Mitigation of cascade failures in complex networks: Theory and Application

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Ryan Ghanbari

BSC (Electrical Engineering), Iran University of Science and Technology
MSC (Electronics Engineering), University of Zanjan

School of Engineering
College of Science, Engineering and Health
RMIT University

October 2019
I would like to dedicate this thesis to my lovely wife, Rose and my adorable daughters, Ronika and Ariana
Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed. I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Ryan Ghanbari
October 2019
Acknowledgements

I hereby appreciate the support of people who made it possible for me to finish this thesis. It is an outstanding lifetime milestone, and I am so excited to finally completed my research PhD program.

First and foremost, I wish to acknowledge my gratitude to my intimate family for their continuous support and encouragement during my PhD program. Mainly, my lovely wife and cute daughter whose significant backing, made me keep on track and finish my work promptly. I also thank my parents and siblings for giving me love and inspiration throughout this work.

I would like to thank my senior supervisor Dr Mahdi Jalili, who gave me this opportunity to work as a PhD candidate at RMIT, Melbourne. Without your advice and drive, I would never have come this far in this thesis. I value your patience, trust, friendly interaction, and professional discussions.

I also would like to thank Prof. Xinghuo Yu, my associate supervisor for his endless support and invaluable suggestions shared with me on my research. Likewise, I appreciate all your endeavour for inviting high-quality scholars to the office so that they can share with us their experience in their area of expertise.

I am also thankful to independent panel members, Prof. Henry Wu and Dr Arash Vahidnia, for examining my progress of three milestones, as well as for their feedback on my study.

Furthermore, I would also like to appreciate Dr Peter Sokolowski for his great attention to details and always making constructive comments to improve the quality of my work. Also, I would thank Dr Paul Cuffe from University College Dublin for presenting and discussing his work on cascade failures in power systems which shed light on the rest of my journey throughout my PhD. I would like to thank Dr Lasantha Meegahapola for sharing with me the Neo Express licensed software to extract the Australian national electricity structure and information.

Also, I wish to thank RMIT University and staff for continued help and support, especially in all administrative matters.
Finally, I would like to thank the Australian Research Council for supporting this research through project No. DE140100620.
Abstract

Complex networks such as transportation networks, the Internet, and electrical power grids are fundamental parts of modern life, and their robustness under any attack or fault has always been a concern. Failure and intentional removal of components in complex networks might affect the flow of information and change balance of flows in the network. This phenomenon may require load redistribution all over the network. Component overloaded can act as a trigger for a chain of overload failures. This overload, could, for example, increase the amount of information a router must transmit and ultimately make internet congestion.

One of the major applications of complex network theory is to study power systems. Power systems are the most complex human-made infrastructures, and almost every individual's life is dependent on electrical energy and resilient functioning of power systems. Recently, there have been many reports about massive power outages leaving vast areas without power that sometimes takes a few days to have the power back. One of the most critical areas in the power system is the root cause analysis of such catastrophes and trying to resolve them. From an electrical engineering point of view, these power outages occur following an initial failure due to problems, such as generators tripping, transformers overheating, faulty power generation units, damage to the transmission system, substations or distribution systems, or overloading of the power system. A faulty protection relay or malicious attack to control centres can also trigger it. In any of these cases, the failed component will be out of service immediately and to keep the robust power delivery to all customers, their loads should be redistributed across the power system, and henceforth some of them might become overloaded as well, and accordingly get out of service. This chain of failures can be propagated all over the system and lead to a catastrophic blackout. This thesis conducts a full study on how to mitigate cascade failures in complex networks.

First, cascade depth is applied to quantify nodes criticality for cascade failures. Then, a wide range of node centrality parameters is considered to find out the relationship between the
node vitality and these centralities. To discover the structure of cascade propagation in complex networks, the edge geodesic distance is considered for computing the structural distance between two arbitrary edges in the network. Then, starting with the single edge removal events, the route that cascade tends to spread is studied. In the next step, the impact of two or three concurrent edge removals on the way the cascade spreads are examined. Besides, the power system vulnerability is studied using the maximum flow algorithm based on Ford-Fulkerson method and critical capacity parameters are identified. A synthetic model with the same properties as a real power system is generated and examined. For a power line, to be overloaded, a new method is developed to overpass across the network and shortlist the busbars for load reduction. Next, a novel sensitivity method is formulated based on AC load flow analysis to rank the loads according to their effect on the lines power flow.
Table of Contents

1 INTRODUCTION ........................................................................................................................................... 1

1.1 Preface .................................................................................................................................................. 1
1.2 Complex Networks ................................................................................................................................. 1
1.3 Power System Overview ......................................................................................................................... 2
1.4 Motivation ............................................................................................................................................. 4
1.5 Thesis Objectives and Contributions ..................................................................................................... 5
  1.5.1 Contributions ................................................................................................................................ 5
1.6 Thesis Structure .................................................................................................................................... 6
1.7 Summary ............................................................................................................................................. 8

2 COMPLEX NETWORK MODELS ................................................................................................................. 9

2.1 Introduction ....................................................................................................................................... 9
  2.1.1 Synthetic Networks ......................................................................................................................... 11
Erdos-Renyi Model ..................................................................................................................................... 11
Watts-Strogatz Model ................................................................................................................................. 12
Barabasi-Albert Model ............................................................................................................................... 13
2.2 Structural Centrality Measures ........................................................................................................... 15
  2.2.1 Neighbourhood-Based Centralities ............................................................................................. 16
Degree Centrality ..................................................................................................................................... 16
Local Rank .................................................................................................................................................. 18
Clustering Coefficient ................................................................................................................................. 19
Coreness Centrality ..................................................................................................................................... 19
2.2.2 Path-Based Centralities ........................................................................................................... 20
   Shortest Path Length ....................................................................................................................... 20
   Average Path Length ....................................................................................................................... 21
   Closeness Centrality ....................................................................................................................... 22
   Node and Edge Betweenness .......................................................................................................... 22
   Information Centrality ................................................................................................................... 23
   Eigenvalue Centrality ..................................................................................................................... 23
   Page Rank ....................................................................................................................................... 24
2.2.3 Flow-Based Centrality (Ford-Fulkerson Algorithm) ................................................................. 25
2.3 Complex Network Models of Power Grids ................................................................................... 26
   2.3.1 Minimum-distance graph ....................................................................................................... 27
   2.3.2 Synthetic Spatial Power Network .......................................................................................... 28
   2.3.3 AC/DC Model ......................................................................................................................... 29
       AC Model .................................................................................................................................... 30
       DC Model .................................................................................................................................... 31
2.4 Benchmark Networks .................................................................................................................. 33
2.5 Conclusion .................................................................................................................................... 37

3 CASCADE FAILURES ....................................................................................................................... 39
   3.1 Introduction ................................................................................................................................. 39
   3.2 The Process of Cascade Failures ................................................................................................ 40
   3.3 Cascade Failure Models in Complex Networks ........................................................................ 42
       3.3.1 Structural Models ................................................................................................................ 43
           Static Model .............................................................................................................................. 43
           Motter-Lai Model ...................................................................................................................... 44
           Effective Efficiency Model ....................................................................................................... 45
       3.3.2 Component Cascading Failures Dependent Model ............................................................... 46
           CASCADE Model ...................................................................................................................... 47
           Branching Process Model ........................................................................................................ 48
   3.4 Dynamic Cascade Failure Models in Power Systems ............................................................... 49
       ORNL-PSerc-Alaska (OPA) Model ............................................................................................... 50
       Hidden Failure Model .................................................................................................................. 50
       Manchester Model ...................................................................................................................... 51
   3.5 Conclusion ................................................................................................................................... 52
# Table of contents

## 4 Cascade Failures and Centrality Measures .............................................. 54

4.1 Introduction ......................................................................................... 54
4.2 Related Works ..................................................................................... 54
4.3 Motivation ............................................................................................. 56
4.4 Cascade Depth ..................................................................................... 57
4.5 Simulation Setup .................................................................................. 59
4.6 Results and Discussion ......................................................................... 60
  4.6.1 Synthetic networks ........................................................................... 61
  4.6.2 Real networks .................................................................................. 65
4.7 Conclusion ............................................................................................ 68

## 5 Cascade Propagation Structure ............................................................... 70

5.1 Introduction .......................................................................................... 70
5.2 Structure of Cascade Propagation ......................................................... 70
  5.2.1 An Illustrative Example .................................................................... 73
5.3 Simulation Results ................................................................................ 76
  5.3.1 Single Edge Removal ....................................................................... 78
  5.3.2 Concurrent Edge Removal ............................................................... 81
  5.3.3 Cascade Propagation Route Based on Ford-Fulkerson Method .......... 88
5.4 Conclusion ........................................................................................... 90

## 6 Vulnerability and Load Reduction in Power Networks ......................... 92

6.1 Introduction .......................................................................................... 92
6.2 Power System Vulnerability .................................................................. 93
  6.2.1 Power System Vulnerability Analysis ........................................... 94
  6.2.2 Simulation Results .......................................................................... 96
6.3 Load Reduction in Power Networks ....................................................... 99
  6.3.1 Sensitivity Analysis and Load Reduction ....................................... 99
  6.3.2 Sensitivity Parameters ................................................................... 100
  6.3.3 Choosing the Loads to Be Reduced ............................................. 103
  6.3.4 Simulation Results ......................................................................... 103
  6.3.5 Kendall’s Tau Correlation ............................................................ 105
6.4 Conclusion ........................................................................................... 108
7 CONCLUSION AND OUTLOOK

7.1 Introduction

7.2 Research Findings Summary

7.3 Future Research Directions

7.3.1 Considering Different Node Types

7.3.2 Studying Dynamical Behaviour Using Cloud Computing

7.3.3 Optimisation of Reduced Loads

References
List of Tables

Table. 2. 1 Comparison of structural properties of IEEE 118 bus and synthetic model .................... 29
Table. 2. 2 Structural features of the studied real networks .................................................................. 33
Table. 2. 3 Demographic information of the real networks under consideration................................. 34

Table. 4. 1 Centrality indices and cascade depth for nodes 6, 21, and 25 for T=1.0, 1.5, and 1.9 .......... 56

Table. 6. 1 Summary of the results from DIgSILENT and Sensitivity methods’ ranking .................. 105
Table. 6. 2 Kendall’s Tau correlation between the DIgSILENT and Sensitivity methods’ rankings 106
List of Figures

Fig. 2. 1 A simple form of complex network consisting of nodes and edges .............................................. 9
Fig. 2. 2 The adjacency matrix representation of a complex network with five nodes and directed edges ........................................................................................................................................... 10
Fig. 2. 3 The Konigsberg problem graph .................................................................................................. 10
Fig. 2. 4 The Erdos-Renyi network with N=9 and P=0.5 ........................................................................ 12
Fig. 2. 5 Comparing the structural properties of a random network and a WS small-world network (N=14, K=4) ........................................................................................................................................ 13
Fig. 2. 6 BA network generated by preferential attachment method (m_0=4, m=4, N=18) ................. 15
Fig. 2. 7 Node degree distributions (a) power-law distribution, (b) Poisson distribution .............. 17
Fig. 2. 8 Ford-Fulkerson based centrality index in a multiple source and sink network (2 sources and 3 sinks) ........................................................................................................................................... 26
Fig. 2. 9 synthetic spatial network constructed by the model introduced in section 2.3.2 ................. 29
Fig. 2. 10 IEEE 118 bus system used in this work as a benchmark network. This figure has been adopted from (Lu, 2008). ........................................................................................................................................ 35
Fig. 2. 11 IEEE 57 test case consisting of 57 buses, 80 lines and 42 loads ............................................. 35
Fig. 2. 12 Degree histogram of the networks considered in this work ..................................................... 36

Fig. 3. 1 Cascade failure process in IEEE 30 Bus network .................................................................... 41

Fig. 4. 1 Impact of node criticality on the depth of cascading failure in IEEE 30 Bus network ......... 57
Fig. 4. 2 Impact of node criticality on the depth of cascading failure in IEEE 30 Bus network ......... 58
Fig. 4. 3 Nodes cascade depth and centralities correlation in Erdos-Renyi networks (N = 1000) ...... 62
Fig. 4. 4 Nodes cascade depth and centralities correlation in watts-Strogatz networks (N = 1000)..... 63
Fig. 4. 5 Nodes cascade depth and centralities correlation in scale-free networks (N = 1000).......... 64
Fig. 4. 6 Nodes cascade depth and centralities correlation in Modular networks (N = 1000) .......... 65
Fig. 4. 7 Nodes cascade depth and centralities correlation in Power networks .................. 66
Fig. 4. 8 Nodes cascade depth and centralities correlation in real networks .................................. 67

Fig. 5. 1 Cascade propagation concept for IEEE 118 bus in the case of single edge removal .......... 74
Fig. 5. 2 Cascade propagation route for Synthetic networks, triggered by single edge removal ...... 77
Fig. 5. 3 Cascade propagation route for power networks triggered by single edge removal .......... 78
Fig. 5. 4 Cascade propagation route for Biological networks triggered by single edge removal ... 79
Fig. 5. 5 Cascade propagation route for real networks triggered by single edge removal ............ 80
Fig. 5. 6 Cascade propagation route in synthetic networks (ER, WS, SF) in concurrent edge failure cases. .................................................................................................................................. 82
Fig. 5. 7 Cascade propagation route in Australian, UK, Spanish, and French power grids in concurrent edge failure cases .................................................................................................................................. 83
Fig. 5. 8 Cascade propagation route in biological networks in concurrent edge failure cases ....... 84
Fig. 5. 9 Cascade propagation route in real (infrastructure) networks in concurrent edge failure cases. .................................................................................................................................. 86
Fig. 5. 10 Cascade propagation route for IEEE-57, IEEE-118, and IEEE-300 triggered by single edge removal.......................................................................................................................... 87
Fig. 5. 11 Cascade propagation caused by (a) 2 and (b) 3 concurrent edge failures in IEEE-57 .... 89
Fig. 5. 12 Cascade propagation caused by (a) 2 and (b) 3 concurrent edge failures in IEEE-118 .... 89
Fig. 5. 13 Cascade propagation caused by (a) 2 and (b) 3 concurrent edge failures in IEEE-300 .... 90

Fig. 6. 1 Influence of single edge random failure on cascade failure in IEEE-118 ...................... 95
Fig. 6. 2 Targeted attack to the edge with the highest normalized centrality in IEEE-118 .......... 96
Fig. 6. 3 Impact of random attack (fault) to one edge of the synthetic networks ....................... 97
Fig. 6. 4 Targeted attack to the edge with the highest normalised centrality in the synthetic network. 98
Fig. 6. 5 Line l connecting buses v_i and v_j .................................................................................. 100
Fig. 6. 6 Line 15\rightarrow 14 current [pu] for a 10 MW increase in the demand of 8 different loads at time t=10s ............................................................................................................................................. 107
Nomenclature

Symbols/Acronyms

\( G \)  Representative graph for complex network
\( N \)  Number of nodes in the network
\( E \)  Number of edges in the network
\( a_{ij} \)  The entry in the i-th row and j-th column of a matrix A
\( k_i \)  Node degree
\( P \)  The probability measure
\( d_{ij} \)  Shortest path length
\( L \)  Average shortest path length
\( <k> \)  Average degree
\( E(G) \)  Network Efficiency
\( S_N \)  Normalized avalanche size
\( C_i \)  Capacity of the node \( v_i \)
\( r \)  Assortativity parameter
\( T \)  Tolerance parameter
\( L_{ij} \)  Load of the link connecting nodes \( i \) and \( j \)
\( \Gamma_i \)  The set of immediate neighbours of \( v_i \)
\( R(k) \)  The number of immediate and the next immediate neighbours of \( v_k \)
\( TH_i \)  A set of top-\( h \) neighbours of node \( v_i \), whose degree is at least \( h \)
\( Cl_i \)  The closeness centrality of node \( v_i \)
\( C_{ij} \)  Capacity of the link connecting nodes \( v_i \) and \( v_j \)
\( e_{ij} \)  Efficiency of edge connecting nodes \( v_i \) and \( v_j \)
\( C_{i-n} \)  Normalized centrality of node \( v_i \)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{pq}$</td>
<td>The number of the shortest path from node $v_p$ to node $v_q$</td>
</tr>
<tr>
<td>$\psi_{pq}(e_{ij})$</td>
<td>The number of shortest paths from node $v_p$ to node $v_q$ making use of $e_{ij}$</td>
</tr>
<tr>
<td>$f_{uv}^{\max}$</td>
<td>Maximum flow from the source $u$ to the sink $v$</td>
</tr>
<tr>
<td>$f_{ij}$</td>
<td>The proportion of the flow $f_{uv}^{\max}$ passing the edge connecting nodes $v_i$ and $v_j$</td>
</tr>
<tr>
<td>$l_{ip}$</td>
<td>The length of the shortest path between nodes $v_i$ and $v_p$</td>
</tr>
<tr>
<td>$INF_i$</td>
<td>The information index of node $v_i$</td>
</tr>
<tr>
<td>$PR_i(t)$</td>
<td>$PR$ value of node $v_i$ at $t$ step</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Cascade propagation route</td>
</tr>
<tr>
<td>$e_0$</td>
<td>An initial single edge failure</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>The list of removed edges</td>
</tr>
<tr>
<td>$e_h$</td>
<td>The highest loaded edge</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>A list of EGDs between two consecutive edges in $\gamma$</td>
</tr>
<tr>
<td>$t$</td>
<td>Cascade order</td>
</tr>
<tr>
<td>$t_h$</td>
<td>The highest cascade order</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Cascade route accumulator</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>Average edge geodesic distance between initially failed edges</td>
</tr>
<tr>
<td>$\lambda_h$</td>
<td>The highest $\lambda_0$ which could be equal to power system graph diameter</td>
</tr>
<tr>
<td>$I_{ij}^*$</td>
<td>The conjugate current in the line connecting busbars $v_i$ and $v_j$</td>
</tr>
<tr>
<td>$S_{ij}$</td>
<td>The apparent power in the line connecting busbars $v_i$ and $v_j$</td>
</tr>
<tr>
<td>$\Delta P_{ij}$</td>
<td>Active power change in the line connecting busbars $v_i$ and $v_j$</td>
</tr>
<tr>
<td>$\Delta Q_{ij}$</td>
<td>Reactive power change in the line connecting busbars $v_i$ and $v_j$</td>
</tr>
<tr>
<td>$\frac{\partial P_{ij}}{\partial P_l}$</td>
<td>Line $i$-$j$ active power, $P_{ij}$, sensitivity to active power change in load $l$:$P_l$</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>EGD</em></td>
<td>Edge Geodesic Distance</td>
</tr>
<tr>
<td><em>ER</em></td>
<td>Erdos-Renyi Random Graph</td>
</tr>
<tr>
<td><em>WS</em></td>
<td>Watts-Strogatz Random Graph</td>
</tr>
<tr>
<td><em>BA</em></td>
<td>Barabasi-Albert Random Graph</td>
</tr>
<tr>
<td><em>OPA</em></td>
<td>ORNL-PSerc-Alska Blackout Model</td>
</tr>
<tr>
<td><em>LR</em></td>
<td>Local Rank</td>
</tr>
<tr>
<td><em>al-index</em></td>
<td>Average Lobby Index</td>
</tr>
</tbody>
</table>
List of Publications

Journal Papers


Conference Papers


Chapter 1

INTRODUCTION

1.1 Preface

The occurrence of cascades of failures leading to whole system collapse like large-scale blackouts in power grids has motivated scholars to develop effective methods to model and analyse the dynamics of this phenomenon. One of the most effective and easy to apply mathematical representatives of real-world systems is a complex network. It consists of number of connected elements, which communicate and interact through the links; where elements and links can be abstracted and modelled as nodes and edges of the network, respectively. It has been shown that many real-world networks have several common properties such as scale-free degree distribution, small-worldness, densification and community structure. In this chapter, the background of cascade failures in complex network and its application in power systems is briefly overviewed. Also, the motivation of this work and the scope of the research are explained. Moreover, the objectives and the contribution of this work are listed and discussed. Finally, the structure of the thesis and significant context of other chapters are briefly illustrated.

1.2 Complex Networks

Complex networks are pervasively everywhere in many fields ranging from engineering to biology, physics, and sociology (Strogatz, 2001, Albert and Barabasi, 2002). Resiliency against cascade failure is a striking property that many real-world networks possess. Human-
made networks should also have high levels of resiliency in order to function correctly. When it comes to an error in network components or targeted attacks to its vital elements, it has been shown that the structure of complex networks plays a critical role in their resilience against such events (Watts, 2002, Moreno et al., 2002). For instance, when degree-based attacks are considered, scale-free networks are vulnerable, despite being resilient against random errors (Albert et al., 2000).

In studying complex networks, we investigate the statistical properties of their nodes with their interconnections. Numerous network properties and performance measures can be determined using network topologies. We need to investigate the networks statistical properties in detail to better comprehend the networks’ behaviour.

1.3 Power System Overview

Power systems are of the most important and probably complicated infrastructures humans have ever made. The existence of a variety of components with diverse operational modes and dynamics, along with its ever-growing structure, make it difficult to study it in detail. So far, many scholars have placed it under scrutiny from their point of view using simplified models. Undoubtedly, in this modern era, all of us are dependent on the proper functioning of the power system, and their continuous power delivery is crucial for our lives and wellbeing. Therefore, power outage at any level has a direct effect on countries’ industrial and commercial services (Bompard et al., 2011). A power grid is vulnerable to natural catastrophes such as earthquakes and Tsunamis as well as to deliberate attacks such as an Electromagnetic Pulse attack or cyber-attack to control centres (Bernstein, 2012). A power grid can be modelled as a complex network with nodes representing the generators, transformers and loads and links describing the transmission lines (Chassin and Posse, 2005, Rosato et al., 2007, Rosa et al., 2007, Solé et al., 2008b, Saniie Monfared et al., 2014, Wang et al., 2011). The edges in this model could be weighed by their loads, impedance or admittance to describe the power system parameters. A failure in one of the network components (nodes and edges) may cause other
components to go beyond their capacity, and consequently fail. Such failures can spread through the network and result in a cascade failure process (Buldyrev et al., 2010, Crucitti et al., 2004b, Ren and Dobson, 2008).

Mirzasoleiman et al. (Mirzasoleiman et al., 2011) investigated the tolerance of cascading failures in weighted networks using three weighting strategies. Saniee Monfared et al. (Saniee Monfared et al., 2014) studied the structural properties of the ultra-high voltage power transmission network of Iran. According to their findings, the Iranian power grid is a small-world network with exponential degree distribution. They also studied the influence of cascading failures on the largest connected component of the network and found it vulnerable. Cuffe (Cuffe, 2017) investigated the impact of a complicated intentional attack against transmission lines of a classic electrical power system.

Despite all the endeavours that scientists and engineers put forward to make the power system work robustly and avoid the faults and power cutoffs, failures take place, and unfortunately, sometimes they happen intentionally by criminals. Ukraine 2015 large-scale blackout caused by cyber-attacks to the power systems (Liang et al., 2017).

