a main pipe bursts or a main source or reservoir has an outage. A new area of research is developing to investigate how sectors can be created dynamically [95, 121, 122]. Such improvements may suggest an active approach towards management of a water distribution network using monitoring (using technologies such as smart meters, remote sensing, environmental sensors, and sensor networks, etc.), simulation (what-if analysis), computation (data processing, multi sensor data fusion, data / big data analytics, etc.), and remote control technologies (such as remotely controllable valves).

Water smart grid. Information and communications technology (ICT) provides two-way communication, data storing and analysis, sensing and system monitoring, and intelligence such as historical and predictive analytics. The *energy smart grid* initiative [43, 77] proposes a highly decentralised and flexible energy supply and distribution by taking advantage of ICT infrastructures. In contrast, many water utilities do not employ technology options beyond basic SCADA [5] systems, though ICT has the potential to provide water management with capabilities such as measuring, sensing, optimising, and detecting the status of water and supporting infrastructure. A *water smart grid* [62] could be a transformation of the current water supply systems using ICT capabilities. Sectorisation may improve manageability of a water distribution network by providing more granular information about different parts of the network. This can be a stepping stone towards a water smart grid.

Additionally, the ideas of this thesis might be generalisable in other domains. As an example, energy networks are another critical urban infrastructure. Partitioning a power network into self-sufficient sub-divisions can stop propagation of disturbances and prevent cascading events [74]. Generalisation of the ideas of this thesis may prove to be useful to further develop the literature in energy domain [74, 107, 125].

It is hoped that the findings of this thesis offer a new perspective on water network sectorisation and improving security of this critical urban infrastructure, and encourage further research in this area.

[5] Supervisory control and data acquisition.

References

[1] ALVISI, S., AND FRANCHINI, M. A heuristic procedure for the automatic creation of district metered areas in water distribution systems. *Urban Water Journal 11*, 2 (Feb. 2014), 137–159.

[2] ALVISI, S., AND FRANCHINI, M. A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering 70* (2014), 41–50.

[3] ANDREEV, K., AND RACKE, H. Balanced Graph Partitioning. *Theory of Computing Systems 39*, 6 (Oct. 2006), 929–939.

[4] ATKINSON, S., FARMANI, R., MEMON, F. A., AND BUTLER, D. Reliability Indicators for Water Distribution System Design: Comparison. *Journal of Water Resources Planning and Management 140*, 2 (Feb. 2014), 160–168.

[5] AWAD, H., KAPELAN, Z., SAVIC, D., AND SAVIĆ, D. A. Optimal setting of time-modulated pressure reducing valves in water distribution networks using genetic algorithms. *Integrating Water Systems (CCWI 2009 Conference)* (2009), 31–37.

[6] BAÑOS, R., RECA, J., MARTÍNEZ, J., GIL, C., AND MÁRQUEZ, A. L. Resilience Indexes for Water Distribution Network Design: A Performance Analysis Under Demand Uncertainty. *Water Resources Management 25*, 10 (Mar. 2011), 2351–2366.

[7] BARBEAU, B., GAUTHIER, V., JULIENNE, K., AND CARRIERE, A. Dead-end flushing of a distribution system: Short and long-term effects on water quality. *Journal of Water Supply Research and Technology-Aqua 54*, 6 (2005), 371–383.

[8] BLONDEL, V. D., GUILLAUME, J.-L., LAMBIOTTE, R., AND LEFEBVRE, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment 2008*, 10 (Oct. 2008), P10008.

[9] BLUM, C., AND ROLI, A. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Comput. Surv. 35*, 3 (Sept. 2003), 268–308.

[10] BOCCALETTI, S., LATORA, V., MORENO, Y., CHAVEZ, M., AND HWANG, D. Complex networks: Structure and dynamics. *Physics Reports 424*, 4-5 (Feb. 2006), 175–308.

[11] BONDY, J. A., AND MURTY, U. S. R. *Graph Theory with Applications*, vol. 28. 1976.

