Faculty of Electrical Engineering, Mathematics and
Computer Science
Network Architectures and Services

Performance of Networks and Systems
TNO, Delft

Vulnerability of Power Grids to Cascading Failures

Trivik Verma
Computer Engineering, Distributed Systems and Networks
(4119576)

Committee members:

Supervisor: Prof.dr.ir. R.E. (Robert) Kooij
Mentor: W. (Wendy) Ellens MSc.
Member: Dr. M.E. (Martijn) Warnier

20th June, 2012
M.Sc. Thesis No: RK 2012-1

Copyright © 2012 by Trivik Verma
All rights reserved. No part of the material protected by this copyright may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without the permission from the author and Delft University of Technology.

Abstract

Power sector vulnerability has been a key issue in society for over a decade. A component failure may trigger cascades of failures across the grid and lead to a large blackout. Complex Network approaches have shown a direction to study some of the problems faced by power grids and it is a continuing challenge thus far. Power grids have been studied for their structural vulnerabilities using purely topological approaches. A purely topological approach assumes that flow of power is dictated by shortest paths. However, this fails to capture the real flow characteristics of power grids. We have proposed a flow redistribution mechanism that closely mimics the flow in power grids using the Power Transfer Distribution Factor (PTDF). With this mechanism we enhance the already existing cascading failure models to study the vulnerability of power grids.

We apply the model to the European high-voltage grid to carry out a comparative study for a number of centrality measures. 'Centrality' gives an indication of the criticality of network components. Our model offers a way to find those centrality measures that give the best indication of node vulnerability in the context of power grids, by considering not only the network topology but also the power flowing through the network. In addition, we use the model to determine the spare capacity that is needed to make the grid robust to targeted attacks.

Keywords: power grids, complex networks, cascading failures, flow redistribution, centrality, tolerance parameter, vulnerability

Acknowledgements

In my master thesis I have studied the vulnerability of power grids to cascading failures. This would not have been possible without the efforts of my Professor, Rob Kooij, and supervisor, Wendy Ellens. I would like to thank you Wendy for finding that hour every week in your busy schedule. Friday was my most productive day. I am really grateful for your time, comments and energy towards my project. Rob, thank you for your enthusiasm, crystal clear vision and especially for this platform at TNO to exhibit my potential. Your lectures inspired me to play with complex networks and it will remain so forever. Your attitude towards research has instigated a very collaborative mindset in me.

This thesis was carried out at TNO. My peers were an important part of my daily routine and I would certainly thank Bas, Robin, Jacob and Dick for making it a wonderful experience. Bas, walks to the coffee machine were a lot of fun. Of course, I lost you half way to the machine. Robin, your funny appearances, even in the most tense moments of my thesis brought a smile to my face. Jacob, thank you for all the brain storming sessions. Dick, your jokes were always lost in translation between Dutch and English. I appreciate the comfort level and friendly environment that you established from day one. Your attitude inspired me even more to focus on my work.

I also thank Miroslav and Yakup for their comments. Miroslav, without your reviews and guidance I would never be writing this acknowledgement. Yakup, collaborating with you was a pleasure and I hope to do more in the future. Your calm demeanour and attitude towards research inspires me to follow a similar path in academia. I thank Piotr, Erik and Timo for all the discussions and help with topics beyond my scope.

Martijn, thank you for being a part of my thesis committee and also for providing a platform for a wonderful workshop at the TBM faculty. Dr. Ray Zimmerman, thank you for all your help with MATPOWER and letting me contribute to your existing efforts.

I would also like to thank my parents, Aruni and Rahul, for being such a great support and giving me this wonderful opportunity to study abroad.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Power grids are the backbone of modern society. A power grid is a network of generation, transmission and distribution sub-stations that transfers power to households and local businesses. Energy is transferred through high-voltage transmission links in order to reduce losses over large distances. The evolution of high-voltage grids has been influenced by social, technical, economic, political and environmental decisions that favour economic prosperity, security and quality of life. A disruption in the functioning of power grids can have severe impact on the social welfare of society [8]. The reliability of power grids is thus very critical for optimal functioning of society.

In recent decades a lot of blackouts have been experienced by power grids across the world because of cascading failures. Power outages are consequences of perturbations that overload the entire system by spreading flows across the network. These perturbations range from severe weather conditions to several human errors. A small failure of a component in the power grid may cause some power to be redirected to its neighbours, which in turn may get overloaded and redirect power to their neighbours and cause them to fail. These failures may initially start with any component of the power grid (faults at power station, transmission links, defects in the distribution network or even a small short-circuit) that propagates its effects to its direct neighbours and so on. This effect is called a *cascading failure*.

Economy is largely dependent on energy. Every year the world economy loses billions of dollars due to power outages and equipment failure in power grids [28].

Disruptions in most infrastructure networks do not lead to a disruption in power grids. On the other hand, disruptions within power grids may spread to other infrastructure networks [38]. The focus must therefore be on safeguarding the reliability of power grids.

Like most complex networks, power grids are also made up of small components that act in a simple manner and together give rise to a behaviour that emerges complex in nature [26]. As the size of power grids is increasing over the years and their connection diversity adds to the complexity, it is becoming very important to understand the emergent behaviour of such systems. The question still remains how complex and vulnerable these systems are and how the small simple components give rise to a larger complex mesh of power. Analysing vulnerability of power grids is a significant issue in society and much has been contributed towards such efforts (we will review past literature in great detail in Chapter 2). This gives rise to an interesting direction of research that not only deals with topological (relating to the structure and interconnection pattern of the grid) vulnerability but also looks into the load flows of power grids [4, 16, 23, 29].