One resolution to deter the catastrophic situations caused by cascade is load shedding which has been a classical and a quite reliable mitigation strategy (Laghari et al., 2015, Reddy et al., 2014, Gu et al., 2014, Ketabi and Fini, 2017, Ketabi and Fini, 2015, Rudez and Mihalic, 2016, Shekari et al., 2016, Dong et al., 2017, Amraee and Saberi, 2017). However, any load curtailment in power systems might cause customer discomfort, and thus such a strategy might not be the ideal solution in some power systems. In this thesis, we focus on two novel methods: The First method is to make the power system resilient by setting the edges’ capacity to their critical value. Finding this value for edges at each attack scenario and setting the parameter higher than that results in making the networks resilient against failure. An alternative solution is to predict the edge(s) which might fail at the next step by tracing the cascade propagation route. These two methods will be discussed in detail in the coming chapters.
1.4 Motivation

In case the cascade failure is triggered in a power system, there are some classical methods to deter it from leading to a blackout such as islanding or load shedding. These methods, unfortunately, always leave a segment of power system unserved. As a result, they become a big challenge for power systems since they cause customers discomfort, and, in some cases, the power system operators are penalised by their sensitive customers. There are some gaps in the literature that this thesis targets to fill.

First, it is worth finding vital nodes in complex networks related to their centrality measures which can be used to mitigate cascade failures in complex networks by putting more protection on critical components and also removing nodes ranked high in terms of those centrality measures showing negative correlations with the cascade depth.

Second, when a single transmission line fails in the power system due to an overload or even a protection relay malfunction, if it could trigger a cascade of failures, the pathway that this failure is propagated is vague and there is no information and analysis as a fast indication for network planners and operators to find out which component is going to fail next; therefore, they can put more protection or preventive measures to deter the cascade from spreading.

Third, it would be beneficial to study the potential vulnerability of power systems in the design and planning stage and find the worst failure cases which can draw the power systems to a total blackout state. It can be helpful to reconsider the power lines carrying power using critical capacity concept.

Fourth, when a power line connecting two substations is working close to its loading limits, reducing the overall load it is directly feeding would be a good measure to lift some pressure from that line. But there is no specific method in the literature to find and target these loads to be reduced, such that the loading level in that particular line does not exceed its limitations.
1.5 Thesis Objectives and Contributions

This thesis is mainly focused on cascade failures in complex networks which aims to discover the root cause of this phenomenon and develop the methods to deter it from occurring. The academic objective of this work is to create a context to analyse complex networks in terms of potential cascade failures and recommend the preventive methods to control its propagation. To be specific, there are four distinct goals for this thesis:

1. **Vital node identification (Chapter 4)**: Conducting a comprehensive analysis to discover the relationship between the centrality parameters of nodes in complex networks and the severity of the cascade by removing the nodes with those properties.

2. **Cascade propagation structure (Chapter 5)**: Determining the cascade pathway in complex networks through an extensive analysis of the cascade propagation route following a single or concurrent edge removals in complex networks and differentiate the cascade trajectory orientations between them.

3. **Power system vulnerability (Chapter 6)**: Developing an approach to analysing the structural vulnerability of power systems when removal of the highest loaded line initiates the cascade. Also, this method determines the critical capacity parameter for power systems which setting the corresponding parameter higher than that, makes the power system resilient against cascade failures.

4. **Selective load reduction (Chapter 6)**: Designing and developing a framework to analyse and determine a set of candidate busbars and rank them to reduce the loading in a specific power line to prevent it from failure and triggering the cascade.

1.5.1 Contributions

Here is the summary of the contributions of this research work:
1. Vital node identification (Chapter 4):
   a. Calculating the correlation between the cascade depth of nodes and their centrality parameters.
   b. Applying the concept cascade depth of a node to study the criticality of a node in a complex network.

2. Cascade propagation structure (Chapter 5):
   a. Introducing the Edge Geodesic Distance parameter to study the cascade propagation route.
   b. Developing an algorithm to compute the cascade propagation route when the cascade is triggered by a single edge failure using the edge geodesic distance.
   c. Developing an algorithm to compute the cascade propagation pathway for the case that cascade is triggered by concurrent edge removal.

3. Power system vulnerability (Chapter 6):
   a) Introducing the critical capacity parameter for power systems and applying it for a real power system.
   b) Developing a synthetic model for the real power system to validate the proposed method to study the power system vulnerability.

4. Selective load reduction (Chapter 6):
   a. Developing A novel traverse method to find the best candidate set of loads in the power system to rank for each transmission line.
   b. Establishing a new sensitivity method to rank the loads based on inverse Jacobian matrix.

1.6 Thesis Structure

This thesis is organized into seven chapters. First, mathematical tools based on complex networks are introduced to model real networks such as power systems. Then, structural
centrality measures, which are used to quantify the importance of components or distinguish the
different network topologies, are defined. The most used structural centrality measures,
neighbourhood-, path-, and flow-based are categorised and explained. Also, in chapter two, the
complex network models of power systems consisting of minimum distance, synthetic spatial
graphs, AC/DC models, are reviewed. This chapter is wrapped up by summarising all
benchmark networks which are going to be used in simulations in this thesis.

In chapter three, initially, the process of the cascade in complex networks is reviewed. Then,
different cascade failure models based on complex networks including structural and
component cascade dependent models and their analysis methods are discussed. Next, the
dynamic cascade failure models for power grids including OPA, Hidden Failure and Manchester
models are inspected.

In a complex network model of real systems represented by nodes and edges; different
nodes have a distinct impact on overall functionality and resiliency of networks against failures.
To study this, in chapter four, several synthetic networks and real networks are put under
scrutiny, and a metric is introduced to measure the cascade depth of a node and hence identify
the most vulnerable nodes which need solid protection measures to deter them from failure.

Following that, in chapter five, we find the trajectory of cascade propagation in complex
networks when it is initialized by single edge removal or concurrent removal of a few edges.
The difference in the cascade spreading trends between these two cases will be pointed out to
be applied in controlling the propagation of cascade failures.

The first section of chapter six is dedicated to studying the vulnerability of power systems
against cascade failures. The very well-known Ford Fulkerson model based on maximum flow
theory is applied to assign the loads to all power lines in a real power system as well as a
synthetic spatial power network consisting different types of nodes for representing generators,
loads and intermediate busbars. In second part of this chapter, to mitigate the cascade failure
triggered by a nearly overloaded line failure, a strategy is proposed that is more efficient than
classical load shedding algorithms. The proposed ‘selective load reduction’ is meant to reduce
the load only in a small portion of the network based on the lines’ flow sensitivity to various loads.

Finally, in chapter seven, all findings of this thesis are summarised, and future research directions are outlined.

1.7 Summary

In this chapter, the background of cascade failures in complex networks was introduced. Also, the motivation of this work, as well as the scope of research was explained. Moreover, the objectives and the contribution of this work were listed and discussed. Finally, the structure of the thesis and the significant context of other chapters were briefly illustrated.
Chapter 2

COMPLEX NETWORK MODELS

2.1 Introduction

Networks are ubiquitous, and we encounter a large number of them whenever any communication conducted, or information interchanged. Generally, a complex network consists of a collection of nodes (vertices) and a selection of edges (links) connecting these nodes. In its simplest case, the nodes in the network are all the same, but more complicated models can also be considered. Sometimes the edges can be weighted by the links’ physical parameters. A simple form of a complex network is shown in Fig. 2.1.

Fig. 2.1 A simple form of complex network consisting of nodes and edges
The simplest mathematical model for a complex network is an adjacency matrix. In Fig. 2.2, an example of an adjacency matrix for a complex network with five nodes and directed edges is displayed.

Complex networks represented by graphs, such as the Internet (Gan et al., 2014), World Wide Web (Fujimura et al., 2005), engineering (Saleh et al., 2018, Saleh et al., 2017), social (Hai-bo et al., 2008) and biological networks (Habibi et al., 2014) have been broadly investigated recently, and this area of study is developed rapidly. Historically, in 1736 when Leonhard Euler, a Swiss mathematician, solved the Konigsberg Bridges problem, the graph theory was introduced. At that time, people in Konigsberg tried to walk around the city by crossing each of its seven bridges exactly once and return to the starting point. After a while,
they came up to the conclusion that such a path does not exist. Euler published his solution where he proved that it was not possible to find such a path. First, he labelled four lands as A, B, C, and D and the seven bridges a, b, c, d, e, f, and g as shown in Fig. 2.3.

Euler represented the walking trip as a sequence of the land letters A, B, C, and D. There is a bridge between neighbouring letters. He noted that land A has five bridges crossing and the others have three bridges. Finally, using the frequency of land letters appear in the sequence of land letters, he finished the proof that there is no solution for the Konigsberg Bridges problem.

2.1.1 Synthetic Networks

When a hypothesis is to be tested, we need appropriate real-world networks for applying the algorithms or methods on them; however, in many cases, there are not such well-fitted networks. Hence, we need to create synthetic networks to construct model networks with similar properties to real ones. In this section, the most frequently studied synthetic complex networks are listed.

Erdos-Renyi Model

The random network theory began with one specific model, initially studied by two Hungarian mathematicians, Paul Erdős and Alfred Rényi (1959). Now the model is called the Erdős-Rényi Random Graph, which is an abstract mathematical model for creating graphs. Erdos-Renyi random networks are constructed by creating a link between any two nodes with probability $P$. It can be shown that the degree distribution in ER model follows Poisson distribution. For a long time, most of the real network studies were considered to be ER random networks. An ER network with $N=9$ and $P=0.5$ is shown in Fig. 2.4.
The ER graphs lack two remarkable properties of many real-world networks:

1. Since there is a constant, and independent probability of connection between each pair of nodes, the clustering coefficient in ER graphs is low.

2. The hub nodes are less likely to be formed in ER graphs, mainly because their degree distribution converges to a Poisson distribution rather than a power law which is the case in plenty of many real-world networks.

Hence, the other types of network models were introduced by other scholars, accounting for mimicking such properties of real networks.

**Watts-Strogatz Model**

The Watts-Strogatz model is a simple model to cope with the first limitation above. It comes up with higher clustering while maintaining the average path lengths of the ER model. It generally interpolates between an ER graph and a ring network. As a result, the model can address "small-world" phenomenon which is observed in a variety of real networks such as food webs, power systems, neural networks of the brain, social influence networks and movie actors’ network. The algorithm to generate a WS network is as follows. Given the number of nodes $N$, an even integer $K$ as the mean degree, and a probability parameter $p$, the algorithm generates an undirected graph consisting of $N$ nodes and $\frac{NK}{2}$ edges using the following steps:
2.2 Structural Centrality Measures

![Diagram showing structural properties of a random network and a WS small-world network](image)

Fig. 2.5 Comparing the structural properties of a random network and a WS small-world network (N = 14 AND K = 4). (a) Random graph: Average degree = 2.71, Average shortest path length = 2.52, Clustering coefficient = 0.05; (b) WS small-world network: Average degree= 4, Average shortest path length = 1.95, Clustering coefficient = 0.37.

1. Generate a regular ring network, a graph with N nodes each connected to \( \frac{K}{2} \) neighbours from each side.
2. For every node \( v_i \) remove the edges \( e_{ij} \) with \( i<j \) and rewire it with probability \( p \) to make the edge \( e_{ik} \) where \( k \) is selected using uniform probability among all possible values avoiding self-loops and duplicated links. In Fig. 2.5. The properties of a small world network are compared to a random network.

**Barabasi-Albert Model**

Two above models result in networks with almost homogeneous degree distribution. In a scale-free network, the degree distribution follows a power law, i.e., the probability of having nodes with degree \( k \), \( P(k) \), has a power-law relationship with the degree, i.e. \( P(k) = k^{-\gamma} \) where \( \gamma \) is the power-law exponent that is typically \( 2 < \gamma < 3 \) for many real systems. However, it has been shown that many real-world networks have heavy-tailed degree distribution, for which many nodes have small degrees, while few have much larger degrees. Such networks are known as scale-free networks, and we use the preferential attachment algorithm to construct them. In the preferential attachment mechanism, the nodes with a higher degree, are more likely to receive new links. Intuitively, if the considered network is a social network connecting people, a link from A to B
means that person $A$ "knows" person $B$. Highly connected nodes represent well-known people. Every newcomer is more likely to become acquainted with people with higher visibility compared to a relatively unknown person. Here is the algorithm for generating BA network (Albert and Barabasi, 2002). Initially, the network has $m_0$ all-to-all connected nodes. Each new node is added to the network and connected to $m \leq m_0$ existing nodes. The probability of connection is proportional to existing nodes degree. i.e.

$$p_i = \frac{k_i}{\sum_j k_j}$$

(2-1)

where $k_i$ is the degree of node $v_i$. In this algorithm, the highly connected nodes ("hubs") will be connected to even more links, while nodes with lower degrees are unlikely to be chosen for connection to new nodes. That means the new nodes "prefer" to get connected to the already massively connected nodes. In Fig. 2.6 the BA network with $m_0=4$, $m=2$ and a total number of nodes $N=100$ is sketched.

In this thesis, we use the algorithm proposed in reference (Jalili et al., 2015), which expands the original preferential attachment algorithm proposed by Barabasi and Albert by adding a tunable real parameter $B$ to demonstrate the initial magnetism of the nodes. At each step, the links are connected to existing nodes with probability

$$p_i = \frac{(k_i+B)}{\sum_{j=1}^{N} (k_j+B)}$$

(2-2)

Later on, many researchers found that a lot of real-world networks from chemical reactions to physics, food web to biology, engineering and sociology share many structural properties (Strogatz, 2001). Putting these properties under scrutiny would give a good insight into the underlying phenomena of the system. For example, studying social networks provides a good idea about how information spread in a network of people or biological network has lots of useful information about the evolution of organisations and their units (Barabasi and Oltvai, 2004).

Also, working on the networks based on the world wide web helps scholars in developing methods to conduct more productive navigation and search on the internet (Kleinberg, 2000).
Fig. 2.6 BA network generated by preferential attachment method $(m_0 = 4, m = 4, N = 18)$.

In this chapter, we review some essential concepts of complex network theory and the methods to apply them to the power systems.

### 2.2 Structural Centrality Measures

Due to the complex structure of synthetic and real networks, we need to summarise and abstract their behaviours by using metrics measuring their structural properties. Generally speaking, a centrality measure assigns a real value to each network component, which is expected to be proportional to the importance of that particular component of the network (Freeman, 1977b, Bonacich, 1987, Borgatti, 2005, Borgatti and Everett, 2006). In this section, we review several critical structural centralities used widely in the complex network area.
2.2.1 Neighbourhood-Based Centralities

In this part of the thesis, we consider unweighted and undirected networks. Let’s denote the complex network as $G = (V, E)$, with $V$ and $E$ being the set of nodes and links, respectively. The network can be represented by its adjacency matrix $A = [a_{ij}]$, where $a_{ij} = 1$ if an edge exists between nodes $v_i$ and $v_j$, and $a_{ij} = 0$ otherwise. We consider networks without any self-loops, i.e., $a_{ii} = 0$. The centrality of a node (or edge) in a network determines its importance in a particular functionality of the network. A centrality index assigns a score to all nodes indicating their vitality in the network. Due to the implicit meaning of ‘importance’, disparate indices have been introduced to cover the concept.

**Degree Centrality**

The most trivial method to evaluate the centrality of nodes accounts for the number of immediate neighbours of a node which is presented as degree centrality. The degree $k_i$ of node $v_i$ is the number of edges incidents on that node:

$$k_i = \sum_j a_{ij}$$  \hspace{1cm} (2-3)

Because the degree centrality is computationally a low complex method compared to more complicated centralities like betweenness, closeness and eigenvector centralities; it is applied in many applications like controlling the animal disease epidemics (Candeloro et al., 2016). In a directed network, the edges have direction on them. This property provides two different degrees for each node; in-degree and out-degree. For example, in a citation network, if paper $A$ cites paper $B$, there is a link from node $A$ to node $B$, and the in-degree of a node shows its popularity and importance.

The *Average Node Degree* is calculated as the average of degrees of all nodes as

$$\bar{k} = \frac{1}{N} \sum_{i=1}^{N} k_i$$  \hspace{1cm} (2-4)

This parameter is an indication of how densely the nodes in the network are connected.

While the average node degree represents a particular structural property of a network, it happens that the networks with the same average node degree have entirely distinct topologies.
To grasp another feature of the network using its topological properties, the *Node Degree Distribution, P(k)*, is applied that is defined as the probability that a randomly chosen node has the degree *k*. It is described as equation 2-5.

\[ P(k) = \frac{N(k)}{N} \]  

(2-5)

where \( N(k) \) represents the number of the nodes with degree *k*.

![Node degree distributions](image)

Two popular node degree distributions in real networks are the power-law distribution and the Poisson distribution. In power-law distribution, \( P(k) \) relates to the degree parameter *k*, with equation 2-6.

\[ P(k) \sim k^{-\gamma} \]  

(2-6)

In this equation, \( \gamma \) is the power-law exponent. In a network with power-law degree distribution, a significant portion of nodes have a small number of connections whereas a small part of nodes has a large number of edges connected; i.e. the network is heterogeneous.

In the Poisson distribution, the probability \( P(k) \) is changed with *k* as equation 2-7.

\[ P(k) = \frac{e^{-\lambda} \lambda^{-k}}{k!} \]  

(2-7)
with $\lambda = \bar{k}$, being the average node degree of the network. In networks with a Poisson degree distribution, no hub node can be formed and most of the nodes have the same degree; i.e. the network is homogeneous.

Poisson and power-law degree distributions are shown in Fig. 2.7. It is essential to know if the nodes with the same degrees are connected (Degree Assortativity). One way to quantify this is by using the degree-degree correlation displayed in equation 2.8:

$$r = \frac{\frac{1}{E} \sum_{j} (k_i k_j a_{ij}) - \left[ \frac{1}{E} \sum_{j} (k_i + k_j) a_{ij} \right]^2}{\frac{1}{E} \sum_{j} (k_i^2 + k_j^2) a_{ij} - \left[ \frac{1}{E} \sum_{j} (k_i + k_j) a_{ij} \right]^2}$$

(2-8)

where $E$ is the total number of edges, $k_i$ is the degree of node $v_i$ and $a_{ij}$ is the entry $(i,j)$ in the adjacency matrix. According to the value of $r$, there are three types of networks:

- $r > 0$: The network is called assortative; i.e. the nodes with higher degree intend to connect and the nodes with lower degrees are connected (rich with rich, poor with poor)
- $r < 0$: The network is called disassortative; i.e. the nodes with higher degree intend to connect the nodes with lower degrees and the nodes with low degrees intend to connect to the nodes with higher degrees (rich with poor)
- $r = 0$: The is no intention for the connection between the nodes regarding their degree

Node degree is a simple centrality measure that needs only local information on the nodes. The degree is indeed the most natural centrality measure to compute and has been shown to control many of the network functions.

### Local Rank

To obtain an effective ranking parameter to overcome computationally complex calculations in large-scale networks, Chen et al. (Chen et al., 2012) proposed another local centrality measure, called Local Rank, which considers information on nodes’ fourth-order neighbours. Local rank is a compromise between the degree centrality and other time-consuming measures. The Local Rank of each node $v_i$ is computed as follows:

$$LR(i) = \sum_{j \in R_i} Q(j)$$

(2-9)
\[ Q(j) = \sum_{k \in I_j} R(k) \]  

(2-10)

where \( I_j \) is the set of immediate neighbours of \( v_i \) and \( R(k) \) accounts for the number of immediate and the next immediate neighbours of \( v_k \).

**Clustering Coefficient**

Another measure which quantifies the interconnection in the network is the clustering coefficient. It measures the local connectivity in the network (i.e., indicating to what extent the neighbours of a node are interconnected). For node \( v_i \) with degree \( k_i \), the maximum number of possible edges among its neighbours is \( k_i(k_i-1) \). The clustering coefficient is the portion of these possible edges that indeed exist and is computed as:

\[ C_i = \frac{|[r,s]|}{k_i(k_i-1)} \]  

(2-11)

if edge \([v_r, v_s]\) exists, and both nodes \( v_r \) and \( v_s \) are also connected to node \( v_i \), then we have \(|[r,s]| = 1\).

**Coreness Centrality**

Recently, Kitsak et al. (Kitsak et al., 2010) argued that the position of a node is more effective than its degree. Specifically, the nodes located at the core part of a network, are more influential compared to the nodes in the periphery. Coreness of a node can be procured by using the k-core decomposition (Dorogovtsev et al., 2006) that repetitively decomposes the network according to the nodes residual degree. To calculate coreness, one needs global topological information of the network, which limits its usage to very large-scale networks (Lu et al., 2016). The k-core decomposition method applied to calculate the coreness centrality can be briefed as follows. Initially the nodes with degree zero are removed with \( C_i=0 \) before we begin the k-core decomposition. In the first step, all nodes with degree equal to 1 are removed, until no more nodes with degree one remain in the network. The nodes removed in this stage have the coreness centrality \( C_i=1 \). Likewise, in the second step, repetitively the nodes with degree equal to 2 are
removed, which are assigned to 2-shell and their coreness centrality is 2. This process keeps on going till all the nodes are removed.

**H-index**

H-index (Hirsch, 2005) is a local centrality in which every node only needs the degrees of its neighbours. H-index of a node is defined as the maximum value $h$ such that there exist at least $h$ neighbours with degrees of at least $h$. In this thesis, we consider the average lobby index (al-index), which is computed as follows (Abbasi, 2013):

$$al-index_i = \frac{1}{h} \sum_{j \in TH_i} k_j$$

where $TH_i$ is a set of top-$h$ neighbours of node $v_i$, whose degree is at least $h$.

**2.2.2 Path-Based Centralities**

The centrality measures mentioned above rely on the nodes’ local neighbourhood. However, when it comes to information propagation, the nodes which can transfer the data faster and more extensive, are more critical. We consider some centrality measures that are based on the length of the shortest path.

**Shortest Path Length**

Many of graph centrality parameters are based on shortest path length between pairs of nodes in the network. A path from node $v_i$ to node $v_j$ is a sequence of edges, directed or undirected, connecting node $v_i$ to node $v_j$. Also, the path length is defined as the number of edges in the path. Now we can explain the shortest path length, $d_{ij}$, as the shortest path from node $v_i$ to node $v_j$ when node $v_i$ and the node $v_j$ are connected. The length of a path is defined as the number of edges included in the path. This can be simply extended to weighted networks by considering their weighted length. It is also known as characteristics path or geodesic path. For instance, in a communication network, the shortest path length is chosen for data transmission between two
different data centres. This problem is also known as the min-delay path problem (Wu et al., 2013). This parameter is also applied in transportation, robotics, and VLSI design (Chen, 1996). Chen et al. (Chen, 1996), Ittai et al. (Abraham et al., 2010, Abraham et al., 2011), and Mohri (Mohri, 2002) introduced different methods to calculate the shortest path in the network. It is worth mentioning that the longest shortest path length in the graph is defined as the graph diameter.