[12] BRAGALLI, C., D'AMBROSIO, C., LEE, J., LODI, A., TOTH, P. Water Network Design by MINLP. *IBM Research Division 24495*, Report RC24495 (2008).

[13] BROTHERS, K. J., REGIONAL, H., AND COMMISSION, W. DMA Design and Application for Non-Revenue Water Reduction in Canada. *Proceedings of the IWA Leakage Management Conference Cyprus 2002 1* (2002), 1–5.

[14] BULUC, A., MEYERHENKE, H., SAFRO, I., SANDERS, P., AND SCHULZ, C. Recent Advances in Graph Partitioning. 1–36.

[15] BURROWS, R., CROWDER, G. S., AND ZHANG, J. Utilisation of network modelling in the operational management of water distribution systems. 83–95.

[16] CAMPBELL, E., AYALA-CABRERA, D., IZQUIERDO, J., PÉREZ-GARCÍA, R., AND TAVERA, M. Water Supply Network Sectorization Based on Social Networks Community Detection Algorithms. In *16th Conference on Water Distribution System Analysis, WDSA 2014* (2014), vol. 00.

[17] CHARALAMBOUS, B. Experiences in dma redesign at the water board of lemesos, cyprus. In *International water association's specialised conference on leakage. Halifax, Nova Scotia, Canada* (2005), pp. 403–413.

[18] CLAUSET, A., NEWMAN, M., AND MOORE, C. Finding community structure in very large networks. *Physical Review E 70*, 6 (Dec. 2004), 066111.

[19] COELLO, C., PULIDO, G., AND LECHUGA, M. Handling multiple objectives with particle swarm optimization. *IEEE Transactions on Evolutionary Computation 8*, 3 (June 2004), 256–279.

[20] COELLO, C. A. C., LAMONT, G. B., AND VELDHUIZEN, D. A. V. *Evolutionary Algorithms for Solving Multi-Objective Problems Second Edition.* 2007.

[21] DAVID W. CORNE, JERRAM, N. R., KNOWLES, J. D., OATES, M. J., AND J, M. PESA-II: Region-based selection in evolutionary multiobjective optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001)* (2001).

[22] DAVIDSON, I. Clustering With Constraints: Feasibility Issues and the k-Means Algorithm. In *Proceedings of the 2005 SIAM International Conference on Data Mining* (2005).

[23] DAVIDSON, I., AND RAVI, S. The complexity of non-hierarchical clustering with instance and cluster level constraints. *Data mining and knowledge discovery* (2007).

[24] DAVIDSON, I., AND RAVI, S. S. Identifying and generating easy sets of constraints for clustering. In *Proceedings of the 21st National Conference on Artificial Intelligence - Volume 1* (2006), AAAI'06, AAAI Press, pp. 336–341.

[25] DAVIS, L. *Handbook of genetic algorithms*, vol. 115. Van Nostrand Reinhold New York, 1991.

[26] DE PAOLA, F., FONTANA, N., GALDIERO, E., GIUGNI, M., AND SORGENTI, G. Automatic multi-objective sectorization of a water distribution network. In *16th Conference on Water Distribution System Analysis, WDSA 2014* (2014), vol. 00, pp. 1–8.

[27] De Paola, F., Fontana, N., Galdiero, E., Giugni, M., Uberti, G. S. D., and Vitaletti, M. Optimal Design of District Metered Areas in Water Distribution Networks. *Procedia Engineering 70*, Lambert 2002 (2014), 449–457.

[28] Deb, K. Multi-objective Optimization. In *Search Methodologies SE - 15*, E. K. Burke and G. Kendall, Eds. Springer US, 2014, pp. 403–449.

[29] Deb, K., and Pratap, A. A fast and elitist multiobjective genetic algorithm: NSGA-II. *..., IEEE Transactions on 6*, 2 (2002), 182–197.

[30] Delgado, D. M. Infrastructure leakage index (ili) as a regulatory and provider tool. *Arizona: Department of Civil Engineering & Engineering Mechanics* (2008).

[31] Deuerlein, J. W. Decomposition Model of a General Water Supply Network Graph. *Journal of Hydraulic Engineering 134*, 6 (June 2008), 822–832.

[32] Di Nardo, A., and Di Natale, M. A heuristic design support methodology based on graph theory for district metering of water supply networks. *Engineering Optimization 43*, 2 (Feb. 2011), 193–211.

[33] Di Nardo, A., Di Natale, M., and Di Mauro, A. *Water Supply Network District Metering - Theory and Case Study*. CISM International Centre for Mechanical Sciences. Springer Vienna, Vienna, 2013.

[34] Di Nardo, A., Di Natale, M., Santonastaso, G. F., Tuccinardi, F. P., Zaccone, G., and Paolo, F. SWANP : software for automatic Smart WAter Network Partitioning. In *7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,* (2014), vol. 3, pp. 1576–1582.

[35] Di Nardo, a., Di Natale, M., Santonastaso, G. F., Tzatchkov, V. G., and Alcocer-Yamanaka, V. H. Water network sectorization based on a genetic algorithm and minimum dissipated power paths. *Water Science & Technology: Water Supply 13*, 4 (Aug. 2013), 951.

[36] Di Nardo, A., Di Natale, M., Santonastaso, G. F., Tzatchkov, V. G., and Alcocer-Yamanaka, V. H. Water network sectorization based on graph theory and energy performance indices. *Journal of Water Resources Planning and Management 140*, 5 (2013), 620–629.

[37] Di Nardo, A., Di Natale, M., Santonastaso, G. F., and Venticinque, S. An Automated Tool for Smart Water Network Partitioning. *Water Resources Management* (Aug. 2013), 4493–4508.

[38] Diao, K., Zhou, Y., and Rauch, W. Automated Creation of District Metered Area Boundaries in Water Distribution Systems. *Journal of Water Resources Planning and Management 139*, 2 (Mar. 2013), 184–190.

[39] Diestel, R. *Graph Theory (Graduate Texts in Mathematics)*. 2006.

[40] Dijkstra, E. W. A note on two problems in connexion with graphs, 1959.

[41] EMPALING, A. J. Sibu Water Board : Challenges in Implementing District Metering Area Pilot District Metering Areas (DMA). 1–8.

[42] ERNST, N. A., MYLOPOULOS, J., AND WANG, Y. *Design Requirements Engineering: A Ten-Year Perspective*, vol. 14. 2009.

[43] FARHANGI, H. The path of the smart grid. *IEEE Power and Energy Magazine 8*, 1 (2010), 18–28.

[44] FARMANI, R., SAVIC, D., AND WALTERS, G. "EXNET" Benchmark Problem for Multi-Objective Optimization of Large Water Systems, Modelling and Control for Participatory Planning and Managing Water Systems. In *IFAC workshop* (Venice, Italy., 2004).

[45] FERNÁNDEZ, A. H. *Improving water network management by efficient division into supply clusters*. PhD thesis, Universitat Politècnica de València, 2011.

[46] FERRARI, G., AND SAVIC, D. Economic Performance of DMAs in Water Distribution Systems. *Procedia Engineering 119* (2015), 189–195.

[47] FERRARI, G., SAVIC, D., AND BECCIU, G. Graph-Theoretic Approach and Sound Engineering Principles for Design of District Metered Areas. *Journal of Water Resources Planning and Management 140*, 12 (Dec. 2014), 04014036.

[48] FLEMING, P. J., PURSHOUSE, R. C., LYGOE, R. J., STREET, M., AND SHEFFIELD, S. Many-Objective Optimization: An Engineering Design Perspective. *Lecture Notes in Computer Science 3410/2005* (2005), 14–32.

[49] FLOYD, R. W. Algorithm 97: Shortest path. *Communications of the ACM 5*, 6 (June 1962), 345.