**Average Path Length**

In a network consisting of $N$ nodes, the total number of pairs of nodes is $2N(N-1)$, where there is a distinct corresponding shortest path length between each pair of nodes. The average path length, $L$, is the average over all the shortest path lengths of these pairs across the network, i.e.,

$$L = \frac{1}{2N(N-1)} \sum_{i \neq j} d_{ij} \quad (2-13)$$

Here, $L$ is an indication of the interconnectivity of the nodes in a network. For example, for a relatively large social network, $L$ is a small value. In late 60’s, small-world experiment was conducted by Stanley Milgram. They examined the United States’ social network by its average path length and concluded that human society has the small-world property with rather short path length. Although this experiment often comes along with ‘six degree of separation’, Milgram never used this phrase himself. Milgram published the result of this experiment in 1967 (Milgram, 1967). Later on, based on the online Facebook database, Backstrom in 2012 showed a 4 degree of separation (Backstrom et al., 2012), and Bhagat in 2016 showed a 3.5 degree of separation (Bhagat et al., 2016).
Closeness Centrality

Closeness centrality quantifies how a component is in the “middle” of network, not too far from the centre, and is defined as the inverse of the average length of the shortest path between a node and the rest of nodes in the network, that is computed as:

\[
Cl_i = \frac{N-1}{\sum_{j \neq i} d_{ij}}
\]  

(2-14)

where \( Cl_i \) is the closeness centrality of node \( v_i \), and \( d_{ij} \) is the shortest path length between nodes \( v_i \) and \( v_j \).

Node and Edge Betweenness

Freeman introduced the node betweenness centrality as a measure to quantify the control of a person on the communication between other people in a social network (Freeman, 1977a). According to his theory, the nodes that have a high probability of getting visited on a randomly picked the shortest path between a randomly chosen pair of nodes to have a high betweenness centrality index. This type of nodes act as the bridge for the propagation of information through the network and are usually essential components for it. The node betweenness centrality for node \( v_i \) is defined as:

\[
\rho_i = \sum_{p \neq q \neq i} \left( \frac{\psi_{pq}(i)}{\psi_{pq}} \right)
\]  

(2-15)

where \( \psi_{pq} \) is the number of the shortest path from node \( v_p \) to node \( v_q \), and \( \psi_{pq}(i) \) is the number of these paths making use of node \( v_i \). It assesses how a node behaves like a bridge in the network. Nodes with high betweenness centrality often connect different parts of the networks. The edges betweenness centrality sometimes are also interpreted as their load. Betweenness centrality of an edge \( e_{ij} \) connecting nodes \( v_i \) and \( v_j \) is computed as:

\[
\rho_{ij} = \sum_{p \neq q} \left( \frac{\psi_{pq}(e_{ij})}{\psi_{pq}} \right)
\]  

(2-16)

where \( \psi_{pq} \) is the number of the shortest path from node \( v_p \) to node \( v_q \), and \( \psi_{pq}(e_{ij}) \) is the number of these paths making use of \( e_{ij} \).
2.2 Structural Centrality Measures

**Information Centrality**

Another frequently considered centrality index is information centrality index that was introduced by Stephenson and Zelen (K. Stephenson, 1998) to quantify the nodes’ centrality in social networks. It is based on transferable information between any pair of nodes in a connected network. Introduction of this measure is motivated by the theory of statistical estimation. Here the term “signal” is defined as a path connecting a pair of nodes, while the variance of this signal is considered as “noise”. It presumes that the information will be lost during every step of propagation in the network, and therefore in longer routes, more information is lost. Accordingly, it scores the importance of a node based on the information contained in all possible paths between every two nodes (K. Stephenson, 1998). The quantity of data that can be transmitted between every two nodes $v_i$ and $v_j$, can be quantified as a summation of information transmitted through every possible route between them, designated as $q_{ij}$. It has been shown that $q_{ij}$ is equivalent to conductance in electrical grids (Altmann, 1993, Poulin et al., 2000) which implies that:

$$q_{ij} = (r_{ii} + r_{jj} - 2r_{ij})^{-1}$$  \hspace{1cm} (2-17)

with $r_{ij}$ are the elements of the matrix:

$$R = (r_{ij}) = (D - A + F)^{-1}$$  \hspace{1cm} (2-18)

where $D$ is a diagonal matrix with nodes degrees as its diagonal elements and $F$ is a matrix whose items are all one. The information index of $v_i$ is then defined as:

$$INF_i = \left[ \frac{1}{n} \sum_j q_{ij} \right]^{-1}$$  \hspace{1cm} (2-19)

**Eigenvalue Centrality**

It has been shown that the importance of nodes obtained based on the eigenvector corresponding to the largest eigenvalue of the adjacency matrix $A$ plays a role in some dynamical processes. Eigenvalue centrality assumes that the impact of a node is determined by both the number of its neighbours and the influence of each neighbour (Bonacich, 1987, Bonacich,
The eigenvalue centrality of a node is proportional to the summation of the centralities of its neighbours (Bonacich, 2007). The importance of a node $v_i$, denoted by $x_i$, is defined as

$$x_i = c \sum_{j=1}^{n} a_{ij} x_j$$

(2-20)

where $c$ is a proportionality constant. Generally, $c = 1/\lambda$ where $\lambda$ is the largest eigenvalue of the adjacency matrix $A$. Let’s define $m(v)$ as the signed component of the maximal magnitude of vector $v$, i.e. $m(v) = |v|$. Let’s consider $x(0)$ as an arbitrary vector. For $k \geq 1$, repeatedly compute $x(k) = x(k-1)A$; normalize $x(k) = x(k)/m(x(k))$; until the desired precision is achieved.

**Page Rank**

PageRank algorithm (Brin and Page, 1998) is a popular variant of eigenvector centrality and is used to rank websites in Google search engine and other commercial scenarios (Langville and Meyer, 2011). The traditional keyword-based websites ranking algorithms are vulnerable to malicious attacks, which improves the influence of a website by increasing the density of irrelevant keywords. PageRank distinguishes the importance of different websites by random walking on the network constructed from the relationships of web pages. Similar to eigenvector centrality, PageRank supposes that the significance of a web page is determined by the quantity as well as the quality of the pages linked to it. Initially, each node (i.e. page) gets one-unit $PR$ value. Then, every node evenly distributes the $PR$ value to its neighbours along with its outgoing links. Mathematically, the $PR$ value of node $v_i$ at $t$ step is

$$PR_i(t) = \sum_{j=1}^{N} a_{ij} \frac{PR_j(t-1)}{d_{jout}}$$

(2-21)

with $N$ being the total number of nodes in the network, and $d_{jout}$ is the out-degree of node $v_j$. The above iteration will stop if the $PR$ values of all nodes reach a steady-state. A significant drawback of the above random walk process is that the $PR$ value of a dangling node (the node with zero out-degree) cannot be redistributed, and then equation 2-21 cannot guarantee the convergence (there are some other cases where equation 2-21 will not converge (Boccaletti et al., 2006)). To solve these problems, a random jumping factor was introduced by assuming that
the web surfer will browse the web pages along the links with probability \( s \), and leave the current page and open a random page with probability \( 1 - s \). Accordingly, equation 2-21 is modified as

\[
PR_i(t) = s \sum_{j=1}^{N} a_{ji} \frac{PR_j(t-1)}{d_{j}^{out}} + (1 - s) \frac{1}{n}
\]  

(2-22)

### 2.2.3 Flow-Based Centrality (Ford-Fulkerson Algorithm)

To calculate the load of each transmission line in power systems, the analytical tool we apply is Ford-Fulkerson method (Ford and Fulkerson, 1956b, Dwivedi and Yu, 2013). Let us assume \( G \) as a directed weighted network. Any transmission line’s load is proportional to its admittance as \( I = VY \) with \( I, V \) and \( Y \) representing the current, bus voltage, the bus admittance matrix. Hence, to evaluate the capability of a link to carry the power flow, it needs to be weighted by its admittance. One can then find the overall power flow from the source nodes (generators) to sink ones (loads). Here, we apply the Ford–Fulkerson maximum-flow min-cut theorem to determine the maximum flow through the network. This theorem implies that the maximum-flow is equal to the sum of flows across “minimum cut” edges set operating at their maximum flow (Ford and Fulkerson, 1956b, Freeman et al., 1991).

Suppose the link \( e_i \) connects nodes \( v_i \) and \( v_j \). For a network with \( m \) sources and \( n \) sinks, we define \( f_{uv}^{\text{max}} \) as the maximum flow from the source \( u \) to the sink \( v \) and \( f_{ij} \) as the proportion of this flow passing from \( e_i \). Then, a flow-based centrality index is defined as (Freeman et al., 1991, Dwivedi and Yu, 2013):

\[
L_{ij} = \sum_{u=1}^{m} \sum_{v=1}^{n} f_{ij}^{uv}
\]

(2-23)

\( L_{ij} \) is indeed the load of the link between nodes \( v_i \) and \( v_j \). To normalise these weights, we simply divide them by the total \( f_{\text{max}} \), as

\[
\tilde{L}_{ij} = \frac{\sum_{u=1}^{m} \sum_{v=1}^{n} f_{ij}^{uv}}{\sum_{u=1}^{m} \sum_{v=1}^{n} f_{\text{max}}^{uv}}
\]

(2-24)
Note that the input to the Ford-Fulkerson maximum flow algorithm is the weighted adjacency matrix of the power grid, where the link weights are the admittance values. The output of this algorithm is another weighted matrix representing the maximum flow for the links. This matrix can be used to rank the edges based on the load they carry and thus identifying the most vulnerable edges. As the flow passing through an edge increases, its weights will also increase. Fig. 2.8 shows the flow distribution for a sample network with two sources and three sinks. The initial admittances of the links are shown in Fig. 2.8a and the normalised centrality indices obtained according to the above method are shown in Fig. 2.8b.

2.3 Complex Network Models of Power Grids

Complex network theory is interested in the characteristic properties of various complex networks and their general analysis methods. One major branch of this theory is the relationship between the topology of the complex network and its performance. The literature is divided into two main categories where some paid attention to detailed modelling of subsystems while some others worked on the overall structure of the complex networks. A detailed network model is a fundamental tool to understand and explore the relationship between the topological
characteristics and the performance of power systems. Mainly, appropriate models of power systems can be applied for testing the control or planning strategies, conducting sensitivity analysis for cascade mitigation and also examining possible scaling by generalising structural metrics.

A significant part of this thesis focuses on the structural analysis of power systems, especially for a chain reaction and failure propagation in their components for mitigation of cascade failures. Hence, in this section, we review the essential complex network models of power systems.

### 2.3.1 Minimum-distance graph

This model aims to generate an undirected graph with the structural properties similar to that of power systems. In the minimum distance method, to connect any pair of nodes, the cost of their distance is taken into consideration. Let’s define $\psi_i$ as the set of neighbours for node $v_i$ and $N$ and $E$ as the number of desired nodes and edges, respectively. First, using a uniform distribution, a random coordinate for each node $v_i$ is generated as $v_i(x_i, y_i)$ where $x_i$ and $y_i$ are the coordinates along $x$ and $y$ axes. Then, $\text{int}(E/N)$ number of edges $[i,j]$ between nodes $v_i$ and $v_j$ is created such that the Euclidean distance between them is minimised:

$$
\min_{v_j} \quad (x_i - x_j)^2 + (y_i - y_j)^2 \\
\text{s.t.} \quad v_j \notin \psi_i \\
$$

with $\text{int}(.)$ being the floor value function. The minimisation can be achieved using the random walk through any possible route with the above property and comparing the computation results of equation 2-25. Finally, an additional edge is generated with the probability $P = E/N - \text{int}(E/N)$. 
2.3.2 Synthetic Spatial Power Network

One of the essential traits of real-world power networks that were neglected in most studies is their spatial constraints (Barthélemy, 2011). Indeed, power networks are spatially extended systems and geographical location of the nodes play essential roles in their connections. Spatial networks show significantly distinct behaviour compared to most of the spatially unrestricted networks (Asztalos et al., 2014). Since none of the methods in the literature can model the power systems accurately; especially most of them ignore their physical properties, in this section, we introduce a synthetic model to construct a spatially extended power network with the source, sink and intermediate nodes. Similar to real power systems, the synthetic model network is also a directed network with flow from the source (generator) to sink (load) nodes. As a spatial network, generators are more likely to feed the loads in their neighbourhood. Likewise, loads are connected to their upstream buses nearby. Not to mention that if busbar \( v_j \) is fed by busbar \( v_i \), then \( v_i \) is upstream busbar for \( v_j \). To have more realistic measures and stronger analogous parameters, we generate a synthetic power network with 118 buses consisting of 91 loads, 19 generators, and 196 lines similar to IEEE 118 bus test set. These 118 buses are randomly distributed in a 118 by 118 plane. Also, 19 generators and 91 loads are randomly distributed on 118 buses as well as intermediate buses. We consider a circle with radius \( r \) around each node (\( r_g \): the radius around generators, \( r_i \): radius for intermediate buses, and \( r_l \): radius for loads) to connect with others in this circle with probability \( p \). Based on these considerations, we assume the connectivity radius for generators as \( r_g = 15 \), for intermediate buses as \( r_i = 20 \), and for loads as \( r_l = 25 \). We also fix the connection probability at \( P = 0.4 \). Fig. 2.9 shows a synthetic spatial power model for 118 buses and 19 generators. Also, to have similar weights in random power network and IEEE 118 bus system, we generate the weights of the model network (admittances) according to a normal distribution with the mean value and standard deviation of IEEE 118 bus system.

Table. 2.1 shows the comparison between IEEE 118 bus system and the synthetic model parameters where the similarity between these two complex networks are quantified.
2.2 Structural Centrality Measures

Table 2. Comparison of structural properties of IEEE 118 bus and synthetic model

<table>
<thead>
<tr>
<th>Feature/Test Case</th>
<th>IEEE 118 Bus</th>
<th>Synthetic Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of nodes</td>
<td>118</td>
<td>118</td>
</tr>
<tr>
<td>No. of edges</td>
<td>179</td>
<td>196</td>
</tr>
<tr>
<td>Network Density</td>
<td>0.0259</td>
<td>0.0283</td>
</tr>
<tr>
<td>Average Path Length</td>
<td>6.308</td>
<td>5.704</td>
</tr>
<tr>
<td>Average Degree &lt;k&gt;</td>
<td>3.033</td>
<td>3.220</td>
</tr>
<tr>
<td>Efficiency</td>
<td>0.210</td>
<td>0.209</td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>0.165</td>
<td>0.158</td>
</tr>
<tr>
<td>Assortativity</td>
<td>-0.1525</td>
<td>-0.145</td>
</tr>
</tbody>
</table>

Fig. 2. 9 synthetic spatial network constructed by the model introduced in section 2.3.2
Red circles: Generators and blue triangles: Loads connected directly with intermediate buses or generators

2.3.3 AC/DC Model

By studying the power flow in power systems, we aim to calculate the value of power in the transmission lines. To do so, we need to consider the actual variations in power system
parameters like voltages, currents and the topology of the network. Also, it should comply with the physical laws and constraints like Kirchhoff’s laws and Ohm’s law. Based on these realisations, in this section, the AC and DC, power flow methods, applied in power systems engineering are described.

AC Model

Studying the AC load flow in a power system enables us to calculate the magnitude and the angle of the busbar voltages given the load specification and the generators real power, \( P \), and the voltage situation (Grainger and Stevenson, 1994). By knowing these data, active and reactive power at each transmission line as well as the generator output reactive power can be calculated. Since the nature of such a problem is nonlinear; the numerical methods should be applied for solving it. This method is only valid in the steady-state of the power systems when there is no transition in power system parameters like the loads’ value, generated power, and frequency.

To begin the formulation, let us model the busbar \( v_i \) with a node at which the voltage is \( V_i = |V_i| \angle V_i \), where \( |V_i| \) is the voltage magnitude, and \( \angle V_i \) is the voltage phase angle. Likewise, \( V_k = |V_k| \angle V_k \) is the voltage at busbar \( k \). Also, the phase difference between nodes \( v_i \) and \( v_k \), is defined by \( \theta_{ik} = \theta_i - \theta_k = \angle V_i - \angle V_k \). Likewise, a transmission line \( e_{ik} \) connecting two busbars \( v_i \) and \( v_k \), can be modelled as a link \( e_{ik} \) connecting corresponding nodes \( v_i \) and \( v_k \). In this link, the current can be evaluated as \( I_{ik} = (V_i - V_k)(G_{ik} + jB_{ik}) \), with \( B_{ik} \) and \( G_{ik} \) being the susceptance and conductance of line \( e_{ik} \), respectively. Thus, the injected power to node \( v_i \) via the transmission line \( e_{ik} \) is

\[
S_{ik} = V_k I_{ik} = P_{ik} + Q_{ik}
\]

where \( P_{ik} \) and \( Q_{ik} \) are the active and reactive powers respectively. \( P_{ik} \) and \( Q_{ik} \) can also be restated as:

\[
P_{ik} = |V_i|^2 G_{ik} - |V_i||V_k|(G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik})
\]

\[
Q_{ik} = |V_i|^2 B_{ik} + |V_i||V_k|(G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik})
\]
2.2 Structural Centrality Measures

For busbar $v_i$, the injected power is defined as $S_i = P_i - jQ_i$, with $P_i$ and $Q_i$ being the active and reactive powers.

Since the injected power to busbar $v_i$ via the transmission lines is zero-sum, two following equations can describe the active and reactive power in it:

$$P_i = -\sum_{k=1}^{N} P_{ik} = \sum_{k=1}^{N} |V_i||V_k|(G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik})$$  \hspace{1cm} (2-29)

$$Q_i = -\sum_{k=1}^{N} Q_{ik} = \sum_{k=1}^{N} |V_i||V_k|(G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik})$$  \hspace{1cm} (2-30)

where $G_{ii} = -\sum_{i \neq k} G_{ik}$ and $B_{ii} = -\sum_{i \neq k} B_{ik}$. In case there is no immediate link between nodes $v_i$ and $v_k$, $G_{ik} = B_{ik} = 0$.

Here, we define the known and unknown parameters of the power system which depend on the type of busbars. If there is a generator connected to a busbar, it is identified as the Generator busbar, otherwise it is a Load busbar. Just one exception is that one of the Generator busbars is selected arbitrarily as the reference bus named slack busbar. In Load busbars, the active power $P_D$ and reactive power $Q_D$ are known values. Hence, we call them $PQ$ busbars. Likewise, for Generator busbars, the active power $P_G$ and the voltage magnitude $|V_G|$ are known. So, they are called $PG$ busbars. Also, for the slack busbar, the voltage magnitude $|V_S|$ and the voltage phase $\theta_S = \angle V_S$ are considered as known parameters. Therefore, the voltage magnitude and phase for the Load busbars, and also the voltage angle for the generator busbars should be treated as unknowns. There is no unknown value for the slack bus as it is the reference busbar. For a system having $N$ busbars and $R$ generators, the number of unknowns is $2(N-1) - (R-1)$, thus there should be the same number of independent equations.

Several algorithms have been proposed to solve the load flow equations, such as the Newton-Raphson method, the Gauss-Seidel method and the Fast-decoupled load flow method.

**DC Model**

Although the AC power flow algorithm can be the most accurate method for calculating the power flow across the grid in the steady-state, it is computationally costly especially when it is supposed to solve a large-scale power system. In this case, large-scale nonlinear equations...
should be solved iteratively at each time step till all the parameters converge to their static values. As stated in the previous section, large-scale nonlinear equations should be solved iteratively at each time step until all of the parameters converge to their static values. Therefore, a linearized version of AC power flow algorithm can be used instead. To do so, some assumptions (Mei et al., 2011) are made:

1. The resistance of each line is much smaller than its reactance ($R_L << X_L$). Thus, we neglect the grid losses, and the line parameters can be rewritten as:

\[
G_L = \frac{R_L}{R_L^2 + X_L^2} \approx 0
\]

\[
B_L = \frac{-X_L}{R_L^2 + X_L^2} \approx \frac{-1}{X_L}
\]

\[
Z_L \approx jX_L, \quad Y_L \approx jB_L
\]

2. There is almost a flat profile for all busbar voltages across the power grid, i.e. the voltage magnitude in per-unit system is equal for all busbars: $|V_i| \approx 1 \text{ p.u.}$

3. The phase difference between two connected busbars is negligible. This assumption will lead to the linearization of sine and cosine terms in the original model:

\[
\sin(\theta_{ik}) = \sin(\theta_i - \theta_k) \approx \theta_i - \theta_k
\]

\[
\cos(\theta_{ik}) = \cos(\theta_i - \theta_k) \approx 1
\]

Now the equation 2-24 can be restated as:

\[
P_{ik} = -(B_{ik} \sin \theta_{ik}) \approx -B_{ik} (\theta_i - \theta_k)
\]

which implies that according to the equation 2-26, for each bus:

\[
P_i = -\sum_{k=1}^{N} P_{ik} = \sum_{k=1}^{N} B_{ik} (\theta_i - \theta_k)
\]

which is the nodal equation for the busbar $v_i$. The matrix format of this equation would be:

\[
P = B\theta
\]

with $P = [P_1 \quad P_2 \ldots P_N]^T$, $\theta = [\theta_1 \quad \theta_2 \ldots \theta_N]^T$ and $B$ being an $N \times N$ matrix containing the transmission line’s admittance. To solve these equations, the consumption load in Load busbars and also the generated active power in Generator busbars are needed. Due to the fact that in DC model the line resistance is neglected, there would be no $RF^2$ loss in the system. Hence, it is called lossless. Therefore, the sum of the consumed power is equal to the amount of the
generated power. Thus, for a power system having $N$ busbars, the power injected for $N-1$ nodes and also the voltage phase for the reference busbar is needed.

### 2.4 Benchmark Networks

In this section, various benchmark networks applied for simulation in future chapters are summarised. First, node vitality measures are examined on benchmark networks.

Benchmark networks consist of both synthetic and real networks. As synthetic networks, we use BA, WS and ER networks, which are explained in section 2.1. Studying model networks can provide insights on how real networks behave; however, they might not capture all properties of real networks. Thus, here several real networks are also studied.