[50] FORTUNATO, S. Community detection in graphs. *Physics Reports 486*, 3-5 (2010), 75–174.

[51] GAREY, M. R., AND JOHNSON, D. S. Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences). *Computers and Intractability* (1979), 340.

[52] GIUSTOLISI, O., BERARDI, L., LAUCELLI, D., SAVIC, D., WALSKI, T., BRUNONE, B., AND WDSA 2014 TEAM. Battle of Background Leakage Assessment for Water Networks (BBLAWN). In *Water Distribution Systems Analysis Conference* (2014), vol. 89, pp. 4–12.

[53] GIUSTOLISI, O., AND RIDOLFI, L. New Modularity-Based Approach to Segmentation of Water Distribution Networks. *Journal of Hydraulic Engineering 140*, 10 (2014), 04014049.

[54] GIUSTOLISI, O., AND SAVIC, D. Identification of segments and optimal isolation valve system design in water distribution networks. *Urban Water Journal 7*, 1 (Feb. 2010), 1–15.

[55] GOLDBERG, D. E. *Genetic algorithms in search, optimization, and machine learning*, vol. 412. Addison-wesley Reading Menlo Park, 1989.

[56] GOMES, R., MARQUES, A. S., AND SOUSA, J. Decision support system to divide a large network into suitable District Metered Areas. *Water science and technology : a journal of the International Association on Water Pollution Research 65*, 9 (Jan. 2012), 1667–75.

[57] GOMES, R., SÁ MARQUES, A., AND SOUSA, J. Identification of the optimal entry points at District Metered Areas and implementation of pressure management. *Urban Water Journal 9*, 6 (Dec. 2012), 365–384.

[58] GRAYMAN, W. M., MURRAY, R., AND SAVIC, D. A. Effects of Redesign of Water Systems for Security and Water Quality Factors. In *World Environmental and Water Resources Congress 2009* (Reston, VA, May 2009), no. 2009, American Society of Civil Engineers, pp. 1–11.

[59] HAJEBI, S., BARRETT, S., CLARKE, A., AND CLARKE, S. Multi-Agent Simulation to Support Water Distribution Network Partitioning. In *27th European Simulation and Modelling Conference - ESM'2013, Lancaster, UK* (Lancaster, United Kingdom, 2013), EUROSIS.

[60] HAJEBI, S., BARRETT, S., CLARKE, A., AND CLARKE, S. Graph Partitioning with Structural Constraints - Poster. In *The 25th International Workshop on Combinatorial Algorithms (IWOCA2014)* (Duluth, Minnesota, USA., 2014).

[61] HAJEBI, S., ROSHANI, E., CARDOZO, N., BARRETT, S., CLARKE, A., AND CLARKE, S. Water distribution network sectorisation using graph theory and many-objective optimisation. *Journal of Hydroinformatics 18*, 1 (2016), 77–95.

[62] HAJEBI, S., SONG, H., BARRETT, S., CLARKE, A., AND CLARKE, S. Towards a Reference Model for Water Smart Grid. *International Journal of Advances in Engineering, Science and Technology 2*, 3 (2013), 310–317.

[63] HAJEBI, S., TEMATE, S., BARRETT, S., CLARKE, A., AND CLARKE, S. Water Distribution Network Sectorisation Using Structural Graph Partitioning and Multi-objective Optimization. *Procedia Engineering 89* (2014), 1144–1151.

[64] HARTIGAN, J. A., AND WONG, M. A. Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics 28*, 1 (1979), 100.

[65] HERRERA, M., IZQUIERDO, J., PÉREZ-GARCÍA, R., AND MONTALVO, I. Multi-agent adaptive boosting on semi-supervised water supply clusters. *Advances in Engineering Software 50* (Aug. 2012), 131–136.

[66] IZQUIERDO, J., HERRERA, M., MONTALVO, I., AND PÉREZ-GARCÍA, R. Agent-based Division of Water Distribution Systems into District Metered Areas. In *ICSOFT (2)* (2009), pp. 83–90.

[67] JAIN, H., AND DEB, K. An Improved Adaptive Approach for Elitist Nondominated Sorting Genetic Algorithm for Many-Objective Optimization. In *Evolutionary Multi-Criterion Optimization SE - 25*, R. Purshouse, P. Fleming, C. Fonseca, S. Greco, and J. Shaw, Eds., vol. 7811 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 307–321.

[68] JAYARAM, N., AND SRINIVASAN, K. Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. *Water Resources Research 44*, 1 (2008).

[69] JUN, H., AND LOGANATHAN, G. V. Valve-Controlled Segments in Water Distribution Systems. *Journal of Water Resources Planning and Management 133*, 2 (Mar. 2007), 145–155.

[70] KAPELAN, Z. S., SAVIC, D. A., AND WALTERS, G. A. Multiobjective design of water distribution systems under uncertainty. *Water Resources Research 41*, 11 (Nov. 2005).

[71] KARYPIS, G., AND KUMAR, V. Multilevel k-way Partitioning Scheme for Irregular Graphs. *Journal of Parallel and Distributed Computing 48* (1998), 96–129.

[72] KIRKPATRICK, S. Optimization by simulated annealing: Quantitative studies. *Journal of statistical physics 34* (1984), 975–986.

[73] KLEIN, D., KAMVAR, S., AND MANNING, C. From instance-level constraints to space-level constraints: Making the most of prior knowledge in data clustering. Tech. rep., Stanford, 2002.

[74] LI, J., LIU, C.-C., AND SCHNEIDER, K. P. Controlled Partitioning of a Power Network Considering Real and Reactive Power Balance. *IEEE Transactions on Smart Grid 1*, 3 (Dec. 2010), 261–269.

[75] LIU, H.-L., GU, F., AND ZHANG, Q. Decomposition of a Multiobjective Optimization Problem Into a Number of Simple Multiobjective Subproblems. *IEEE Transactions on Evolutionary Computation 18*, 3 (June 2014), 450–455.

[76] MACDONALD, G., AND YATES, C. D. Dma design and implementation, a north american context. In *Leakage 2005, Specialized Conference Proceedings* (2005).

[77] MASSOUD AMIN, S., AND WOLLENBERG, B. Toward a smart grid: power delivery for the 21st century, 2005.

[78] METROPOLIS, N., ROSENBLUTH, A., ROSENBLUTH, M., TELLER, A., AND TELLER, E. Simulated annealing. *Journal of Chemical Physics 21* (1953), 1087–1092.

[79] MORRISON, J., TOOMS, S., AND ROGERS, D. District metered areas guidance notes. *International Water Association (IWA), Water Loss Task Force* (2007).

[80] MURRAY, R., GRAYMAN, W. M., SAVIC, D. A., FARMANI, R., BOXALL, J., AND MAKSIMOVIC, C. Effects of DMA redesign on water distribution system performance. In *Proceedings of the Tenth International Conference on Computing and Control in the Water Industry* (2009).

[81] NARDO, A., NATALE, M., GUIDA, M., AND MUSMARRA, D. Water Network Protection from Intentional Contamination by Sectorization. *Water Resources Management 27*, 6 (Sept. 2012), 1837–1850.

[82] NATIONAL RESEARCH COUNCIL. *Improving the Nation's Water Security: Opportunities for Research.* The National Academies Press, 2007.

[83] NEWMAN, M. E. J. The Structure and Function of Complex Networks. *SIAM Review 45*, 2 (jan 2003), 167–256.

[84] NEWMAN, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America 103*, 23 (2006), 8577–8582.

[85] NEWMAN, M. E. J., AND GIRVAN, M. Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 69* (2004).

[86] ORMSBEE, L. E., AND REDDY, S. L. Nonlinear Heuristic for Pump Operations. *Journal of Water Resources Planning and Management 121*, 4 (July 1995), 302–309.