<table>
<thead>
<tr>
<th>Network/Parameter</th>
<th>N</th>
<th>E</th>
<th>&lt;d&gt;</th>
<th>APL</th>
<th>$\bar{C}$</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia power grid</td>
<td>1130</td>
<td>1426</td>
<td>2.5221</td>
<td>12.3545</td>
<td>0.0768</td>
<td>0.00112</td>
</tr>
<tr>
<td>US Power Grid</td>
<td>4941</td>
<td>6594</td>
<td>2.6690</td>
<td>18.9891</td>
<td>0.1031</td>
<td>0.00027</td>
</tr>
<tr>
<td>European Power Grid</td>
<td>1494</td>
<td>2156</td>
<td>2.8862</td>
<td>18.8847</td>
<td>0.0947</td>
<td>0.00096</td>
</tr>
<tr>
<td>UK Power Grid</td>
<td>327</td>
<td>407</td>
<td>2.4892</td>
<td>12.9994</td>
<td>0.0677</td>
<td>0.00381</td>
</tr>
<tr>
<td>IEEE 300-Bus System</td>
<td>300</td>
<td>409</td>
<td>2.7266</td>
<td>9.9353</td>
<td>0.0960</td>
<td>0.00455</td>
</tr>
<tr>
<td>French Power Grid</td>
<td>146</td>
<td>223</td>
<td>3.0547</td>
<td>6.6081</td>
<td>0.2601</td>
<td>0.01051</td>
</tr>
<tr>
<td>IEEE 118-Bus System</td>
<td>118</td>
<td>179</td>
<td>3.0338</td>
<td>6.3087</td>
<td>0.1355</td>
<td>0.01294</td>
</tr>
<tr>
<td>Spanish Power Grid</td>
<td>98</td>
<td>175</td>
<td>3.5714</td>
<td>4.9168</td>
<td>0.3532</td>
<td>0.01842</td>
</tr>
<tr>
<td>Internet Autonomous</td>
<td>3233</td>
<td>5773</td>
<td>3.5712</td>
<td>3.7624</td>
<td>0.0152</td>
<td>0.00055</td>
</tr>
<tr>
<td>Melbourne Natural Water</td>
<td>2145</td>
<td>2644</td>
<td>2.4652</td>
<td>16.8134</td>
<td>0.0922</td>
<td>0.00057</td>
</tr>
<tr>
<td>Melbourne Roads Grid</td>
<td>1830</td>
<td>2678</td>
<td>2.9267</td>
<td>7.6054</td>
<td>0.0469</td>
<td>0.00082</td>
</tr>
<tr>
<td>US Airports Network</td>
<td>1574</td>
<td>17215</td>
<td>21.8742</td>
<td>3.1151</td>
<td>0.3841</td>
<td>0.00695</td>
</tr>
</tbody>
</table>

We consider eight power networks including Australian high voltage power grid, Western US power system, European power network, UK, French and Spanish power systems, IEEE 300 Bus and 118 Bus benchmark networks. We also study some other real networks, including the Internet in the level of the Autonomous system, US Airport network, Melbourne road network and Melbourne water networks. Table 2.2 represents the structural features of the real networks considered in this work.
Table 2. Demographic information of the real networks under consideration.

$N$: number of nodes, $E$: number of edges

<table>
<thead>
<tr>
<th>Networks</th>
<th>$N$</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Synthetic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>200</td>
<td>983</td>
</tr>
<tr>
<td>WS</td>
<td>200</td>
<td>1000</td>
</tr>
<tr>
<td>BA</td>
<td>200</td>
<td>975</td>
</tr>
<tr>
<td><strong>Power Grid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish Power Grid</td>
<td>98</td>
<td>175</td>
</tr>
<tr>
<td>French Power Grid</td>
<td>146</td>
<td>223</td>
</tr>
<tr>
<td>UK Power Grid</td>
<td>327</td>
<td>407</td>
</tr>
<tr>
<td>Australian Power Grid</td>
<td>1130</td>
<td>1426</td>
</tr>
<tr>
<td><strong>Biological</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macaque Cortical Network</td>
<td>71</td>
<td>438</td>
</tr>
<tr>
<td>Cat Cortex Network</td>
<td>95</td>
<td>1170</td>
</tr>
<tr>
<td>C. elegans Network</td>
<td>297</td>
<td>2359</td>
</tr>
<tr>
<td><strong>Real (Infrastructure)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Airports</td>
<td>332</td>
<td>2126</td>
</tr>
<tr>
<td>Melbourne Natural Water</td>
<td>2145</td>
<td>2644</td>
</tr>
<tr>
<td>Melbourne Roads</td>
<td>1830</td>
<td>2678</td>
</tr>
</tbody>
</table>

In the following chapters, cascade propagation methodology is applied on some synthetic networks as well as real-world networks ranging from biological networks to road networks, from power grids to airport and natural water distribution network. Table 2.3 briefs the number of nodes and edges in the network under consideration. To examine the power system vulnerability against cascade failure, both real and synthetic model networks are applied. As a real network, we consider IEEE 118 bus network (Chen and McCalley, 2005, Machowski et al., 2011), consisting of 19 generators, 9 transformers, 186 lines and 91 loads as shown in Fig. 2.10.

We also introduce a synthetic model to construct a spatially extended power network with the source, sink and intermediate nodes which is explained in section 2.3.2. We apply the load reduction algorithm on IEEE 57 bus network, which consists of 7 generators, 80 lines and 42 loads and is shown in Fig. 2.11. The IEEE 57 busbar test case is a representation of a section of the American Electric Power Grid (in the Midwestern US) in the 1960s.
2.4 Benchmark Networks

Fig. 2. 10 IEEE 118 bus system used in this work as a benchmark network. This figure has been adopted from (Lu, 2008).

Fig. 2. 11 IEEE 57 test case consisting of 57 buses, 80 lines and 42 loads
Fig. 2.12 Degree histogram of the networks considered in this work. It consists four power grids (Australian, US, French and IEEE 300 bus networks) and four real networks (US Airport, Melbourne water, cat cortex and Melbourne roads networks).

In fig. 2.12, the degree distribution of four power systems (Australian, US, French and IEEE 300 bus networks) as well as four other real networks (US Airport, Melbourne water, cat cortex and Melbourne roads networks) are depicted. Generally, in power networks, there are not any high connected busbars (nodes), whereas in some other real networks like US Airport and Melbourne roads networks there are plenty of hubs.
2.5 Conclusion

In this chapter, the complex network fundamentals, models, and structural centralities were reviewed. When it comes to model complex networks such as power systems, synthetic models are advantageous candidates. Three different synthetic models, i.e. ER, WS, and BA models, and the methods to generate them along with their properties were discussed.

Then, the structural centralities of complex networks were explained. We can assign a score to any component in a complex network, e.g. nodes or edges using centrality measures. Structural centralities can be categorized into three main groups: centralities based on the local neighbourhood which only takes the nodes’ neighbours into consideration, path-based centralities which concentrate on the information flow in the network. It means the components which appear more in information transfer across the network, get higher centrality values.

Finally, flow-based centrality was reviewed which mainly applies the very well-known Maximum-flow based Ford-Fulkerson method to compute the load of each component in power system.

Afterwards, the complex networks models of power grids which can be adopted to mimic the power systems behaviours were reviewed. In minimum-distance method, a cost function is evaluated to connect any two nodes.

In some cases, to confirm a proposed analytic method and its results on a real complex network, another similar network is necessary to conduct the same analysis. This twin network with similar structural properties can hardly be found in real-world in many cases. Hence, we try to generate them using synthetic networks generation methods. In the case of power networks, we introduced the synthetic spatial power network which considers the power grids physical properties like their spatial extension while connecting the busbars. Besides, it assumed that there are different types of nodes in the network, e.g. in power system case, we have generators, intermediate and load nodes in the model.

AC power system load flow analysis is a complicated and also more accurate method for computing the power system’s parameters compared to other methods in the expense of costly
calculation and divergence risk, which can be traded off by using its approximation DC load flow. In the next chapter, the cascade failure phenomenon in complex networks and also in the power system as a specific complex network is studied.
Chapter 3

CASCADE FAILURES

3.1 Introduction

Power grids, road networks, airports, water distribution infrastructures and the Internet are examples of critical networks playing essential roles in modern communities. Their proper functioning and resiliency are of extreme importance. Such critical networks might be subject to failures in their components (nodes/edges). Errors (i.e., random failure) and attacks (i.e., intentional failures) have been studied in complex networks (Albert et al., 2000, Mirzasoleiman and Jalili, 2011, Jalili, 2011, Perc, 2009). It has been shown that the resiliency of complex networks against errors and attacks depends on their structure. For example, scale-free networks – which are characterised by heavy-tailed degree distribution – are fragile to attacks, while being resilient against errors (Albert et al., 2000). Networks with more homogeneous degree distribution (e.g., random networks) behave similarly against degree-based errors and attacks.

In some cases, a more catastrophic situation can happen, and a failure in one component may propagate through the network resulting in failing some other components. It is worth mentioning that the cascade failure proceeds based on load redistribution in the network and components’ individual parameters do not significantly affect it. This is referred to as cascade failure in the literature (Crucitti et al., 2004a, Wang et al., 2008, Li et al., 2012, Simonsen et al., 2008). Such cascade failures were responsible for large-scale blackouts in power grids (Buldyrev et al., 2010). Power grids can be modelled as networked structures with generators, loads and transformers as nodes and wirings as links. Any failures in one of the network components may give rise to an overload in other components and as a result a failure in them. Such failures can pervade the network and result in a cascade failure process (Ren and Dobson, 2008, Crucitti et al., 2004a, Jalili, 2015), which can, in turn lead the whole
or a fundamental part of the network to crash. In next section, the process of cascade failure will be discussed in detail.

3.2 The Process of Cascade Failures

Cascade failures have been frequently reported in real networks (Xu and Wang, 2005, Kinney et al., 2005). They are often responsible for large-scale blackouts in power systems. In this section, we review a widely-studied process (Mirzasoleiman et al., 2011) applied in cascade failures analysis in a complex network.

Network components (nodes/edges) might undergo random failure and targeted attacks. In some cases, this could lead to a cascade of failures in the network, where a failure might trigger a failure(s) in other components. The cascade failure process is as follows.

Initially, a set of components (nodes/edges) fail due to an overload, cyber-attack or possibly any other malfunction in the protection system. To compensate for the flow shortage caused by this failure and keep the information or power flowing in the network, the loads on impacted components are redistributed through the network. In other words, as the initial failure happens, the failed edge(s) are removed from the network, and the load redistribution algorithm is applied on the modified network to calculate the new loads for components.

Here, we assume that each edge $e_{ij}$ has bounded capacity $C_{ij}$, demonstrating the maximum load which can be delivered by the edge. Hence, those lines loaded higher than their capacity are overloaded, and thus removed from the network. We suppose that $C_{ij}$ is proportional to its initial load $W_{ij}$, $(C_{ij} = TW_{ij})$, where $T > 1$ is the capacity parameter (Mirzasoleiman et al., 2011, Wang and Chen, 2008) and the initial loads for all the edges are calculated at the outset. As $T$ increases, the network becomes more tolerant, and the failures have less impact on that. Often, for any edge, there exists a constant threshold value $T_c > 1$ such that if additional load is larger than its capacity, then the edge cannot tolerate the increased flow. A phase transition occurs at this critical point, and the network faces further redistribution of the flows. Consequently, more edges might break down and this process will keep running till every remaining edges’ load gets less than its capacity. This process continues until no further edge is removed from the network, i.e. no edge is overloaded as a result of load
redistribution. Such a cascade failure process might lead to failures in a large number of the links, and thus large-scale blackout in power grids. A pseudo-code of the cascade failure process is shown in Procedure 3.1.

It should be noticed that two load redistribution methods are applied in this work: edge betweenness (defined by equation (2-16)) and maximum-flow based method (defined by equation (2-24)).
Procedure 3.1: Cascade Failure Procedure
Given a connected graph $G = (N, E)$, a threshold value $T$, and an initial edge failure $e_0$ event:

1. For each edge $e_k \in E$ connecting nodes $v_i$ and $v_j$
   - Calculate the initial load in $e_k (L_{ij})$

2. Initially $C_{ij} = L_{ij}$

3. $(G - e_0) \rightarrow G$ remove $e_0$ from the graph

4. For each remaining edge $e_k \in E$
5.   Calculate $L_{ij}$
6.   If $L_{ij} / C_{ij} > T$
7.     $(G - e_k) \rightarrow G$

8. Repeat 4 until no edge gets removed, and the cascade stops

9. Return $G$

In Fig. 3.1, the process of cascade failure in IEEE benchmark 30-bus network is displayed. At the beginning (a), the network is working normally in steady-state and all the nodes in the network are connected through edges. Not to mention that the thickness of the edges is proportional to their capacity. Initial failure is planted on node 6 which makes that node and all edges connected to it get removed from the network (b). Because of change in the network structure, the flow of power is redistributed in the network and every overloaded edge is removed from the network (c). This procedure keeps going (d, e) until no more edge gets overloaded and removed. At this stage cascade stops and the network reaches to another steady-state (f).

3.3 Cascade Failure Models in Complex Networks

It has always been crucial for scholars to understand the way complex networks behave during a cascade of failures; especially the mechanism of cascade spreading throughout the complex networks is very important. To do so, many researchers preferred to study the cascade by considering the macroscopic parameters of a number of complex networks such as power systems, whereas; some others decided to investigate the microscopic component properties. One of the major branches in this area is to find out the role of structural vulnerability in the spreading of a cascade in complex networks. Among a number of models introduced for studying the cascade failures, some only investigated the static structural
models which are not that successful in explaining the real behaviour of power systems. On the other hand, there are some dynamic models trying to comprehensively describe the complex networks; for example, in power system case, both slow dynamics evolutions like network augmentation and upgrades as well as the fast dynamics like short circuits or any unanticipated faults.

In this section, different cascade models for the complex networks are explained. First, the structural models are studied consisting of static, Motter-Lai and Effective Efficiency models. Following that, the component cascading failure dependent models are reviewed. This category consists of CASCADE model and branching process model.

### 3.3.1 Structural Models

The structural models characterise the macroscopic attributes of complex networks. They apply topological properties of the complex networks to measure their vulnerability to multiple attack scenarios and evaluate their resiliency against cascade failures (Holme, 2002, Holme et al., 2002, Motter and Lai, 2002, Wang et al., 2008, Crucitti et al., 2004a, Holme and Kim, 2002). There are three major structural models; the static model, Motter-Lai and effective efficiency models. These models are further described in the following sections.

#### Static Model

The first and less complicated cascade model for complex networks is the static model of network vulnerability to failures or attacks. In this model, the cascade is mainly triggered by the removal of a single or multiple components, nodes or edges, from the network which may lead to dysfunction of the total network. After some chain reactions in the network and removal of more components, the failure stops. Then, to quantify the extremity of the destruction, the remaining network is compared with the original network. Here, some metrics have been proposed to measure the severity of the damage. For instance, Crucitti et al. (Crucitti et al., 2005) applied the network global efficiency $E(G)$ as an indication of the effectiveness of the connections in the network.

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}}$$ (3-1)
with $d_{ij}$ being the shortest path length between nodes $v_i$ and $v_j$. The lower the proportion of global efficiency of the network after and before the attack is, the more vulnerable the power system is to the cascade of failures. Also, Albert and Solé (Albert et al., 2004, Solé et al., 2008a) proposed that the largest connected component can successfully quantify the extremity of the attack on the network.

To study the effect of removing a single edge, $e_{ij}$, Mirzasoleiman et al. (Mirzasoleiman et al., 2011) calculated the number of broken edges, $s_{ij}$, which can be interpreted as the normalised avalanche sizes after the cascade process is stopped. Then, they proposed the Normalized average size of the removed edges ($S_N$) which is the expected value of the normalised avalanche sizes ($s_{ij}$) corresponding to each edge $e_{ij}$ removal.

$$S_N = \frac{\Sigma_i \Sigma_j s_{ij}}{E}$$  \hspace{1cm} (3-2)

To capture the dynamical properties of complex networks and discover the procedure of cascade propagation through these systems, a number of dynamical models have been introduced. Here, we summarise these dynamical models used for studying the cascade mechanism of complex networks.

**Motter-Lai Model**

To consider the dynamics of the networks, Motter and Lai (Motter and Lai, 2002) introduced a model to explain the procedure of cascade failure in complex networks using discrete-time, step-by-step, analysis method. Suppose that in a network with the representative graph $G$, at each time step one unit of energy is transferred between each pair of nodes through the shortest path connecting them. Now, the load of each node can be defined as the total number of shortest paths making use of that node (Newman, 2001, Goh et al., 2001, Holme and Kim, 2002) i.e. the node’s betweenness. The maximum load a node can handle is defined as the node capacity. Since in real networks constructed by humans, the capacity is limited by the budget; naturally it is assumed that the capacity $C_j$ of node $v_j$ is a multiple of its load $L_j$.

$$C_j = (1 + T)L_j(0)$$  \hspace{1cm} (3-3)
where $T > 0$ is a tolerance parameter, and $L_j(0)$ is the initial load of node $v_j$ with the complex network being in its normal mode.

In this model, the initial failure is planted in the network by removal of a single node. This failure changes the structure of the network and the way the shortest paths are distributed, which in turn might lead to overload and failure of other nodes. This procedure of node removal and load redistribution goes on until no node is overloaded. Now the damage can be measured by the size of the giant component, the largest connected component, which is quantified relative to the number of nodes in the network before node removal ($N$). This algorithm was applied to Western America power system, and the authors concluded that the attack based on the nodes’ load makes more severe damage than the degree-based attack (Motter and Lai, 2002). Problem with this model is that in reality, the power between any pair of nodes does not necessarily flow in the shortest path between them, which is a structural parameter. Instead, according to Kirchhoff’s law, the physics of the power system and parameters like the impedance and admittance, determine the path and amount of the power flow between nodes.

**Effective Efficiency Model**

Crucitti et al. (Crucitti et al., 2004a) proposed a method to transform the networks to a weighted, undirected graph, $G$, where the weights on edges are allocated by the parameter called effective efficiency. $G$ is described by an $N \times N$ adjacency matrix. If there is an edge $e_{ij}$ between nodes $v_i$ and $v_j$, there will be a real number in the range $(0,1]$ attached to that element in the adjacency matrix. This number is a representative of flow efficiency between nodes $v_i$ and $v_j$. Initially, for all existing edges, we have $e_{ij} = 1$, then the time evolution procedure is conducted to emulate the flow redistribution dynamic after a node removal. The algorithm assumes that the flows between a pair of nodes using the most efficient path linking them. For each path, the efficiency is defined as the superposition of effective efficiencies of all links forming that path.

The load of node $v_i$, displayed by $L_i(t)$, is defined as the number of the most efficient paths making use of that node at time $t$ (Goh et al., 2001). The capacity of each node, $C_i$, is
the highest load it can handle which is assumed to be proportional to its initial load $L_i(0)$ (Motter and Lai, 2002).

$$C_i = T.L_i(0)$$  \hspace{1cm} (3-4)

Parameter $T$ is the tolerance parameter in the network. The evolution happens at each time step. Another feasible assumption which brings up the same results is to set:

$$C_i = T.max_j L_j(0), \; \forall i = 1, 2, ..., N$$  \hspace{1cm} (3-5)

When a node is removed, due to a change in the structure of the network, the most efficient paths between pairs of nodes and also the load distribution is changed which might lead to an overload in some nodes. If the load in the node $v_i$ is less than its capacity $C_i$, its efficiency remains unchanged. Otherwise, the effective efficiencies for all the links connected to the node $v_i$ will be derated as the following iterative rule at time $t$:

$$e_{ij}(t) = \begin{cases} 
    e_{ij}(0) \frac{c_i}{L_j(t)} & \text{if } L_i(t) > C_i \\
    e_{ij}(0) & \text{if } L_i(t) \leq C_i
  \end{cases}$$  \hspace{1cm} (3-6)

with $e_{ij}(0)$ being the initial efficiency of edge connecting nodes $v_i$ and $v_j$. After the initial node failure in the network, the iterative rule presented in equation 3-6, will be kicked off and iterated until the network reaches a steady state. As an indication for the severity of cascade failures, the most efficient paths will be averaged. The authors summarised that the removal of highly loaded nodes could lead to a severe cascade of failure which in turn leads to a catastrophic collapse across the complex network (Crucitti et al., 2004a).

Despite two latter models, Motter-Lai and the effective efficiency models, intend to study the dynamics of the networks, the flow redistribution algorithm in both cases, fail to capture the electrical parameters of power systems, since they only consider the structural parameters of the networks. Some other inaccurate works were conducted in Wang and Rong (2009) where a load of the failed node is distributed amongst its neighbours or in (Yagan, 2015) after a line removal, its load is equally dispatched across the network lines.

### 3.3.2 Component Cascading Failures Dependent Model

The second category of cascade models is the component cascading failures, including the CASCADE model (Dobson et al., 2007, Dobson et al., 2003, Dobson et al., 2005b) for transferring load and the model for the branching process (Dobson et al., 2004, Dobson et
3.3 Cascade Failure Models in Complex Networks

In these models, no specific structure of complex networks is taken into consideration. Instead, the cascade of failure happens as a result of overloading, which in turn leads to component failures in complex networks.

**CASCADE Model**

Here, the fundamentals of the CASCADE model is described. Suppose that the network has $N$ identical components, which initially work on steady-state and are loaded by independent load values $L_1, L_2, \ldots, L_N$, chosen randomly from the range $[L_{\text{min}}, L_{\text{max}}]$. Then, a disturbance $D$ is applied indiscriminately on all the loads in the network. Every component whose load is more prominent than its threshold fails and its load is redistributed among other components, which are working normally. Sometimes this procedure may yield a cascade of failures. In a study (Dobson et al., 2005b), the procedure of cascade failures evolution leading to the whole complex network collapse was studied as a function of operational parameters like loading level. The authors found that when all the components are lightly loaded, the faults across the network are independent and the cascade distribution is exponential which means the system rarely experiences the large-scale collapse. On the contrary, when the loading level reaches to a critically high value, the fault shows the power-law distribution where if the loading exceeds that threshold, large-scale collapse might be inevitable. Although CASCADE model simulates the overall outage, there are some restrictions in its assumptions since all the components are assumed to be identical and while redistributing the loads, it doesn’t consider the network topology.

CASCADE model is an idealistic model of what happens during the cascade of failures as it ignores the network structure and how the components are connected. It also assumes that the faulty components make the overall load to be redistributed across the network. In real complex networks, however, there are a lot of dynamics and evolutions during the cascade, which are simply ignored in the CASCADE model. Scholars study this model as a comparison with their introduced models.
Branching Process Model

In probability theory, a branching process is known as a stochastic process and includes collections of random variables. It has been initially designed to mathematically model a population where each item in generation $t$ generates an arbitrary number of elements in the generation $t+1$ according to a fixed probability distribution (Athreya, 2006). For example, the item could be bacteria which produce 0, 1, or 2 new children with a probability in a single period.

Branching process model based on branching processes introduced by Dobson et al. (Dobson et al., 2004, Dobson et al., 2005a, Dobson, 2004) and the commensurate models, were designed to accurately model the cascade failures in complex networks. The system is considered to have infinite components. Also, a component which was faulty at time step $t$ can be faulty again at time step $t+1$ with a probability $p$. The procedure ends when there is no further faulty components. Analysing the faulty components overall number statistically shows that when the fault propagation rate, which is the average probability over all faults, meets a specific threshold, the distribution of fault probability will be the power law with the coefficient equal to -1.5, which shows a significant risk especially when the system is in critical state (Dobson et al., 2004). Two issues should be considered while applying the branching process model on power systems. The number of elements is limited and also the system fault level should be a function of the fault spread rate. This process is an estimation of CASCADE model (Dobson et al., 2005a, Dobson, 2004). Further, Dobson et al. (Dobson et al., 2005a) showed that the relation between fault level and its propagation rate is linear. Although the branching process is physically significant, i.e. such an approximation captures some salient features of loading dependent cascading failure and suggests an approach to reduce the risk of large cascading failures by limiting the propagation of failures; it fails to consider the power systems dynamics and is limitedly applied to real power systems.

For more clarifications, in CASCADE model in which all items are considered identical and independent from the network structure, the disturbance is distributed indiscriminately. The loading level of each component after disturbance is compared to its value and identifies the state of each element being normal or failed. Likewise, in the branching process, the number of components is considered to be infinite. Each component
generates its next state stochastically. The limitation of this procedure is that it ignores the network structure and the fault spread rate.

### 3.4 Dynamic Cascade Failure Models in Power Systems

The electric power system as a pillar of the economy is a complex infrastructure. Delivery of electric power across long distances is made possible via its inter-connectivity. But, at the same time, it could be turned into power system flaw since the disturbances can be propagated through the same network. In other words, the occurrence of blackouts raised by cascading failures is the direct result of electrical and physical properties of power systems. Cascading failure in power systems as a complex dynamical event is usually triggered by ever-rising demand, increasingly renewables penetration, power system component failures sudden load redistribution in transmission lines, complicated interconnections which sometimes gives rise to inevitable control systems malfunction. Even if this event is managed by protection relays’ immediate reaction by disconnecting the faulty line, it can make other lines overloaded, or busbars work higher than their standard limitations. Dwivedi et al. (Dwivedi and Yu, 2013) proposed a new centrality measure for ranking the power lines in a power system based on the maximum flow algorithm (Ford and Fulkerson, 1956a). They showed that in power systems, there are some power lines more vulnerable to attacks which if removed, there would be a significant drop in power system functionality by shifting their loads to adjacent components.

In practice, usually, a blackout in power systems is a result of a combination of consecutive failures occurring as the subsequent of the initial failure (Baldick et al., 2008, Georgilakis and Hatzigiayriou, 2015, Bialek et al., 2016, Dwivedi et al., 2010). The power system topological vulnerability evaluation models are limited as they only consider the static properties of the networks. To capture the dynamical properties of power systems and discover the procedure of cascade propagation in power systems, a number of dynamical models for power systems have been introduced. Here, we summarise these dynamical models used for studying the cascade mechanism and vulnerability analysis of power systems. The cascade failure dynamic model of power systems is divided into three sub-
categories. OPA model, which is mainly based on optimal DC power flow analysis, the model based on hidden faults and estimation of DC power flow analysis, and also Manchester model which fundamentally applies AC power flows and load shedding.

**ORNL-PSerc-Alaska (OPA) Model**

Load flow analysis methods, AC and DC power flow, are mainly applied to simulate the power system in the steady-state. The OPA model, ORNL-PSerc-Alaska, is a model to study the blackout proposed by scholars from three research institutes (Dobson et al., 2001) using the DC power flow method. In this model, two different dynamics for power systems are considered: slow and fast dynamics. Slow dynamics takes the power system’s long-term extension into consideration and studies the interaction between the load increase and the power system developments. Fast dynamics put the power system day to day operation under scrutiny (Carreras et al., 2001, Carreras et al., 2002a, Carreras et al., 2004, Carreras et al., 2002b, Mei et al., 2009).

In this model, when a single failure happens, the power of each busbar’s intakes is redistributed across the network to make a new steady-state. The only condition is to minimise the difference between power injected in all nodes before and after the failure. Then, the DC power flow is applied to calculate the power in all transmission lines and determine new failures. One of the most critical limitations of OPA model is that since all the calculations are based on DC power flow, to have accurate results, it is applicable only on lightly loaded power systems and is not able to assign the reactive power and voltage profiles. This is a considerable disadvantage since sometimes cascade failures occur in massively loaded power systems.

**Hidden Failure Model**

The idea behind the hidden failure model comes from a fault in a relay protection equipment which cannot be identified in normal operation but appears in special conditions when the system gets faulty (Chen et al., 2005, Tamronglak, 1994). It assumes that any line’s failure could be a result of a protection relay abnormality in its neighbouring lines. In this
model, if a transmission line fails, with a probability proportional to their new loadings, its adjacent lines are likely to fail too. The DC power flow analysis method is applied to assign the lines’ flow. Applying this model on power systems (Tamronglak, 1994), shows that the distribution of blackout exhibits power-law properties. It is worth mentioning that the hidden failure model is very similar to the fast dynamics OPA model. Hence, it is only capable of modelling the lightly loaded power systems and cannot be applied to simulate the power system long-term development.

**Manchester Model**

Despite the previous dynamic model applying the DC power flow, in the Manchester model, the cascade failure is modelled using AC power flow analysis (Nedic et al., 2005, Rios et al., 2002, Kirschen et al., 2004). Because the equations in DC power flow analysis are primarily linear, usually the convergence of this method is not an issue. On the contrary, since calculating AC power flows in power systems needs nonlinear programming with nonlinear constraints, which in some cases does not converge to any solution. Thus, to make sure that the solution exists, the load shedding algorithm is added to this method. The most significant disadvantage of this model is to ignore the slow dynamics like the power system evolitional improvement and considering only the fast dynamics.

To have an idea about these dynamic models, when a failure happens in the network, OPA model applies DC power flow model to redistribute busbars’ incoming power across the network to reach the steady-state. Hidden failure model considers the protection system relays’ hidden faults that may cause the line outage and as a result in this model, neighbouring lines could be faulty. It also applies DC power flow model for the analysis. Contrary to OPA and hidden failure models, the Manchester model uses AC power flow along with the load shedding schemes to overcome the divergence issues. The major point to keep in mind when finding a solution for these models is that divergence is always an issue in nonlinear programmings, such as AC power flow analysis, which highly depends on initial conditions and parameters constraints setups.
3.5 Conclusion

In this chapter, we focused on cascade failures in complex networks and power systems as special case. First, the process of cascade failure in complex networks from initial failure following by load redistribution stage and approaching to final equilibrium point was reviewed. Likewise, cascade failure in complex networks was studied and different models of cascade were reviewed.

There are generally two main categories including structural and component cascading dependent models. In latter section, static, Motter-Lai and Effective efficiency models were summarized. Then in the later category, CASCADE and branching process models were reviewed and pros and cons of each model were stated. Following that, the dynamic cascading models of power systems consisting OPA, Hidden failure and Manchester models were summed up. In OPA model, in case of any failure occurrence, to reach to another steady-state, the incoming power of each busbar is redistributed across the network. Hidden failure model considers the protection system relays hidden faults that may cause the line outage and as a result in this model, neighbouring lines could be faulty. Despite the OPA and hidden failure model which apply DC power flow to calculate power system parameters, Manchester model uses AC power flow along with the load shedding schemes to overcome the divergence issues.

The following two chapters will focus on discovering behaviour of general complex networks during the cascade of failures in their components. The main purpose in these chapters is to figure out how to react to a failure in case it is critical enough to trigger a cascade of failures. To tackle these possible criticalities, we should always have some strategies ready to provide fast responses to such initial failures. The research summarised in chapter four shows that since degree centrality is negatively correlated with the cascade depth, the nodes with such attributes can act as fuses in a critical situations and their removal can be effective to preserve a more significant portion of affected network. To avoid unnecessary complexity and better achieve this objective and fairly compare the results for various complex networks, the dynamics of real networks, such as power networks, are neglected for simplicity, and the fundamental components for all complex networks are
3.5 Conclusion

considered to be the same. On the contrary, in chapter six, the major focus is on power networks, where based on power systems’ dynamic models, a method is developed to deter the cascade from driving the whole network to the blackout state.
Chapter 4

CASCADE FAILURES AND CENTRALITY MEASURES

4.1 Introduction

By discovering the critical nodes in complex networks, one can control the propagation of information in social networks, spreading epidemics in society, and prevent disastrous cascade failures leading to blackouts in power grids or internet outages (Motter and Lai, 2002, Motter, 2004). However, finding vital nodes in a network is not trivial (Lü et al., 2016). At first, the term ‘vital node’ has a vast meaning. Sometimes, we are interested in finding the small set of people whose vaccination can stop the disease from spreading, while sometimes we need to find the most critical busbars which if failed, the power system will experience a cascade of failures. Secondly, to establish a reasonable balance between local and global parameters is challenging. The behaviour of a cascade failure depends on the location of the initial failure, i.e., the node(s)/edge(s) that initially fail. Such a study is missing in the literature, which will be addressed in this work. In this work, we examine a number of node centrality metrics on synthetic and real networks to study the role of each centrality measure in the cascade failure. To this end, we first obtain some centrality measures for the nodes, including degree, betweenness, closeness, clustering coefficient, eigenvalue, information index, lobbying index and local rank. We also consider cascade depth, which is defined as the number of failed nodes as a result of initial failure in a node.

4.2 Related Works

Cascade failures have been studied in the literature including on power grids (Saniee Monfared et al., 2014, Crucitti et al., 2004c, Kinney et al., 2005, Chen et al., 2010). However,
the previous works have mainly focused on modelling cascade failures (Dobson et al., 2005c), designing efficient protection strategies (Koch et al., 2010, Vaiman et al., 2013) or studying the dependence of cascade failures on the structural properties (Wu et al., 2006, Buldyrev et al., 2010) of networks.

Information cascade is a similar topic to cascade failure in the literature of network science. A piece of information, starting from a set of seed nodes, disseminates through the network. Influence maximisation is a well-known problem in this field, which is to find the optimal set of seed nodes such that initially activating them has the maximum influence, i.e., the largest number of finally activated nodes. Corley et al. (Corley Jr, 1974) defined the most vital nodes in a network as those whose failure engenders the highest decline in maximum flow between a particular node pair. In another work, Corley et al. (Corley and Y.S. David, 1982) found the $k$-most vital nodes whose failure makes the shortest distance between two arbitrary nodes the highest possible. Real networks often have community structure. He et al. (J.-L. He, 2015) introduced an approach to find the top-$k$ influential spreaders in networks with community structure. Recently, Jalili and Perc studied the correlation between influence range, i.e., the number of activated nodes as a result of initially activating a node, and centrality measure (Jalili and Perc, 2017). They used independent cascade model and identified the centrality measures showing strong positive correlations. In this work, as a node fails, it might trigger a cascade of failures according to explanations in chapter three. When the cascade stops, the number of failings as a result of initially failed node, is called its cascade depth. Cascade depth is indeed a measure indicating vitality of nodes about cascading failures and the larger values of cascade depth imply the higher criticality of node in cascade failure. Although the role of various node centrality measures on network functions has been studied in the literature (see reviews in (Lü et al., 2016, Jalili and Perc, 2017)), there is no existing work to examine the relationship between centrality measure and cascade failure. The significant contribution of this work is to consider correlation of nodes’ vitality (i.e., their centrality) and their influence on the cascade failures measured by the cascade depth.
4.3 Motivation

In this work, we study how the choice of the initial failure (based on the centrality values of the nodes) affects the cascade depth. Let us give a simple example to show how different choices for the location of the initial failure influences the outcome. Fig. 4.1 and 4.2 graphically show the result of cascade failure in IEEE benchmark 30-bus network. The cascades are performed in two different cases: when a non-critical node (node 25 in Fig 4.1) initially fails and when the failure happens in a critical one (node 6 in Fig 4.2). As it is seen, by failing the non-critical node, majority of the nodes are survived from the failures; whereas removing a critical node results in a dramatic outcome by failing two-thirds of the network. It is worth adding that \( T = 1 \) represents the system’s initial conditions, but the cascading failure is not trivial under this situation. On the contrary, most of the components’ failure cannot trigger the cascade of failures and the cascade chain will stop after spreading only in few (mostly local) steps. In Figs. 4.1 and 4.2, only those few components (nodes 25 and 6) are removed to first examine that if they can trigger the cascade of failure, and second how critical their failure is in terms of preserving the rest of the network. As it can be witnessed from Figs. 4.1 and 4.2, removing node 6, totally dysfunctions the network. But many other nodes cannot trigger this cascade to proceed even for a few steps.

Table 4.1 compares the centrality indices and the percentage of removed nodes as a result of the failure of nodes 6, 21, and 25 for \( T = 1.0, 1.5, \) and \( 1.9 \).

<table>
<thead>
<tr>
<th>Node#</th>
<th>( k_i )</th>
<th>( \rho_i )</th>
<th>( C_l )</th>
<th>( x_i )</th>
<th>( C_i )</th>
<th>al_index</th>
<th>LR</th>
<th>INF_i</th>
<th>T=1</th>
<th>T=1.5</th>
<th>T=1.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7</td>
<td>0.95</td>
<td>1.81</td>
<td>0.003</td>
<td>0.154</td>
<td>4.66</td>
<td>1.0</td>
<td>0.68</td>
<td>0.7</td>
<td>0.63</td>
<td>0.6</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>0.69</td>
<td>1.75</td>
<td>0.0008</td>
<td>1.0</td>
<td>4.49</td>
<td>0.33</td>
<td>0.54</td>
<td>0.33</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>25</td>
<td>4</td>
<td>0.59</td>
<td>1.09</td>
<td>0.107</td>
<td>0</td>
<td>3.0</td>
<td>0.16</td>
<td>0.41</td>
<td>0.3</td>
<td>0.3</td>
<td>0.26</td>
</tr>
</tbody>
</table>
4.3 Motivation

The capacity parameter is set to $T = 1$ for simplicity. A) when a non-critical node (node 25) is removed. (a) The original network with the link width proportional to their load (betweenness centrality measure), (b) initially, node 25 is removed, and as a result the node 26 gets disconnected which are shown in grey, (c) the overall load is redistributed through the network, and as a result of overloading in some edges (displayed as thicker links than the original case), and their removal, the isolated nodes (shown in red), get removed, (d) further load redistribution happens in the network and (e) node 21 gets removed, (f) Finally, in the equilibrium point, the cascade stops and no edge is removed.

4.4 Cascade Depth

To study cascade failures in networks, as is explained in chapter three, we assume that each edge $e_{ij}$ has bounded capacity $C_{ij}$, demonstrating the utmost load which can be delivered by the edge. We suppose that $C_{ij}$ is proportional to its initial load $W_{ij}$, ($C_{ij} = TW_{ij}$), where $T > 1$ is the capacity parameter (Mirzasoleiman et al., 2011, Wang and Chen, 2008). At the outset,
The capacity parameter is set to $T = 1$ for simplicity. When a critical node is removed, (a) the network is working normally and the width of the links is proportional to their load, (b) initially, node 6 (shown in grey) is removed, (c) the overall load is redistributed through the network and as a result of overloading in some edges and their removal, the isolated nodes (shown in red) are removed, (d) further load redistribution happens in the network and (e) node 21 gets removed, (f) Finally, in the equilibrium point, the cascade stops, and no edge is removed.

the initial loads (betweenness centrality in this work) are calculated for all edges. Then, the cascade is triggered by removing the very first node. We choose the initial node for removal based on different centrality measures (e.g., degree, betweenness or closeness); the node with the highest centrality is removed (if there are more than one node with maximal centrality, one of them is randomly removed). Note that when a node fails, all edges connecting to that node are also removed from the network. As the initial node is removed, the betweenness centrality is recomputed, and all nodes with betweenness centrality higher
than their capacity are removed. This process continues until no further removal is required and a steady state is obtained. As a measure of the cascade depth, we calculate the number of nodes removed from the network; the larger the number of the removed nodes, the more severe the cascade. It is worth mentioning that at each step the nodes - to be failed - are discovered and removed at the same time at the end of that time step. We then apply the Spearman rank correlation (Spearman, 1904, Fieller et al., 1957) to find the relation between the cascade depth and the centrality indices. A pseudo-code of the procedure is presented in Procedure 4.1. In next section, we systematically study components’ criticality for cascade failures.

Procedure 4.1: Cascade Failure and Centrality Measures

For a given graph $G$ with $N$ nodes

1: For a certain value for the capacity parameter $T$
2: Calculate the initial load (betweenness centrality) for all edges ($W_j$) and set ($C_j = TW_j$); also calculate all nodes’ centrality parameters (e.g. degree, betweenness). All the centralities are normalised using:

$$C_{i-n} = \frac{C_i - C_{\text{min}}}{C_{\text{max}} - C_{\text{min}}}$$

where $C_i$ is the centrality of node $v_i$ and $C_{i-n}$ is its normalised centrality. $C_{\text{max}}$ and $C_{\text{min}}$ are the maximum and minimum centrality of nodes, respectively. The normalized centralities scale in the range $[0, 1]$.

3: For each node $v_i$, $i=1, ..., N$
4: Remove node $v_i$ and all the edges connected to $v_i$.
5: Recalculate the load of the remained edges and remove those with load larger than their capacity.
6: Repeat 4 until no component gets removed, and the cascade stops.
7: Calculate the cascade depth for node $i$:

$$(\text{Cascade depth})_i = \frac{\text{number of failed nodes as node } i \text{ initially fails}}{N}$$

8: Calculate the Spearman correlation between the cascade depth and centrality metrics.

4.5 Simulation Setup

In this section, we provide details of the numerical simulation set up. As mentioned, the cascade failure process is based on computing loads for edges. In this work, we take betweenness centrality of edges as their loads. To obtain cascade depth for a node, it is initially removed from the network. When a node is removed, all its connecting edges are
also removed from the network. Then, the load (betweenness centrality) of the edges is recalculated, and those with load higher than their capacity are removed from the network. This process continues until no further removal is performed. The cascade depth is the number of failed nodes as a consequence of such a process. Here, we first obtain the cascade depth of all nodes. Then, we obtain the Spearman rank correlation between the cascade depth and centrality measures. To calculate the Spearman correlation between two variables, the Pearson correlation is assessed between their rank values. Unlike Pearson correlation which computes linear relationships, Spearman's correlation computes monotonic relationships. If the observations of two variables have a similar rank, the Spearman correlation will be high with the value close to 1. In case their ranks are dissimilar, this value will be around -1.

We repeat the process for different values of the capacity parameter $T$. Naturally, as $T$ increases, the networks become more tolerant against cascade failures. However, this means that the capacity of the network also increases, which requires a higher cost. Thus, in many cases, it is desired to make the network resilient against cascade failures by designing a minimum $T$.

A positive (or direct) correlation between the cascade depth and centrality measures means that the higher the centrality of a node, the more severe effect its failure has on the network. Whereas, negative (or inverse) correlation indicates that failure in nodes with higher centrality has less effect (i.e., lower cascade depth) than failure in those with lower centrality values.

All the program coding and calculation are performed in Python. To reduce the complexity, the group centrality measures are applied for edges betweenness calculation at each iteration, which is computationally efficient, especially for large-scale networks.

### 4.6 Results and Discussion

In this work, we consider both synthetic and real networks. As synthetic networks, we use BA, WS and ER networks which are explained in section 2.1. Here we also study a number of real networks. We consider eight power networks including Australian high voltage grid, Western US power grid, European power network, UK power grid, French
power grid, Spanish power grid, IEEE 300 Bus and 118 Bus benchmark networks. We also study some other real networks including the Internet in the level of the Autonomous system, US Airport network, Melbourne road network and Melbourne water networks. Table 2.1 represents the structural features of the real networks considered in this work.

In this section, we carry out several numerical simulations to study how the choice of the location of the initial failure impacts the cascade depth. The simulation is conducted on both synthetic and real networks. The results and discussion about the findings are as follows.

### 4.6.1 Synthetic networks

At this stage, to have enough structural diversity over synthetic networks, each network is generated 100 times and the results show averages over these 100 realizations. Also, for each network type, three probability values are considered ($P=0.001$, $P=0.01$, $P=0.1$) to get diverse density values. Another major structural diversity is achieved using various setups over three main synthetic network categories (ER, WS, and BA). For example, a combination of parameters $m$ and $P$ for WS networks or combining $B$ and $m$ for BA networks, is considered.

The synthetic networks are constructed with 1000 nodes and different parameters. As explained above, the simulations results are the average over 100 realizations. Fig. 4.3 shows the results for Erdos-Renyi networks with $N = 1000$ and two connection probabilities ($P = 0.01$ and $P = 0.001$).

Different centrality measures are tested including degree, betweenness, closeness, eigenvector, clustering, average lobby index, local rank, and information index. As it can be seen, three centrality measures have higher correlation values among others. The degree centrality is correlated negatively with the cascade depth, while betweenness centrality and local rank are positively correlated. Information index shows a positive correlation for denser networks (Fig. 4.3b), while its correlation is close to zero for sparse networks. Note that positive correlation indicates that as the value of the centrality measure increases, the cascade depth also increases (i.e., more nodes will fail as a result of a failure in a node).
Cascade Failures and Centrality Measures

Fig. 4. 3 Nodes cascade depth and centralities correlation in Erdos-Renyi networks (N = 1000)
Erdos-Renyi network is constructedler by creating a link between any two nodes with probability P. Different centrality measures are tested including degree, betweenness, closeness, eigenvector, clustering, average lobby index, local rank, and information index. The networks are with N = 1000 and (a) P = 0.01; (b) P =0.001. The data shows average over 100 realisations.

Figs. 4.4 and 4.5 show the results for Watts-Strogatz small-world and scale-free networks, respectively. Again, the degree centrality is negatively correlated with the cascade depth, while betweenness centrality shows a positive correlation. While closeness and eigenvector centrality show almost no relation with cascade depth, there is a small trend of oscillation in clustering centrality in scale-free case. As the connection probability decreases in Watts-Strogatz networks, the average lobby index gets more negative correlation with cascade depth. In higher values of P, there is no meaningful relationship between these two parameters. Local rank shows a positive relationship with the cascade depth, which is in a good agreement with results on another dynamical process (information propagation) (Eguiluz and Klemm, 2002, Thomas and Paolo, 2004). However, as the networks become denser, this local interrelation is overwhelmed by global distribution of edges, and as a result the influence of the local rank in denser networks is lower than their corresponding sparser networks. It is worth mentioning that for Watts-Strogatz networks, the local connectedness (and clustering coefficient) decreases by increasing P. Except a small swing between
information index correlation with cascade depth in lower probabilities of connection in Watts-Strogatz, there is no evidence of connection between these two values in Figs. 4.4 and 4.5.

Fig. 4.4 Nodes cascade depth and centralities correlation in watts-Strogatz networks (N = 1000)

To create the watts-Strogatz network, first, a ring graph in which each node is connected to its m-nearest neighbor is considered. Then, the links are rewired with probability P. (a) m = 2, P = 0.1; (b) m=2, P=0.01; (c) m=2, P=0.001; (d) m=5, P=0.1; (e) m=5, P=0.01; (f) m=5, P=0.001; (g) m=10, P=0.1; (h) m=10, P=0.01; (i) m=10, P=0.001.