[87] OSTFELD, A., UBER, J. G., SALOMONS, E., BERRY, J. W., HART, W. E., PHILLIPS, C. A., WATSON, J.-P., DORINI, G., JONKERGOUW, P., KAPELAN, Z., DI PIERRO, F., KHU, S.-T., SAVIC, D., ELIADES, D., POLYCARPOU, M., GHIMIRE, S. R., BARKDOLL, B. D., GUELI, R., HUANG, J. J., MCBEAN, E. A., JAMES, W., KRAUSE, A., LESKOVEC, J., ISOVITSCH, S., XU, J., GUESTRIN, C., VANBRIESEN, J., SMALL, M., FISCHBECK, P., PREIS, A., PROPATO, M., PILLER, O., TRACHTMAN, G. B., WU, Z. Y., AND WALSKI, T. The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms. *Journal of Water Resources Planning and Management 134*, 6 (Nov. 2008), 556–568.

[88] OSYCZKA, A. Multicriteria optimization for engineering design. *Design optimization 1* (1985), 193–227.

[89] PAAR, C., AND PELZL, J. *Understanding cryptography: a textbook for students and practitioners*. Springer Science & Business Media, 2009.

[90] PAPANIKAS, D. G., PANTAZIS, V., PAPAGIANNIDIS, P., BITZAS, M., AND PROTOPSALTIS, A. A System for the Engineering Design of Transmission and Distribution Pipe Networks. 1992, pp. 91–114.

[91] PARKINSON, A. R., BALLING, R., AND HEDENGREN, J. D. Optimization methods for engineering design. *Brigham Young University* (2013).

[92] PEARL, J. *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1984.

[93] PERELMAN, L., AND OSTFELD, A. Topological clustering for water distribution systems analysis. *Environmental Modelling & Software 26*, 7 (July 2011), 969–972.

[94] PERELMAN, L., AND OSTFELD, A. Water-Distribution Systems Simplifications through Clustering. *Journal of Water Resources Planning and Management 138*, 3 (May 2012), 218–229.

[95] PERELMAN, L. S., ALLEN, M., PREIS, A., IQBAL, M., AND WHITTLE, A. J. Flexible Reconfiguration of Existing Urban Water Infrastructure Systems. *Environmental Science & Technology 49*, 22 (nov 2015), 13378–13384.

[96] POHL, I. S. *Bi-Directional and Heuristic Search in Path Problems.* Ph.d. dissertation, Stanford University, 1969.

[97] PONS, P., AND LATAPY, M. Computing communities in large networks using random walks. *Computer and Information Sciences-ISCIS 3733* (2005), 284–293.

[98] PORCO, J. Municipal Water Distribution System Security Study: Recommendations for Science and Technology Investments. *Jour. AWWA 102*, 4 (2010), 30–32.

[99] PRASAD, T. D., AND PARK, N.-S. Multiobjective Genetic Algorithms for Design of Water Distribution Networks. *Journal of Water Resources Planning and Management 130*, 1 (Jan. 2004), 73–82.

[100] PROGRAMME, U. I. H. Plan for the International Hydrological Programme of UNESCO Phase VI. Tech. Rep. April 2012, 2012.

[101] RAO, S. S. *Engineering Optimization: Theory and Practice*, 4th ed. John Wiley & Sons, 2009.

[102] RECA, J., AND MARTÍNEZ, J. Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resources Research 42*, 5 (may 2006).

[103] ROSSMAN, L. A. *Epanet 2 users manual.* No. September. 2000.

[104] SAVIC, D., AND FERRARI, G. Design and performance of district metering areas in water distribution systems. In *16th Conference on Water Distribution System Analysis, WDSA 2014, Bari, Italy.* (2014).

[105] SCIBETTA, M., BOANO, F., REVELLI, R., AND RIDOLFI, L. Community detection as a tool for complex pipe network clustering. *EPL (Europhysics Letters) 103*, 4 (Aug. 2013), 48001.