There is another trend in Fig 4.4 (g-i), which is worth mentioning in terms of impact of the capacity parameter on correlation trends. This figure shows that in these WS networks (m=10) and for higher density values, as T increases, the betweenness centrality shows significant correlation with the cascade depth. On the contrary, with the density values decreasing, the betweenness shows less relationship with the cascade depth. This trend could be interpreted as saturation in the relationship between the betweenness and cascade depth when the capacity parameters exceed a threshold value as the density decreases.
Fig. 4.5 Nodes cascade depth and centralities correlation in scale-free networks (N = 1000)
Scale-free networks are created using preferential attachment method (a) B=0, m=2; (b) B=10, m=2; (c) B=0, m=5; (d) B=10, m=5; (e) B=0, m=10; (f) B=10, m=10.

Fig. 4.6 shows the correlations for modular networks, where the same phenomenon is observed. To construct modular networks, we first create five distinct modules each comprising 200 nodes, with one of the synthetic methods stated above (Erdos-Renyi, Watts-Strogatz or scale-free). Then, with probability $P$, these modules are connected to form a unified network with $N = 1000$. It can be observed that increasing the probability of connection between the modules weakens the influence of the local rank, which is mainly due to the reduction of local connectedness. The same trend exists for node’s betweenness
4.6 Results and Discussion

4.6.1 Centrality

There is a higher correlation between betweenness centrality and cascade depth in modular networks, whereas it almost fades in lower values of $P$. For all other centrality indices, no connection with cascade depth can be witnessed.

4.6.2 Real networks

Figs. 4.7 and 4.8 show the correlation of the cascade depth with other centrality measures for real-world networks. In Fig. 4.7 we simulate the cascading failure process in eight distinct power grids with a different number of nodes and edges. The results show that four centrality measures are dominant with higher correlation values. While degree centrality has a high negative correlation with the cascade depth in all power grids, the clustering coefficient is negatively correlated with the cascade depth in rather sparse networks. On the other hand, betweenness and local rank are positively correlated, with local rank having higher correlation values than betweenness. Although average lobby index exhibits a noticeably negative correlation with cascade depth especially for smaller power networks, clustering centrality shows negative connection with cascade depth almost in all cases of power grids. Again, there is a slight swing in information index around zero for less dense grids.

![Fig. 4.6 Nodes cascade depth and centralities correlation in Modular networks ($N=1000$)](image)

The networks comprise 5 modules each consisting 200 nodes. These modules are connected with probability $P$. (a) $P=0.1$; (b) $P=0.01$; (c) $P=0.001$. 


Fig. 4. 7 Nodes cascade depth and centralities correlation in Power networks
(a) Australia power grid N = 1130, E = 1426 (total number of links); (b) US Power Grid N = 4941, E = 6594; (c) European Power Grid N = 1494, E = 2156; (d) UK Power Grid N = 327, E = 407; (e) IEEE 300-Bus N = 300, E = 409; (f) French Power Grid N = 146, E = 223; (g) IEEE 118-Bus System N = 118, E = 179; (h) Spanish Power Grid N = 98, E = 175.
Fig. 4.8 shows the results for some other real networks. Again, while the degree is highly negatively correlated with cascade depth of a node, except the airport case, betweenness shows a light correlation with it. A similar pattern is correct for local rank where except the roads network which demonstrates the high connection between local rank and cascade depth, in other cases they are less correlated. Also, it is interesting that information index is almost highly positively correlated in US airport network and the Internet in the level of the autonomous system with less relation for other networks. Finally, clustering
coefficient among other centrality indices displays a small correlation swing around zero for these real networks.

The above results show that in some cases, there are patchy correlations between cascade depth and some particular centrality measures. In Erdos-Renyi networks, for small values of the connection probability, the relationship between the information index centrality and the cascade depth jumps to 0.7. Also, in Watts-Strogatz networks with small rewiring probability, the average lobby index interestingly shows a negative correlation with cascade depth. In power grids, similar behaviour is observed for clustering centrality where it shows a negative correlation in almost all the cases. There are however three centrality measures with almost consistent behaviour across all synthetic and real networks. Although with varying strength, the degree centrality shows negative correlation with the cascade depth, while the betweenness and local rank centrality measures show positive correlations.

These findings can be used to design networks that are resilient against cascade failures. A possible approach to minimise the effects of cascade failure is to do some structural intervention in the network. One can structurally intervene in a network by adding, removing and restructuring components (nodes/edges). Our results indicate that a network can be made resilient against cascade failures by removing the highest-degree nodes. Indeed, one can make a network more resilient by removing nodes that are ranked high regarding those centrality measures showing a negative correlation with cascade depth.

4.7 Conclusion

In this chapter, in order to investigate the impact of initial components’ break on the level of destruction caused by cascade failures, we set up a simulation by studying the effect of nodes’ attributes on the surviving network size when the cascade stops. To have a better scale for comparison, we considered a centrality measure based on cascade failure process (i.e., cascade depth) and studied its correlation with other centrality measures including degree, betweenness, closeness, clustering coefficient, local rank, eigenvalue, lobbying index and information index. To obtain the cascade depth for a node, we first triggered a failure from that node and then got the number of subsequently failed nodes as a result of
cascade failures in the network. The study was carried out in a number of models (Erdos-Renyi, Barabasi-Albert, Watts-Strogatz, and modular) and real (power grids, roads network, natural water network, airport, and the Internet in the autonomous level) networks.

Not all centrality measures have remarkable impacts on cascade failure criticality of a node. Indeed, a considerable number of these parameters do not show any significant correlation with cascade depth and are removed from the final shortlisted centrality variables. The remaining node centrality measures are concluded to have the most definitive effect in controlling the cascade of failures when the strategies for deterring them from happening are planned. Based on our simulation results, it is evident that among all node centrality measures, only three of them, i.e. node degree centrality, betweenness and local rank, show high correlation with cascade depth. Hence, these three centralities can be applied to control the criticality of cascade failure before and after it is triggered. Interestingly, we found that the cascade depth is negatively correlated with the degree, indicating that initial failure in higher-degree nodes will result in lower cascade depth (i.e., less failed nodes as a result of cascade failure process). While, betweenness centrality and local rank showed a positive correlation with the cascade depth, indicating that nodes with higher betweenness and local rank centrality (and not those with high degrees) should be further protected to be resilient against cascade failures. Indeed, to some extent, the nodes with higher degrees, behave like fuses which if removed during the cascade, end up preserving a more substantial part of the network from failure with the price of disconnecting a small partition of it. Our results can be used to mitigate cascade failures in complex networks by removing nodes ranked high in terms of those centrality measures showing negative correlations with the cascade depth.
Chapter 5

CASCADE PROPAGATION STRUCTURE

5.1 Introduction

A solution to deter the cascade failures is to predict the edge(s) which might fail at the next step by tracing the cascade propagation route. Chen et al. (Chen and McCalley, 2005) introduced an approach for identifying the high-risk contingencies known as contingency set as a fundamental stage in monitoring the power system security level (Jian and McCalley, 2001). Phadke and Elizondo et al. (Phadke and Thorp, 1996, Elizondo et al., 2001) introduced a method to detect the hidden failure in protection devices like relays in an online manner to avoid cascade failures. To the best of our consideration, there is no systematic investigation of what exactly occurs when a single or multiple edge removals, trigger a cascade of failure. In this work, we first investigate the impact of single edge failures and discover the spreading route of the cascades in networks. The relation between the cascade propagation pathway and the cascade stage is examined. Then, concurrent edge failures are put under scrutiny to find out how the cascade propagation depends on the topological relationships between the edges from which the cascade initially starts. According to our findings, the complex networks’ response to the single edge failure is entirely different from that of multiple edge failure. It mentions that the strategies to manage these two types of components breakdowns should be designed accordingly.

5.2 Structure of Cascade Propagation

An essential issue in studying cascade failures in complex networked systems is the dynamics of cascade propagation and to analyse the sequence of consecutively failed edges.
This section addresses this issue and examines the pathway of cascade propagation. To have a better understanding of the cascade propagation mechanism, let us first define \( EGD \) (Edge Geodesic Distance), which is a metric to quantify the topological distance between two edges. Let us consider edge \( e_i \) between nodes \( v_i \) and \( v_j \) and edge \( e_2 \) between nodes \( v_p \) and \( v_q \). \( EGD \) between these two edges is defined as:

\[
EGD(e_1, e_2) = \min(l_{ip}, l_{iq}, l_{jp}, l_{jq})
\]  
(5-1)

where \( EGD \) is the Edge Geodesic Distance and \( l_{ip} \) is the length of the shortest path between nodes \( v_i \) and \( v_p \).

Depending on the network topology and the criticality of the initially failed edges, the cascade could last even more than 50 steps, at which the highest loaded edges will be removed from the network, and the cascade process evolves through the next steps. Let us define the cascade step as ‘cascade order’. Under this definition, the initially failed edges have a cascade order of zero. The edge failing in the first step after the initial load redistribution has cascade order of one; the failing edge in the second step has a cascade order of two and so on. Our aim in this work is to study the edge geodesic distances as a function of the cascade order.

To have a reliable statistical estimate, we run the simulations for all possible single edge failures or two-edge and three-edge failures and show the averages over all these values i.e. this is an exhaustive search and the order of removal does not matter. For example, for Spanish power grid with 98 edges, we first remove the edges one by one and study the cascade failure as a consequence of this initial failure which results in utmost 98 distinct lists of \( EGD \) values depending on whether removal of an edge can trigger a cascade failure. These lists generally have different lengths since, for each initial failure, the cascade may last for different cascade orders. Finally, for each cascade order \( r \) the average over all lists is calculated as the propagation route.

A pseudo-code of the cascade propagation algorithm when a single edge removal triggers a cascade of failures is provided in procedure 5.1. The algorithm begins by initially failing an edge \( e_i \). All edges removed in next steps of the algorithm (including the initially removed edge \( e_i \)) are recorded in \( \gamma \) that is a list which keeps track of the removed edges according to the cascade order. When the initial failure happens, the load is redistributed all
over the network, and the new loads are calculated for the edges. The load computation method for all lines in complex networks is edge betweenness centrality as per equation 2-16. Edge betweenness centrality is one of the most common parameters for the weight allocation (Mirzasoleiman et al., 2011, Zhu et al., 2012). Intuitively, the higher the loads in a line is, the bigger will be the value of edge betweenness for that line. The highly loaded lines in the complex networks demonstrate bridge-like components between two different sections of the network and their failure may affect the information delivery at substantial portion of the network. Following the load calculation for each edge, the edge with the highest load ($e_h$) is identified, and if its load is higher than the predefined threshold value $T$, it is removed from the network and added to $\gamma$. Let us denote the edge added to $\gamma$ at time $t$ by $\gamma(t)$. This process is iterated until no more edges are removed, where the cascade stops. Then, $EGD$, as defined in equation (5-1), is calculated between edges removed in consecutive orders. Let us define $\sigma_i(t)$ as the $EGD$ calculated between $\gamma_i(t)$ and $\gamma_i(t+1)$. This cascade failure process is repeated by selecting all edges, one-by-one, as the initially failed edge. Let’s suppose that the highest cascade order is $t_h$. Indeed, $t_h = \text{maximum cardinality of } \sigma_i(t)$. We then calculate the cascade propagation route as

$$\mu(t) = \frac{1}{m} \sum_{i=1}^{m} \sigma_i(t) ; \quad t = 1 \ldots t_h$$

(5-2)

Let us make an example of how this method works. Suppose that we have three edges ($e_1, e_2, e_3$) initially failed and removed from the network, one at a time, and the corresponding lists of consecutively removed edges for them ($\sigma_1, \sigma_2, \sigma_3$) are as $\sigma_1= [0, 1, 1], \sigma_2= [0, 2, 2], \text{and } \sigma_3= [1, 2, 4, 4]$. For example, the values in $\sigma_1$ show that after removal of $e_1$, the $EGD$ between the edge which is immediately failing and $e_1$ is zero implying that they share a node and so on. It is clear that for this example $t_h = 4$ representing the highest cascade order in $\sigma_3$. These explanations lead the cascade propagation route to be as:

$$\mu(t) = \left[\frac{0+0+1}{3}, \frac{1+2+2}{3}, \frac{1+2+4}{3}, \frac{4}{1}\right] = [0.33, 1.67, 2.33, 4]$$

It implies that for example at cascade order $t = 1$ the average of $EGDs$ for all failed edges (those in $\gamma$) is 0.33. At cascade order $t = 2$ this value is 1.67 that is higher than very first cascade order.
5.2 Structure of Cascade Propagation

Procedure 5.1: Cascaded Propagation Algorithm for Single Edge Failure
Given a connected graph $G = (N, E)$ and a threshold value $T$

1: $\mu$ is an empty array
2: For any $i$ and $e_i \in E$
3: $\gamma$ is an empty array to record the list of removed edges
4: $e_i \rightarrow \gamma$ save the initially failed edge in $\gamma$
5: $(G-e_i) \rightarrow G$ removes the $e_i$ from the graph
6: Recalculate load of each edge ($L$) using equation 2-16,
7: Find the highest loaded edge $e_h = \text{argmax} (L)$
8: If $L_h > T$
9: $e_h \rightarrow \gamma$ record the next –to be failed – edge
10: $e_h$ is removed from the graph $(G- e_h) \rightarrow G$
11: $\sigma$ is an empty array to record the $EGDs$ between the edges consecutively saved in $\gamma$
12: For item $\gamma_i(t)$ corresponding to the cascade order $t$
13: $\sigma(t) = \text{EGD} (\gamma_i(t), \gamma_{i+1}(t))$
14: Arrange the $\sigma_i$ lists according to their length (The longest first)
15: For $t \in [0: t_h]$, $(\text{number of lists (}$ $\sigma$ $\text{) with length } \geq t) \rightarrow m$
16: Calculate $\mu(t)$ using equation 5-2

$\mu$: cascade propagation route
$e_0$: an initial single edge failure
$\gamma$: the list of removed edges
$e_h$: the highest loaded edge
$\sigma$: a list of $EGDs$ between two consecutive edges in $\gamma$
$t$: cascade order
$t_h$: the highest cascade order

5.2.1 An Illustrative Example

In this section, an example is provided to explain how the cascade is propagated when a single edge is removed from the network. Let us consider the IEEE 118-bus benchmark network as an example. Fig. 5.1a shows the portion of the network that is affected by the failure. First, the transmission line connecting the buses 59 and 63 fails and is removed from the network. This failure causes the shortfall power to be compensated by the rest of the network, which in turn causes an overload in other transmission lines, resulting in their
possible failure. Not to mention that failure of any overloaded line depends on its tolerance against overload. This parameter is identified as a threshold value. In this example, as it can be seen, the line between buses 63 and 64 is the first line affected by this overload; it fails in response to the initial failure of the links between buses 59 and 63, with cascade order of 1.

Likewise, at the next cascade order, another power line connected to the busbar 64 is overloaded and fails, having cascade order 2. Fig. 5.1 b shows $\mu$ as average $EGD$ against the cascade order for this example network. This figure indeed describes the topological relationships within the consecutively affected edges at each cascade order when the failure initially starts from the edge between buses 59 and 63. Explicitly, by removal of this edge, the edge geodesic distances for the first order is zero as lines 63-59 and 63-64 share busbar 63. At the second cascade order, the removed edge shares the busbar 64 with the line 63-64, and thus the edge geodesic distance is also zero for this case. Note that in studying the cascade propagation route, the edge geodesic distance always is calculated between two
5.2 Structure of Cascade Propagation

Consecutively failed edges. It is evident that initially the cascade is propagated locally in the close neighbourhood of the removed edge with lower cascade orders. However, as cascade evolves to higher orders, it is spread in areas much farther than the vicinity of the initially failed edge.

Procedure 5.2: Cascaded Propagation Algorithm for Concurrent Edge Failure

A connected graph $G = (N, E)$ and a Threshold value $T$

1: $\zeta$ is an empty list, $\mu$ is an empty array
2: For $\lambda \in [0: \lambda_h]$ ($\lambda_h$: the highest possible EGD which could be the graph diameter)
3: $\zeta(\lambda) = \emptyset$, $\zeta$ contains many empty lists
4: For any $F_0 \subseteq E$
5: Calculate $\lambda_0$ (average EGD between edges in $F_0$)
6: $\gamma$ is an empty array to keep track of removed edges
7: $F_0 \rightarrow \gamma$
8: $(G-F_0) \rightarrow G$ $F_0$ is removed from the graph
9: $e_h = \text{argmax} \; (L)$
10: If $L_h > T$
11: $e_h \rightarrow \gamma$
12: $(G-e_h) \rightarrow G$ The highest loaded edge is removed
13: Repeat 9 until no more edge fails, and the cascade stops.
14: $\sigma$ is an empty array to record the EGDs between the edges consecutively saved in $\gamma$
15: For item $\gamma(t)$ corresponding to the cascade order $t$
16: $\sigma(t) = \text{EGD} \; (\gamma(t), \gamma(t+1))$
17: $\zeta(\lambda_0) + \sigma \rightarrow \zeta(\lambda_0)$ according to $\lambda_0$, the $\sigma$ is aggregated to the corresponding list
18: For $\lambda \in [0: \lambda_h]$
19: $m = \text{length} \; (\zeta(\lambda))$
20: Calculate $\mu (\lambda)$ using equation 5-3

$\zeta$: cascade route accumulator

$\lambda_0$: Average edge geodesic distance between initially failed edges

$\lambda_h$: the highest $\lambda_0$ which could be equal to power system graph diameter

It is worth clarifying that we examine the way in which a cascade of failures propagates in a complex network. The network selection is independent of the network type. The study in this chapter is conducted in general complex network context and is not limited to power
networks. No real complex network dynamics are reflected in simulations as all the nodes are considered to be identical. The aim is to discover the next area where the cascade might be propagated into. Therefore, Fig. 5.1 can be simply replaced by the cascade propagation route diagram for any other real or synthetic complex network.

A procedure similar to the above is performed for the cases of simultaneous two- and three-edge removals. The algorithm starts by initially failure $F_0$ on the network, where $F_0$ can be any possible combination of two or three edges over the network. Here, we calculate $\lambda_0$ which is the average edge geodesic distance between the edges in $F_0$. The rest of this process is similar to procedure 5.1 except at the final stage where we calculate the average edge geodesic distance for the cascade propagation versus $\lambda_0$. We then compute the cascade propagation route as

$$\mu(\lambda) = \frac{1}{n} \sum_{\lambda=0}^{\lambda_h} \zeta(\lambda)$$

(5-3)

### 5.3 Simulation Results

In this section, the cascade propagation methodology is applied on some synthetic networks as well as real-world networks ranging from biological networks to road networks, from power grids to airport and natural water distribution network. In this section, the edge betweenness is used to evaluate the centrality of edges as an identical meter. Table 2.3 briefs the number of nodes and edges in the network under consideration.

This section provides the numerical simulation results for the cascade propagation route where the cascade failure is induced by single or multiple edge removal. Since the trend for single edge removal and concurrent edge removal are entirely different, their simulation results are presented separately. For both cases, the simulations are conducted on benchmark networks introduced in section 2.4.
5.2 Structure of Cascade Propagation

Fig. 5. 2 Cascade propagation route for Synthetic networks, triggered by single edge removal. The networks are ER, WS, and BA random networks. The error bars show the standard deviation. After iterating the single edge removal algorithm (procedure 5.1) on all edges, for each cascade order, the average and standard deviation of the edge geodesic distance is calculated.
The networks consist of Spanish, French, UK and Australian power grids. Other designations are as in Fig. 5.2.

5.3.1 Single Edge Removal

In this section, the local propagation pattern of cascade failures caused by single edge removal is studied. To this end, for each edge in the network initially an edge is removed at a time, and the edge betweenness is applied to calculate the load of all remaining edges. Then, the highest loaded edge ($e_h$) is identified, and if its load exceeds its nominal capacity (which is set using the threshold value), it fails and consequently gets removed from the
5.2 Structure of Cascade Propagation

This process keeps on going until no further edge is overloaded, where the cascade stops. We then use the algorithm introduced in Procedure 5.1 to compute the edge geodesic distance, as expressed by equation (5-2), which identifies how the cascade propagates. This algorithm is repeated for every single edge in the network, and finally, the average EGD is reported as the route the cascade propagates. A small value for the average EGD implies that the cascades propagates primarily into the local neighbourhood, while its large value indicates global propagation of the cascade.

Fig. 5.4 Cascade propagation route for Biological networks triggered by single edge removal
Networks: Macaque cortical connectivity, Cat cortex, and C. elegans neural network. Other designations are as Fig. 5.2
Fig. 5.5 Cascade propagation route for real networks triggered by single edge removal. The networks consist of Melbourne road and natural water, and US Airport networks. Other designations are as Fig. 5.2.

Fig. 5.2 to 5.5 represent the cascade propagation profile for Synthetic, power grids, biological and real networks, respectively. It is seen that the cascades propagate locally for small cascade orders, while it tends to propagate globally throughout the network by increasing the cascade order. For lower cascade orders, small average EGD implies that when the failure of an edge triggers a cascade, the edges in its neighbourhood are the most vulnerable ones and are more likely to fail. Therefore, the local vicinity needs further protection. From the power system perspective, to keep the power flowing to the loads in the affected areas, power should be delivered from the generators closest to the loads. This process would minimise the unwanted effects of cascade propagation and is only possible
5.3 Simulation Results

through the transmission lines adjacent to the failed line. As the cascade keeps ongoing and more extensive areas are involved, it becomes difficult to predict which part of the complex network is the next to be affected. In this stage, the cascade aggressively propagates almost all over the network, which may result in an uncontrollable sequence of failures, and large-scale blackouts in power system case.

In Fig 5.2, the diameter of both ER and BA networks is 4. Any of these numbers are averages over lots of EGDs and are generally independent. An initial small average $EGD$ implies that when the failure of an edge triggers a cascade, the edges in its neighbourhood are the most vulnerable ones and are more likely to fail. As the cascade keeps on going and more extensive areas are involved, it becomes difficult to predict which part of the network is the next to be affected. In this stage, the cascade aggressively propagates almost all over the network. Regarding the oscillation in BA networks, the smaller values belong to highly connected nodes which takes time (cascade order) for the cascading failure procedure to remove all their connection, and when the cascade propagates to other areas, there is an increase in average $EGD$.

In Fig 5.3, in higher orders around 12 and 13, a spike in average edge geodesic distance can be seen, which belongs to the last connections in the network and happens when the network has lost its connectivity and become totally dysfunctional.

5.3.2 Concurrent Edge Removal

In some cases, more than one edge failure might happen at the same time. Due to the fact that two failures cannot be exactly concurrent in real-time, one could be the consequence of another one that has not been managed properly. This could be a result of initial single edge failure which is not managed properly and leads to another failure or a fault in control system which gives rise to protection systems failures. For such conditions, we use the pseudocode provided in Procedure 5.2. Depending on the edge geodesic distance between the initially failed edges, the behaviour of the affected network might change. Fig. 5.6 to 5.9 show the result for two and three concurrent edge failures in benchmark networks. These networks exhibit an interesting behaviour; when the edge geodesic distance between initially
failed edges are small, meaning that the fault is happening in a small area by removing two or three edges at the same time, the average cascade steps grow in the opposite direction, which implies that the cascade propagates in farther areas.

Fig. 5. 6 Cascade propagation route in synthetic networks (ER, WS, SF) in concurrent edge failure cases. All the synthetic networks consist of 200 nodes and almost 1000 edges. The structural properties of this network can be found in Table. 2.3. Distance between initially failed edges depends on the network diameter which is the highest distance between any pair of nodes in the network. As the initially failed edges are close together, the cascade spreads in farthest areas as it is evident from higher average globality values. Then up to a breakpoint, this trend keeps ongoing. The cascade behaviour totally changes after this point where two or three concurrent failures happen too far from each other and cannot boost each other’s destructive impact. That’s the reason the cascade propagation route follows the single line failure inclination.
5.3 Simulation Results

Fig. 5. 7 Cascade propagation route in Australian, UK, Spanish, and French power grids in concurrent edge failure cases. In some power systems, the network’s diameter is a bigger value, so in horizontal axis there is distance between initially failed edges up to 25. Cascade tends to propagate in farther areas as the edges are so close together up to a threshold point (8 for Australian, 9 for UK, 3 for Spanish, and 6 and 5 for the French power grid in two and three contingency cases respectively). Thereafter, the cascade shows the single edge failure behaviour.

But for each network, it seems that there is a threshold distance between initially failed edges and increasing this distance from zero up to the threshold value will decrease the area that could be affected by the cascade. After this threshold, the trend is totally changing as the cascade tends to spread in bigger area compared to the values smaller than the threshold.
To quantify how aggressively the cascade triggered by concurrent edge removal could grow, we define the average cascade globality parameter. This parameter is the indicator of the route and distance the cascade will go at next step. Not to mention that the average globality value is calculated using $EDG$ introduced in equation (5-1).

Fig. 5. 8 Cascade propagation route in biological networks in concurrent edge failure cases. In all three networks (C. elegans, Cat cortex, and Macaque cortical), the cascade breakpoint is 3 for two concurrent edge removals and 2 for three concurrent edge removals. After this threshold the network treats them like two or three single-edge removals, which the trend in all figures shows such behavior.
A large cascade globality value signifies that from one cascade order to the next one, the cascade takes big jumps to spread, making the anticipation of next – to be failed – component or even area almost impossible. In other words, the higher is the value of this parameter, the more global (and severe) is the cascade. By increasing the average distance between the initially failed edges, the cascade propagation radius decreases up to a particular point where the average initial contingency hits the breakpoint (e.g. 2 for ER case).

By increasing the distance between the initially failed edges further than this, a trend similar to that of single edge removal case is observed. Indeed, when the distance between the concurrently failed edges is more than this breakpoint, their concurrent failure is not boosted by superimposing of these two or three single-edge failures at the same time.

Fig. 5.6 shows the cascade propagation route in synthetic networks (ER, WS, SF) when concurrent edge failure happens in the network. There are 200 nodes and around 1000 edges in all three types of synthetic networks. In table 2.3 the structural features of these networks are summed up. Not to mention that the distance between initially failed edges can be as high as the network diameter which is the highest distance between any pair of nodes in the network. While the initially failed edges are structurally near each other, the cascade spreads in farther areas since it is obvious from higher average globality values. Then up to a breakpoint (which is 2 for all cases except two concurrent edge failures in WS which this value is 3) this direction keeps going. The cascade behaviour totally changes after this breakpoint where two or three concurrent edge removals happen too far from each other and are not able to uplift each other’s destructive impact. That is the reason the cascade propagation route follows the single line failure trend.

Fig 5.7 shows the cascade propagation route in four different power systems (Australian, UK, Spanish, and French power grids) while two and three concurrent edge failures are imposed on them. Since in some cases the network diameter is as big as 25, in this figure, there are bigger distances between initially failed edges. Again, cascade tends to propagate in farther areas as the edges are so close together up to a threshold point (8 for Australian, 9 for UK, 3 for Spanish, and 6 and 5 for French power grid in two and three contingency cases respectively). Thereafter the cascade shows the single edge failure inclination.
Fig. 5. Cascade propagation route in real (infrastructure) networks in concurrent edge failure cases. In some networks, the distance between initially failed edges is so significant which implies that the network diameter is a high value. The threshold in US Airport network is 3 and 2 for two and three concurrent edge removals respectively. These values for Melbourne natural water are 13 and 10, whereas that of Melbourne roads are 10 and 8. After these breakpoints, the cascade behaves like single edge removal.

Fig. 5.8 pictures the route the cascade is propagated in biological networks (C. elegans, Cat cortex, and Macaque cortical) when they are exposed to two and three concurrent edge failures. For all three networks, the cascade breakpoint is 3 for two concurrent edge removals and 2 for three concurrent edge removals. After this threshold the network treats them like two or three single-edge removals, which the trend in all figures show such a behaviour.
5.3 Simulation Results

Fig. 5. 10 Cascade propagation route for IEEE-57, IEEE-118, and IEEE-300 triggered by single edge removal. In this case, despite the rest of this chapter so far, the Ford-Fulkerson method is applied to calculate the edges' loads. Red bars represent the average edge geodesic distance for different cascade orders. The error bars show the standard deviation. After iterating the single edge removal algorithm (procedure 5.1) on all edges, for each cascade order, the average and standard deviation of the edge geodesic distance is calculated.
Fig. 5.9 shows the cascade propagation route in concurrent failure cases in real-infrastructure-networks (US Airport, Melbourne natural water and roads). For US airport network when two concurrent edge failures happen, the breakpoint is 2, while that of Melbourne natural water and roads are 13 and 10 respectively. Also, in three concurrent edge removal case, these values are 2, 10 and 13 respectively depending on the network structure. Past this critical point, the concurrent failure propagation shows single edge removal trends. When the initially failed edges are close to each other, the cascade spreads in farther areas as it is evident from higher average globality values. Then, up to a breakpoint (dip point), this trend keeps going. The cascade behaviour changes after this point where two or three concurrent failures happen too far from one another and cannot boost each other’s destructive impact. That is the reason why the cascade propagation route follows the single line failure inclination.

5.3.3 Cascade Propagation Route Based on Ford-Fulkerson Method

In this section, the cascade propagation methodology introduced in section 5.2 is applied on IEEE 57, 118 and 300 (University, 2003) bus networks. Despite the previous sections that for monitoring the cascade propagation route, the edge betweenness is applied to evaluate the load of edges; in this section, the maximum flow method based on Ford-Fulkerson explained in section 2.3.5 is used.

Simulation results for cascade propagation route when it is triggered by single edge removal is presented in figures 5.10 to 5.13. Again, the most prominent yield from this set of simulations is their conformation to the trends previously observed from cascade propagation route in benchmarks introduced in section 2.4, in spite of the difference in the mechanism of load calculation. According to these results, in all three IEEE test cases, the cascade started by the removal of a single line tends to spread in the vicinity of the initially failed edge; then in higher cascade orders it is propagated in farther areas across the network. On the other hand, when the cascade is triggered by concurrent edge removal, as the initially removed edges are topologically close enough, the cascade tends to spread in farther areas since they can boost each other’s destructive impact and cause severer cascades.
5.3 Simulation Results

Fig. 5. 11 Cascade propagation caused by (a) 2 and (b) 3 concurrent edge failures in IEEE-57. For this network as it can be seen, the breakpoints for two and three concurrent edge removals are 6 and 5 respectively and after this point, the network shows a single edge removal sentiment.

For IEEE 118 bus system, the breakpoints for two and three concurrent edge removals are 8 and 7 respectively, implying that before these points, the concurrent edge removal incidents uplift each other's effect and thereafter the networks show a single edge removal impression.

For large distances between the initially failed busbars, they can be treated as two separate cascades started from different locations of the network, and thus will also result in a severe cascade. Therefore, it is crucial that when a failure is detected in one of the links, its immediate neighborhood is further protected. This may effectively stop the cascade in its early stages.
Cascade Propagation Structure

Fig. 5. 13 Cascade propagation caused by (a) 2 and (b) 3 concurrent edge failures in IEEE-300
Breakpoints for both concurrent edge removals are 8, meaning that from zero up to these points the concurrent edge removal events increase each other’s destructive impact. Then, after that, the networks show single edge removal trend

5.4 Conclusion

In this chapter, we studied the cascade propagation patterns in networks. It investigated how cascade failures triggered by initially a single or two- and three-edge concurrent failure propagates throughout the network. To this end, first a set of edges fails, then the overall load of the network is recalculated, and the edge with the highest load fails if its load is more than the predefined threshold value. To calculate the components’ loads in the network, both edge betweenness and maximum flow-based centrality measure were adopted. Both these methods are adopted by many scholars to allocate the edge loads. Not to mention that highly loaded lines behave as bridges in complex networks and their removal will potentially disconnect a significant part of the network. The process continues until no edge is overloaded, where the cascade stops. As the number of edges sequentially failing as a result of this process increases, the cascade becomes severer. We adopt a wide range of synthetic and real networks as benchmark networks. When a single edge initially fails, the cascade evolves locally up to a particular point, and thereafter it tends to spread globally in farther areas all over the network. It gives the network planners and operators the insight that as an edge initially fails, the edges surrounding its neighbourhood should be protected as soon as possible before the contagion infects the farther areas and get out of control. When two or three edges concurrently fail at the initial step, somewhat different behaviour was observed.
If the initially removed edges are close, they can boost each other’s destructive role, and cause the cascade to spread in farther areas. Also, as the initially failed edges get distant from each other, the cascade pervades in these edges’ close neighbourhood. To conclude, in order to make a complex network resilient to cascade of failures caused by single edge breakdown, it is vital to manage such a failure carefully and deter another component from failure to avoid the situation from becoming more complicated. This can be accomplished by introducing new protection policies or even putting tighter protection constraints in its neighbourhood. However, the network’s condition is already intricate when concurrent edge removal happens. To protect a complex network from cascade of failures when such excursions happen, more comprehensive studies should be conducted to scrutinize all the possible dynamics in complex networks.
Chapter 6

VULNERABILITY AND LOAD REDUCTION IN POWER NETWORKS

6.1 Introduction

When a power system is in operating condition, it might sustain many changes in its states (Dobson et al., 2005c, Motter, 2004, Shekhtman et al., 2016). The power grid may experience a generator disruption, demand increasing, outage of transmission lines or even equipment failure (Talukdar et al., 2005). Modern networks are increasingly getting connected to less predictable intermittent energy sources, i.e. wind energy which was not considered in initial power grid design. As a result of all these situations, the power grid may have to operate close to its stability limits (Koch et al., 2010). An immediate consequence of this could be a sudden change in the power flow in particular transmission lines, bus voltage angles change and so on (Pahwa et al., 2013b). Some types of overloads in transmission lines might lead to lines’ failure. Repetitive redistribution of the load all over the network may push the power system to end up with a blackout (Carreras et al., 2002b).

Power transmission lines are much more vulnerable than busbars since they are physically exposed to the public access in much broader areas, and also remote cyber-attack happens to them more often. On the other hand, since the substations are always protected physically in a closed area, there are narrower types of attacks to them compared to transmission lines. This makes the literature limited in the field of vulnerability study of busbars (Solé et al., 2008a, M. C. Rosa, 2007, Albert et al., 2004, Holmgren, 2006, Wang and Rong, 2009). In the first section of this chapter, we analyze the vulnerability of a synthetic and a real power system against a cascade of failures and conclude with a solution to prevent the cascade failure to lead to a catastrophic blackout.
Shedding some loads is a very effective measure to return the power grid to the stable state. This helps maintain the generation-load balance (Hines et al., 2009). Otherwise, the infected part of the network should be islanded (Koch et al., 2010, Vaiman et al., 2013, Spiewak, 2016, Qi et al., 2015). For a long time, the load shedding has been applied as a method of mitigating the overload cascade failure (Aponte and Nelson, 2006, Pahwa et al., 2013a). But, all these methods lead to customers discomfort and power cut off. In the second section of this chapter, we propose a method to mitigate cascade of failures in case a power line is getting overloaded.

### 6.2 Power System Vulnerability

A number of methods have been proposed to calculate realistic loads for power networks (Zhu et al., 2014, Zhu et al., 2012, Zhu et al., 2013). The method based on maximum flow (Freeman et al., 1991, Ford and Fulkerson, 1956b), explained in section 2.2.3, is an appropriate way of obtaining a load of the networks, which will be used in this work. Dwivedi et al. (Dwivedi and Yu, 2013) investigated the vulnerability of power systems by proposing a maximum flow-based complex network approach. They also introduced a centrality index to identify the vulnerable lines based on maximum flow and ranking them according to their removal effect on deteriorating the power system performance and triggering a pervasive cascade of failures. Moussa et al. (Moussa et al., 2018) investigated the vulnerability of power system interdepending with the communication network to failure in their single or multiple components. They found that the power system is highly vulnerable to the loss of a small set of links. Yang et al. (Yang et al., 2017) identified, quantified and analyzed a small but topologically central set of power system components which are vulnerable under multiple conditions. If the initial failure happens in the vicinity of this set, the power system is more likely to undergo a cascade of failures. Albert et al. (Albert et al., 2004) studied the power systems from the network perspective and investigated the power delivery from generators to loads when specific nodes are attacked. They simulated different attack scenarios including random, degree based, and load-based scenarios and concluded that removal of the highest loaded nodes by recalculating the load
in remaining nodes at the end of each ten steps has the most destructive effect on the power system functionality. Duenas et al. (Dueñas-Osorio and Vemuru, 2009) modelled the overloads due to cascading failures using the tolerance parameter $T$ which is the networks elements flow capacity relative to their flow demand. They also followed the cascade in the power systems step by step and found that at the first time-step right after the disruption; the most significant drop in the network functionality occurs. They derive that to improve the power system robustness only updating the $T$ is not enough and the topology of the network should be improved to decongest, decentralise the system and increase the number of alternative paths in case of any failure occurrence.

In this section, to study cascade failures in power networks, we consider the IEEE 118 bus network, shown in Fig. 2.10, and a proper synthetic spatial network model, introduced in section 2.3.2, as benchmarks. We then apply the Ford-Fulkerson method, to calculate the maximum flow of lines and find the vulnerable edges in the network. Then, the influence of random and targeted cascaded failures is investigated.

6.2.1 Power System Vulnerability Analysis

To examine the power system vulnerability, both real and synthetic model networks are applied. As a real network, we consider IEEE 118 bus network (Chen and McCalley, 2005, Machowski et al., 2011), consisting of 19 generators, 9 transformers, 186 lines and 91 loads as shown in Fig. 2.10. We also introduce a synthetic model to construct a spatially extended power network with the source, sink and intermediate nodes which is explained in section 2.3.2.

To study cascade failures in the networks, we apply the method introduced in section 3.2, $(C_{ij}=TW_{ij})$, where $T>1$ is the capacity parameter. To study the network robustness outline in cascading failures, its behaviour is investigated as a function of the threshold parameter $T$. To do so; we first remove an edge (randomly or the one with the highest centrality). Then, the maximum flow algorithm is performed again on the network, and the loads are obtained. If a load of an edge is higher than its capacity, the edge fails and is
removed from the network. The process continues until no further failure happens in the network.

We study a number of network parameters in the process of cascading failures. To this end, the average maximum flows, the normalised number of removed edges $S_N$, and the size of the giant component (i.e., the number of nodes in the largest connected component of the network) and global efficiency are studied. The flow is delivered through a network via the connecting links.

An important parameter specifying the influence of cascading failures for a network is the critical capacity parameter $T_c$. When $T > T_c$, no cascading failure occurs, and the system functions normally. Indeed, $T_c$ is the minimum value of protection strength to evade cascading failure, i.e., the network is resilient against cascading failures with minimum cost by setting the capacity parameter as $T_c$.

---

Fig. 6. 1 Influence of single edge random failure on cascade failure in IEEE-118
Top left: SN the normalised average number of the removed edges. Top right: Maximum flow calculated using the Ford-Fulkerson method. Bottom left: Normalized number of survived bus bars after cascade stops (the size of the giant component). Bottom right: Network global efficiency. In all figures, the x-axis is the capacity parameter T. Data shows averages over 10 realisations.
6.2.2 Simulation Results

First, the influence of cascade failures in the IEEE 118 bus network is studied by initially removing an edge randomly from the network. The process is repeated 10 times, and the average profiles are reported in the results (Fig. 6.1). As per the figure’s caption, the x-axis is the capacity parameter $T$. Vertical axis parameters are explained in figure titles as the average number of the removed edges, maximum flow, the normalised size of the giant component, and the network global efficiency. The fitted line is simply a general indicator of the parameters change trends before and after the critical capacity $T_c$. For a specific value of $T$, after one of the links is randomly disconnected, the cascading failures are studied, and the performance measures are calculated. As it is seen from Fig. 6.1, at $T_c=1.019$ the phase transition occurs, and the number of failed edges ($S_n$) suddenly falls from more than 40% to zero. Likewise, the maximum flow rises to around 1 from zero. Eventually, the network efficiency reaches the initial value ($\approx 0.21$) after $T_c$.

![Graphs showing simulation results](image)

Fig. 6.2 Targeted attack to the edge with the highest normalized centrality in IEEE-118. All designations are as for Fig. 6.1.
In the second experiment, we study how a targeted attack triggers a cascade of failures in the network. To this end, the edge with the highest maximum flow-based centrality, as calculated by equation (2-24), is removed from the network, and the cascaded failure procedure is run. The results are shown in Fig. 6.2. As can be seen, at $T_c = 1.041$, the phase transition occurs, and $S_N$ suddenly falls from more than 60% to zero. The maximum flow rises to around 100% from zero. For $T < T_c$, after the attack to the most vulnerable edge, hardly half of network nodes could survive, but with $T > T_c$, almost none of the busbars fail. The critical capacity parameter for a targeted attack is higher than a random failure, which means that the cost of network protection against attacks is higher than random failures.

Fig. 6.3 Impact of random attack (fault) to one edge of the synthetic networks (see text for explanation of the model and the parameters set up). Other designations are as for Fig. 6.1.
Fig. 6.3 displays the behaviour of the synthetic model network when one of its edges is randomly removed. For this case, the critical capacity parameter is $T_c=1.007$, which demonstrates better robustness compared to $T_c = 1.019$ for IEEE 118 bus network.

![Graphs](image1)

**Fig. 6.4** Targeted attack to the edge with the highest normalised centrality in the synthetic network. Other determinations are the same as Fig. 6.1.

Fig. 6.4 shows the performance of the same model network exposed to a cascaded failure caused by a targeted attack on the most central edge in the model network. For this case, we discover the critical capacity parameter as $T_c = 1.018$, which implies that the model network is less tolerant than IEEE 118 bus network with $T_c = 1.041$. Not to mention that again, in the case of targeted attacks, the general behaviour of the metrics with respect to the capacity parameter is similar in two networks.
These results show that IEEE 118 bus network has better tolerance than the synthetic spatial power networks, which indeed indicates intelligent engineering design for this network. However, its robustness can still be improved by some structural interventions, which can be a relevant subject for future studies.

6.3 Load Reduction in Power Networks

Here we propose a method to rank the loads to be reduced to tackle the overload in a particular line which otherwise might drag the power system to instability. First, we perform the load flow analysis to evaluate initial conditions just before the occurrence of an overload in the transmission lines. Then, for a particular line working close to its loading limitation, the algorithm will nominate a few loads which are fed through that line. Now, using the tree traverse method, only those loads which absorb the higher portion of their power from that line will be finalized to be ranked. Then, using a set of a partial differential algorithm, the sensitivity of line power to the power change in those loads will be evaluated. The higher is the sensitivity value for a load, the more sensitive would be the line power to the power change in that specific load and also the higher would be the rank of that load. Load reduction algorithm begins from the top-ranked loads and reduces them hierarchically until the power flow in the line gets back to a normal range.

6.3.1 Sensitivity Analysis and Load Reduction

In this section, the basis for sensitivity analysis and loads ranking is explained. We begin with power flow in a single transmission line and then using power flow equations the algorithm is extended to find the particular line’s power sensitivity to any arbitrary load’s change.
6.3.2 Sensitivity Parameters

To consider an accurate flow dispatch, (Peschon et al., 1968) proposed the first-order sensitivity equations which then adopted by others. Huang et al. (Huang and Yao, 2012), introduced a new Jacobian-based distribution factor to solve the complex power flow in real-time. In this section, in order to deter a nearly overloaded power system transmission line from failure, a sensitivity analysis method is introduced to rank the loads to be reduced.

Fig. 6.5 Line I connecting buses vi and vj

Fig. 6.5 shows the diagram of a single transmission line connecting busbars vi and vj. Iij shows the current delivered from busbar vi to busbar vj. R and X represent the line’s resistance and reactance respectively. The line impedance and admittance is calculated from circuit theory as:

\[ Z_{ij} \angle \theta_z = R + jX \]  
\[ Y_{ij} \angle \theta_y = \frac{1}{Z_{ij} \angle \theta_z} \]

where \( Z_{ij} \) and \( \theta_z \) are the magnitude and angle of line impedance. Also \( Y_{ij} \) and \( \theta_y \) stand for line admittance amplitude and angle respectively. The current in the line can be calculated as:

\[ I_{ij} = \frac{v_i - v_j}{R + jX} = \frac{v_i \angle \delta_i - v_j \angle \delta_j}{Z_{ij} \angle \theta_z} = Y_{ij} \angle \theta_y \left( V_i \angle \delta_i - V_j \angle \delta_j \right) = Y_{ij}.V_i \angle (\delta_i + \theta_y) - Y_{ij}.V_j \angle (\delta_j + \theta_y) \]

where \( V_i \) and \( V_j \) are the voltage magnitudes of busbars vi and vj. Likewise, \( \delta_i \) and \( \delta_j \) are the voltage angles of busesbars vi and vj respectively. The conjugate of the current is calculated from:

\[ I_{ij}^* = Y_{ij}.V_i \angle (-\delta_i - \theta_y) - Y_{ij}.V_j \angle (-\delta_j - \theta_y) \]

and the apparent power can be calculated as:

\[ S_{ij} = V_i \angle \delta_i \cdot I_{ij}^* = Y_{ij}.V_i^2 \angle (\delta_i - \delta_i - \theta_y) - Y_{ij}.V_i.V_j \angle (\delta_i - \delta_j - \theta_y) \]
6.3 Load Reduction in Power Networks

\[ Y_{ij}V_i^2 \angle (-\theta_y) - Y_{ij}V_iV_j \angle (\delta_i - \delta_j - \theta_y) \quad (6-5) \]

From this equation, the active and reactive power delivered by the line can be evaluated by (Kundur, 1994, Machowski, 2008):

\[ P_{ij} = \text{Re}(S_{ij}) = Y_{ij}V_i^2 \cos(-\theta_y) - Y_{ij}V_iV_j \cos(\delta_i - \delta_j - \theta_y) \quad (6-6) \]
\[ Q_{ij} = \text{Im}(S_{ij}) = Y_{ij}V_i^2 \sin(-\theta_y) - Y_{ij}V_iV_j \sin(\delta_i - \delta_j - \theta_y) \quad (6-7) \]

\( P_{ij} \) and \( Q_{ij} \) are the line’s active and reactive powers respectively. \( \text{Re(.)} \) and \( \text{Im(.)} \) stand for real and imaginary parts of a complex number.