[106] SEMPEWO, J., PATHIRANA, A., AND VAIRAVAMOORTHY, K. Spatial Analysis Tool for Development of Leakage Control Zones from the Analogy of Distributed Computing. In *Water Distribution Systems Analysis 2008* (Reston, VA, Apr. 2009), American Society of Civil Engineers, pp. 1–15.

[107] SEN, A., GHOSH, P., VITTAL, V., AND YANG, B. A new min-cut problem with application to electric power network partitioning. *European Transactions on Electrical Power 19*, 6 (Sept. 2009), 778–797.

[108] SNYMAN, J. A. *Practical Mathematical Optimization*, vol. 97 of *Applied Optimization*. Springer-Verlag, New York, 2005.

[109] SYCARA, K. P. Multiagent Systems. *AI magazine 19*, 2 (1998), 79–93.

[110] TALBI, E.-G. *Metaheuristics: From Design to Implementation.* Wiley Publishing, 2009.

[111] TARJAN, R. Depth-First Search and Linear Graph Algorithms. *SIAM Journal on Computing 1*, 2 (June 1972), 146–160.

[112] TARJAN, R. Depth-First Search and Linear Graph Algorithms. *SIAM Journal on Computing 1*, 2 (June 1972), 146–160.

[113] TODINI, E. Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water 2*, 2 (June 2000), 115–122.

[114] TZATCHKOV, V. G., ALCOCER-YAMANAKA, V. H., AND BOURGUETT ORTÍZ, V. Graph Theory Based Algorithms for Water Distribution Network Sectorization Projects. In *Water Distribution Systems Analysis Symposium 2006* (Reston, VA, Mar. 2006), American Society of Civil Engineers, pp. 1–15.

[115] UK WATER AUTHORITIES ASSOCIATION. Report 26 Leakage Control Policy & Practice. Tech. rep., 1980.

[116] UNITED NATIONS UNIVERSITY. *Water Security and the Global Water Agenda.* 2013.

[117] U.S. ARMY CORPS OF ENGINEERS. Water Supply: Water Distribution. 2004.

[118] WAGSTAFF, K., CARDIE, C., ROGERS, S., AND SCHRÖDL, S. Constrained k-means clustering with background knowledge. In *Eighteenth International Conference on Machine Learning/* (2001), pp. 577–584.

[119] WALSKI, T., CHASE, D., SAVIC, D., AND GRAYMAN, W. Advanced water distribution modeling and management.

[120] WANG, Q., GUIDOLIN, M., SAVIC, D., AND KAPELAN, Z. Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front. *Journal of Water Resources Planning and Management 141*, 3 (Mar. 2015), 04014060.

[121] WRIGHT, R., STOIANOV, I., AND PARPAS, P. Dynamic Topology in Water Distribution Networks. *Procedia Engineering 70* (2014), 1735–1744.

[122] WRIGHT, R., STOIANOV, I., PARPAS, P., HENDERSON, K., AND KING, J. Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics 16*, 6 (nov 2014), 1280.

[123] YANNAKAKIS, M. Node-and edge-deletion NP-complete problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing - STOC '78* (New York, New York, USA, 1978), ACM Press, pp. 253–264.

[124] YUAN, Y., XU, H., AND WANG, B. An improved NSGA-III procedure for evolutionary many-objective optimization. In *Proceedings of the 2014 conference on Genetic and evolutionary computation - GECCO '14* (New York, New York, USA, 2014), ACM Press, pp. 661–668.

[125] YUSOF, S., ROGERS, G., AND ALDEN, R. Slow coherency based network partitioning including load buses. *IEEE Transactions on Power Systems 8*, 3 (1993), 1375–1382.

[126] ZHANG, Q., AND LI, H. MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition. *IEEE Transactions on Evolutionary Computation 11*, 6 (Dec. 2007), 712–731.