Now, we suppose that any change in active and reactive power in an arbitrary load in the power system has an effect on active and reactive power of line \( e_{ij} \) (between busbars \( v_i \) and \( v_j \)) (Huang and Yao, 2012). Therefore, we can formulate this change as:

\[ \Delta P_{ij} = \sum_{k=1}^{NBus} \frac{\partial P_{ij}}{\partial P_k} \Delta P_k + \sum_{l=1}^{NBus} \frac{\partial P_{ij}}{\partial Q_l} \Delta Q_l \quad (6-8) \]

In the above equation, \( \Delta P_{ij} \) is the line’s active power change. Also, \( \Delta P_l \) is active power change of load \( v_l \). In this equation, we denote the quantities \( \frac{\partial P_{ij}}{\partial P_k} \) and \( \frac{\partial P_{ij}}{\partial Q_l} \) as the sensitivity of line \( e_{ij} \) active power to active and reactive power change in load \( v_l \) respectively.

Using partial differential equation, we can evaluate these sensitivity parameters as functions of buses’ voltages and angles.

\[ \frac{\partial P_{ij}}{\partial P_l} = \sum_{k=1}^{NBus} \frac{\partial v_k}{\partial P_l} \frac{\partial P_{ij}}{\partial v_k} + \sum_{k=1}^{NBus} \frac{\partial \delta_k}{\partial P_l} \frac{\partial P_{ij}}{\partial \delta_k} \quad (6-9) \]
\[ \frac{\partial P_{ij}}{\partial Q_l} = \sum_{k=1}^{NBus} \frac{\partial v_k}{\partial Q_l} \frac{\partial P_{ij}}{\partial v_k} + \sum_{k=1}^{NBus} \frac{\partial \delta_k}{\partial Q_l} \frac{\partial P_{ij}}{\partial \delta_k} \quad (6-10) \]

where \( NBus \) is a number of buses in the power system. Since line \( e_{ij} \) is linking the bus \( v_i \) and the bus \( v_j \), the flow \( P_{ij} \) will only be related to \( V_i, V_j, \delta_i \) and \( \delta_j \). Thus,

\[ \frac{\partial P_{ij}}{\partial P_l} = \frac{\partial v_i}{\partial P_l} \frac{\partial P_{ij}}{\partial v_i} + \frac{\partial v_j}{\partial P_l} \frac{\partial P_{ij}}{\partial v_j} + \frac{\partial \delta_i}{\partial P_l} \frac{\partial P_{ij}}{\partial \delta_i} + \frac{\partial \delta_j}{\partial P_l} \frac{\partial P_{ij}}{\partial \delta_j} \quad (6-11) \]
\[ \frac{\partial P_{ij}}{\partial Q_l} = \frac{\partial v_i}{\partial Q_l} \frac{\partial P_{ij}}{\partial v_i} + \frac{\partial v_j}{\partial Q_l} \frac{\partial P_{ij}}{\partial v_j} + \frac{\partial \delta_i}{\partial Q_l} \frac{\partial P_{ij}}{\partial \delta_i} + \frac{\partial \delta_j}{\partial Q_l} \frac{\partial P_{ij}}{\partial \delta_j} \quad (6-12) \]

From load flow analysis we know that:

\[ \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} \frac{\partial P}{\partial \delta} & \frac{\partial P}{\partial V} \\ \frac{\partial Q}{\partial \delta} & \frac{\partial Q}{\partial V} \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta V \end{bmatrix} \quad (6-13) \]
in the above equation, the partial derivative matrix is called the Jacobian matrix. If we multiply the inverse of Jacobian matrix by equation 6-13 from left, we have a new equation as:

\[
\left[ \frac{\Delta \delta}{\Delta V} \right] = \left[ \begin{array}{cc} \frac{\partial \delta}{\partial P} & \frac{\partial \delta}{\partial Q} \\ \frac{\partial V}{\partial P} & \frac{\partial V}{\partial Q} \end{array} \right] \left[ \begin{array}{c} \Delta P \\ \Delta Q \end{array} \right]
\]

(6-14)

which is the key equation for values needed in equation 6-8.

By differentiating equations 6-6 and 6-7, we have:

\[
\frac{\partial P_{ij}}{\partial V_i} = 2Y_{ij} V_i \cdot \cos(-\theta_y) - Y_{ij} V_j \cdot \cos(\delta_i - \delta_j - \theta_y)
\]

(6-15)

\[
\frac{\partial P_{ij}}{\partial V_j} = -Y_{ij} V_i \cdot \cos(\delta_i - \delta_j - \theta_y)
\]

(6-16)

\[
\frac{\partial P_{ij}}{\partial \delta_i} = Y_{ij} V_i V_j \cdot \sin(\delta_i - \delta_j - \theta_y)
\]

(6-17)

\[
\frac{\partial P_{ij}}{\partial \delta_j} = -Y_{ij} V_i V_j \cdot \sin(\delta_i - \delta_j - \theta_y)
\]

(6-18)

\[
\frac{\partial Q_{ij}}{\partial V_i} = 2Y_{ij} V_i \cdot \sin(-\theta_y) - Y_{ij} V_j \cdot \sin(\delta_i - \delta_j - \theta_y)
\]

(6-19)

\[
\frac{\partial Q_{ij}}{\partial V_j} = -Y_{ij} V_i \cdot \sin(\delta_i - \delta_j - \theta_y)
\]

(6-20)

\[
\frac{\partial Q_{ij}}{\partial \delta_i} = -Y_{ij} V_i V_j \cdot \cos(\delta_i - \delta_j - \theta_y)
\]

(6-21)

\[
\frac{\partial Q_{ij}}{\partial \delta_j} = Y_{ij} V_i V_j \cdot \cos(\delta_i - \delta_j - \theta_y)
\]

(6-22)

These values are all we need in equations 6-11 and 6-12 to calculate the sensitivities. Likewise, we can write all these equations for the line’s reactive power sensitivity to an arbitrary load’s active or reactive power change:

\[
\Delta Q_{ij} = \sum_{l=1}^{\text{NBus}} \frac{\partial Q_{ij}}{\partial P_l} \cdot \Delta P_l + \sum_{l=1}^{\text{NBus}} \frac{\partial Q_{ij}}{\partial Q_l} \cdot \Delta Q_l
\]

(6-23)

Again, \(\Delta Q_{ij}\) is the line’s reactive power change. Here, in this equation, we call \(\frac{\partial Q_{ij}}{\partial P_l}\) and \(\frac{\partial Q_{ij}}{\partial Q_l}\) as the sensitivity of line \(e_{ij}\) reactive power to active and reactive power change in load \(l\), respectively. Also, we can evaluate these centrality indices as functions of buses’ voltages and angles.

\[
\frac{\partial Q_{ij}}{\partial P_l} = \sum_{k=1}^{\text{NBus}} \frac{\partial V_{kj}}{\partial P_l} \cdot \frac{\partial Q_{ij}}{\partial V_k} + \sum_{k=1}^{\text{NBus}} \frac{\partial \delta_{kj}}{\partial P_l} \cdot \frac{\partial Q_{ij}}{\partial \delta_k}
\]

(6-24)
The partial differential $\frac{\partial Q_{ij}}{\partial P_{l}}$ explains how the reactive power of line $e_{ij}$ evolves when the active power in load $l$ changes. Now that we have all sensitivity equations ready, for any nominated line, we can easily rank the loads according to their impact on that particular line’s power flow.

### 6.3.3 Choosing the Loads to Be Reduced

For any network with $NB_{us}$ nodes and $NP_{Q}$ loads, suppose that the power flow in the line $e_{ij}$ is close to its limitation, the algorithm aims to select a set of loads whose power reduction can bring the power flow in that individual line back into a normal range. The algorithm could be summed up as follows. First, the load flow analysis is performed to evaluate the lines power flow or loading level. According to this information, a directed graph is defined as a model of power grid where the weights on edges are equal to the flows in the corresponding transmission lines. Since the graph is directed, all the weights are positive. Now, by implementing the graph theory pathfinding fundamentals, the algorithm will traverse inside the tree to find the loads fed from the transmission line $e_{ij}$. Then, from this set, only those loads whose more significant part of power come from the line $e_{ij}$ will remain in the set. This is the final set of loads whose elements should be prioritised from $1$ to $n_{L}$ according to their effect on the line $e_{ij}$ power flow. In this set, $n_{L}$ is the number of items. To do so, the sensitivity analysis method which comprehensively explained in section $A$, will be implemented to compute the line $i-j$ power flow sensitivity to the change in the loads’ power and according to those indices, the loads will be ranked from $1$ to $n_{L}$. The higher the rank is, the more effective the load on line $i-j$ power flow is.

### 6.3.4 Simulation Results

We apply the algorithm on IEEE 57 bus network shown in Fig. 2.11. As discussed in the previous sections, before running the algorithm in this test case, we perform the load flow analysis on the network to get the initial values for the voltage and angle of the buses, lines.
active and reactive delivered power and also lines loading levels to check if any line is working close to its loading limitation. Now, we nominate a line which is about to be overloaded which in this case, we choose the line connecting the busbars 15 and 14. The flow direction is from bus 15 to 14. We suppose that an overall load rise has happened which is about to overload the line $15\rightarrow 14$. At the next stage, we propose a method to discover the loads responsible for its overloading. According to the algorithm explained in previous section, a set of loads to be reduced is chosen and ranked using the sensitivity analysis method. The nominated loads getting at least a portion of their power from line $15\rightarrow 14$ are as follows:

$$[14, 56, 46, 48, 38, 47, 37, 31, 32, 33, 34, 35, 36, 40, 39, 57]$$

After traversing the network tree, only the following set is finalized for load reduction:

$$[14, 38, 47, 32, 33, 35, 57, 31]$$

To have a better assessment of the results, we compare them with that of DIgSILENT simulation which is applied as the ground truth. Power Factory DIgSILENT is an eminent power system investigation software used in generation, transmission, distribution and industrial systems analysis. It also offers a wide variety of power system functionalities ranging from classic properties to highly complicated and modern applications consisting wind power generation, distributed generation, simulation in a real-time manner and monitoring the performance of the system as means of system test and supervision. Being easy to apply and windows environment-friendly application, it is a combination of credible and pliable capabilities for modelling the power systems with cutting edge algorithms and a unique database significance. It is also highly flexible for interfacing which makes it a perfect option in automated solutions for business cases.

In Fig. 6.6, the results from DIgSILENT are presented. At $t=10s$, a $10MW$ jump is applied for the loads mentioned in the previous section. As it can be observed, the $10MW$ increment in load #12, has the highest effect on line $15\rightarrow 14$ current as it jumps from $0.995[pu]$ to $1.018[pu]$ which shows $8.6\%$ increase from its initial condition. Likewise, the change for other loads is sketched in Fig. 6.6 and the final level of line current are shown in the sections. The loads are ordered according to their influence on the line flow. As can be seen, the loads are ranked according to their sensitivity as follows:
Also, their sensitivity flow vector is

\[ [1.081, 1.047, 1.020, 1.018, 1.018, 1.015, 1.015, 1.011] \]

where the value 1.011 corresponds to the less sensitivity pertains to load #31.

The comparison between the proposed algorithm performance and the DIgSILENT is summarized in table 6.1. It is evident that the results of the proposed method based on sensitivity analysis are in a good match with the result from DIgSILENT in ranking the loads according to their effect on the line’s power flow. The only discrepancy is at ranks 7 and 8, which the algorithm implies that despite the results from DIgSILENT, at the 7th priority, it is preferred that the load #31 is chosen to be reduced instead of the load #33. It is worth mentioning that the sensitivity parameters in table 6.1 have no unit since in all equations 6-11, 6-12, 6-23, 6-24, both numerator and denominator are in per unit (pu). In the next section, the ranking correlation is calculated between the results from the algorithm and that of DIgSILENT to quantify the performance of this method.

<table>
<thead>
<tr>
<th>Load #</th>
<th>Current [pu]</th>
<th>DIgSILENT rank</th>
<th>Sensitivity</th>
<th>Sensitivity rank</th>
<th>Load #</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>1.081</td>
<td>1</td>
<td>0.1179</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>47</td>
<td>1.047</td>
<td>2</td>
<td>0.0957</td>
<td>2</td>
<td>47</td>
</tr>
<tr>
<td>38</td>
<td>1.020</td>
<td>3</td>
<td>0.0762</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>57</td>
<td>1.018</td>
<td>4</td>
<td>0.0708</td>
<td>4</td>
<td>57</td>
</tr>
<tr>
<td>35</td>
<td>1.018</td>
<td>4</td>
<td>0.0684</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>32</td>
<td>1.015</td>
<td>6</td>
<td>0.0622</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>33</td>
<td>1.015</td>
<td>6</td>
<td>0.0578</td>
<td>8</td>
<td>33</td>
</tr>
<tr>
<td>31</td>
<td>1.011</td>
<td>8</td>
<td>0.0608</td>
<td>7</td>
<td>31</td>
</tr>
</tbody>
</table>

6.3.5 Kendall’s Tau Correlation

We apply Kendall’s Tau Correlation to find the correlation between rankings \( R \) computed using the algorithm and the ground truth \( \sigma \). This correlation can be calculated as:

\[
\tau(\sigma, R) = \frac{n_c - n_d}{0.5n(n-1)}
\]  

(6-26)
Table 6. 2 Kendall’s Tau correlation between the DlgSILENT and Sensitivity methods’ rankings

<table>
<thead>
<tr>
<th>DlgSILENT rank</th>
<th>Sensitivity rank</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \tau = \frac{27 - 1}{0.5 \times 8 \times (8 - 1)} = 0.9285 \]

with \( n_c, n_d \) and \( n \) being the number of concordant and discordant pairs and the size of ranking array respectively. Two pairs \((\sigma_i, r_i), (\sigma_j, r_j)\) are concordant if \((\sigma_i < \sigma_j \text{ and } r_i < r_j)\) or \((\sigma_i > \sigma_j \text{ and } r_i > r_j)\) at the same time. Otherwise, they are discordant. Now, we are ready to calculate the correlation between rank calculated by DlgSILENT and the rank calculated by sensitivity algorithm. Table 6.2 summarizes the calculation.

Hence, for this case, the Kindall’s Tau correlation between the DlgSILENT and Sensitivity methods’ rankings will be:

which shows a very good match between ground truth and algorithm rankings.

It is worth mentioning that when it comes to larger networks, the proposed sensitivity analysis is much more time-efficient than traditional iterative methods based on numerical solution methods for problem-solving. Since the iterative methods highly depend on initial conditions and solution settings, it is more likely that they get trapped in a local solution and suffer from divergence; where the proposed sensitivity-based method, solves the problem only according to the system differential equations and there will not be any iteration because it applies the system last working information or load flow analysis before the contingency or overloading happens.
Fig. 6.6 Line 15→14 current [pu] for a 10 MW increase in the demand of 8 different loads at time t=10s. The order of the loads (14, 47, 38, 35, 57, 32, 33, 31) shows the extent of their effect on overloading the line 15→14. x-axis shows the time (seconds) and y-axis represents the per unit current in the line 15→14.
6.4 Conclusion

In the first part of this chapter, the vulnerability of power systems to cascading failures was put under scrutiny. IEEE 118 busbar network was considered as a benchmark network. Also, a synthetic spatial power network with a controllable number of generators, loads and network density was generated and investigated. A recently proposed maximum flow-based model was used to weight the links, resulting in a proper centrality measure for the links. The capacity of the links is proportional to their centrality value. A number of performance measures (the size of the connected component, the number of failed edges, the maximum flow and the global efficiency of the network) were studied as functions of the capacity parameter. The critical capacity parameter was found for each case, setting the parameter higher than that results in making the networks resilient against cascade failure.

In the second part of this chapter, we studied a novel method to mitigate the cascade failures in power grids. The power grid is one of the most vital engineering networks in the society, and any failure in even one component (a transmission line or a busbar) can lead to a cascade of failures in all over the network which is known as a blackout. To protect the power grids from such catastrophes, we proposed a sensitivity analysis method for ranking the loads -to be reduced- to deter a particular nearly-overloaded line $i \rightarrow j$ from breakdown. In this method, we first chose a set of loads which get fed from this line. Then by traversing in the tree beginning from node $j$ and including all the loads in that set and also their connecting busbar, we chose the final set of loads by shortlisting the nodes whose the more significant portion of the power came from the line $i \rightarrow j$. According to the proposed method, we can calculate the sensitivity of the flow in line $i \rightarrow j$ to the load increase in any arbitrary load in the network. By applying this method to the shortlisted loads, we can rank the loads according to their effect on the special line current.

We compared our findings with the simulation results from the DIgSILENT software designating that the proposed method can effectively rank the loads to be reduced (to deter the cascade failure in the power grid) even faster, primarily when the algorithm deals with bigger networks and traditional methods suffer from divergence. This method can be effectively implemented with much less calculation, and complexity compared to
conventional methods even on the slower processors and lower available memories. It might be worthwhile to further conduct an extensive study on the effect of parameter variations on the cascade failure evolution.
Chapter 7

CONCLUSION AND OUTLOOK

7.1 Introduction

In this chapter, first, the results from previous chapters are summarised in section 7.2. Then, section 7.3 draws a bigger picture of the cascade failure field in complex networks especially power systems and gives some recommendations for the future works in this subject.

7.2 Research Findings Summary

Cascade failures in complex networks make the communication fail in their components and lead to a major loss in network functionality. Especially in the case of power systems, cascade failure is the main reason for significant power outages which most of the times is triggered by a single component failure and gets propagated across the network.

In this thesis, first, the correlation between cascade failures and the structural centrality measures was discussed by studying the cascade depth structural metric for each node in the network. Three structural centralities among others were pointed out to have the most impact on nodes vitality and be applied to control the complex network functionality during the cascade of failures. We found that the nodes with higher degree centralities show negative correlation with cascade depth and to some extent behave like fuses which despite disconnecting a small portion of the network can preserve a substantial part of it.

Next, the cascade propagation pathway was studied when it got triggered by two distinct failure events, single or concurrent edge removal, and the difference between preventive measures was discussed. It was concluded that to make a complex network
resilient to cascade of failures caused by single edge break down, it is crucial to manage such a failure carefully before another edge gets removed. This can be done by putting more protection on the vicinity of that edge. However, concurrent edge removal is much more complicated and protecting the network under such events needs comprehensive studies.

Also in the first part of the last chapter, the vulnerability of power systems against different attack scenarios to power transmission lines was put under scrutiny, and the critical capacity for transmission lines was computed. The methodology was examined on both synthetic and real power systems. It is shown that by setting the capacity parameter higher than its critical value, the network will be resilient against cascade failures. Finally, in the second part of last chapter, a novel load reduction method for the case that a single transmission line is about to get overloaded was introduced and simulated. The results were compared to ground-truth results. The simulation results were compared with that of DlgSILENT software implying that the proposed method can effectively rank the loads to be reduced (to deter the cascade failure in the power grid) even faster hence this method can be effectively implemented with much less calculation, and complexity compared to conventional methods. The main contributions can be summarized as below:

1. For power system vital node identification, in chapter four
   (a) Introduced a new metric to find the node cascade depth which is the number of removed nodes as a result of that particular node’s failure
   (b) Conducted a correlation analysis between cascading failures and node structural centralities

2. For the cascade pathway analysis in chapter five,
   (a) Introduced an algorithm to calculate the cascade propagation route when the cascade is triggered by a single edge failure using the edge geodesic distance
   (b) Developed an algorithm to calculate the cascade propagation pathway for the case that cascade is triggered by concurrent edge removal

3. For the vulnerability analysis in chapter six,
   (a) Generated a synthetic power network according to the existing real power system structural properties
(b) Computed the transmission lines critical capacity to avoid the cascade of failures.

4. For the selective load reduction method proposed in chapter six,
   (a) Introduced a novel traverse method for finding the best candidate set of loads in the power system to rank for each transmission line
   (b) Formulated a new line power flow sensitivity method for rating the loads based on inverse Jacobian matrix

7.3 Future Research Directions

In further research, there are many aspects to the cascade failure in power systems which could be examined subsequently. Some of the possible directions for future research are listed as follows.

7.3.1 Considering Different Node Types

In chapter four, we analysed a wide range of synthetic and real networks to discover the vital nodes in complex networks. In all those simulations, the nodes were assumed to be similar for simplicity. In future work, different types of nodes depending on the network functionality can be considered. Considering the variety of component types in simulations rather than having them all identical can lead to first, a better understanding of complex networks dynamics, second, achieving better and more realistic strategies for catastrophe management, and third, building future networks more resilient to various failures. For instance, in power systems, there are three main busbar types: generators, intermediate busbars, and loads. Also, the capacity of nodes can be considered while delivering the flow across the network.

7.3.2 Studying Dynamical Behaviour Using Cloud Computing

One of the biggest limitations in studying dynamical power systems behaviour is computational boundaries. Most of simulations in this thesis, even when the considered
models are not complicated, were run on desktop computers and in some cases especially for bigger networks it took weeks to come up with the results.

Undoubtedly, when more dynamics are to be considered like power systems short term response to generators augmentation, big loads curtailment, or the power system oscillatory behaviour investigation before reaching a steady-state; needs a huge computational resource which is only available using cloud computing facilities. For the latter case, if the high-frequency oscillation gets simply ignored, they can trigger protection relays and make some unforeseen trips to power system components. For instance, when in chapter five, the cascade propagation route is scrutinized, we assume that the power system reaches a new steady-state by simply ignoring the oscillations right before that. Another example could be the time when large power systems with thousands of busbars are modelled and also different types of generation including traditional power plants as well as renewables like wind farms and aggregated solar farms each with their own technical specifications, exist in the power system, any simplification makes the simulations less accurate which shows the need for such computational facilities.

7.3.3 Optimisation of Reduced Loads

In chapter six, the best candidate loads were determined to be ranked for reduction so that the load in a particular power line does not exceed its boundaries. Depending on the sensitivity values, an objective function can be defined, and the reduction parameters can be optimised to minimise the overall curtailed load across the network. In this way, the optimised amount that each load should shed is computed with subject to returning the power line’s flow to its limitations. This would be helpful to better operate power systems. First, minimising the overall amount of load - to be shed - will increase the power system reliability as a smaller quantity of loads needs to be disconnected. Second, this method will combine the busbars, which need to react and probably lowers the pressure on few loads by more contributions from various load centres. Third, by conducting a distributed overload control, the chance of malfunctioning in all metering and protection controls is decreased, and hence the probability of triggering a cascade of failure will be reduced.
References


BIALEK, J., CIAPESSONI, E., CIRIO, D., COTILLA-SANCHEZ, E., DENT, C., DOBSON, I., HENNEAUX, P., HINES, P., JARDIM, J., MILLER, S., PANTELI,


HIRSCH, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the USA* 102, 16569-72.


MACHOWSKI 2008. Power system dynamics : stability and control


References