1.2 Research Problem

Disruptions (accidents, human errors, component failures, etc.) in a power grid may lead to a failure that propagates through the grid and results in a large-scale power outage. This *cascading failure* phenomena has been researched extensively [11, 16, 23, 31, 35, 38] but many properties of power grids are still missing from the analyses and need to be included for a more accurate study. This leads us to study cascading failures by incorporating properties of flows in power grids in order to deal with the problem of power outages in high-voltage power grids. Through this study we want to identify *vulnerable* nodes of high-voltage power grids. We use the following definition for vulnerability,

Definition 1. *(Vulnerability) The most vulnerable node of a power grid is one which, upon removal from the grid together with its incoming and outgoing links, causes the maximum amount of damage (def. 2) to the grid.*

A removal strategy could either be a random failure or an attack targeted to a specific node and will be specified explicitly wherever necessary.

In order to quantify the damage in a power grid we use the following two quantities relevant for the context of this work. Both definitions will be used and the distinction will be explicitly mentioned wherever necessary,

Definition 2. *(Damage)*

(i) $\underline{D_1}$
The fraction of nodes that have exceeded their capacity and are rendered out-of-service.

(ii) $\underline{D_2}$
The fraction of energy demand that cannot be satisfied after nodes have exceeded their capacity and are rendered out-of-service.

The above discussion about power outages naturally leads us to the following two questions,

Primary Research Question

How to identify the most vulnerable nodes in a high-voltage power grid?

Specifically, we try to find out which of the centrality measures (measures to portray the importance of a node in a network) proposed in the theory of Complex Network Analysis (CNA) is best at quantifying the vulnerability of nodes in a power grid.

Secondary Research Question

How much capacity should exist for each node of a power grid in order for it to sustain attacks/failures?

We use the answer of the first question to find the most disruptive *removal strategy*, i.e. removing the most vulnerable node according to the centrality measure that best captures the vulnerability of nodes in a power grid.

More concretely, through this question we aim at finding the minimum tolerance level α (defined in Chapter 3) such that a power grid sustains a random failure or an attack targeted to its most vulnerable node.

1.3 Research Direction

In order to address the above questions we need to realistically model the flow of power grids and therefore be able to predict the evolution of cascading failures in a high-voltage power grid with more accuracy than existing theoretical CNA models. The existing cascading failure models [16, 23] assume that power flows between two nodes via the shortest path that exists between them. It is of very little significance whether the length of the shortest path is based on distance or impedance of each link in a path. Power, in reality, flows through multiple paths that exist between two nodes and is divided between these paths based on electrical properties (impedance, conductance, capacitance) of the transmission links as depicted in Fig. 1.1. Hence, we try to converge towards a realistic flow mechanism. We will use measures from the theory of CNA to carry out a more comprehensive study.

Every power grid has evolved under different conditions and policies. Yet there are underlying universal characteristics that allow these networks to show similar topological patterns [33] and CNA is a very appropriate tool to study power grids as a complex infrastructure network.

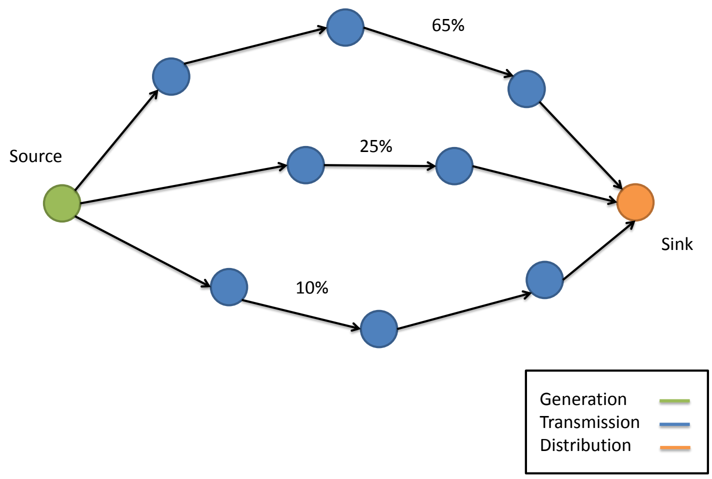


Figure 1.1: A network diagram showing the distribution of power across multiple paths between a pair of generation and distribution nodes of a power grid.

1.4 Outline

Chapter 2 talks about the context in which this work is carried out. Power grids are studied from complex networks approach and the context defines power grids well enough to fit discussions for analysis. Section 2.2 describes the relevant literature in the field of vulnerability analysis of power grids. In Section 2.3, we give a detailed overview of important concepts from the theory of CNA which will be used in further chapters.

Chapter 3 describes our *Cascading Failure* model and its constituent parts in great detail. We describe the flow characteristics of power grids in section 3.1. To study cascading failures differently from the already existing efforts [16, 23, 31, 35, 38], we design a new flow redistribution mechanism and implement performance measures to compare different removal strategies (both theoretical[1] and context[2] based). Section 3.2 explains the basis for our proposed flow redistribution mechanism. In section 3.3 we build the model around the new flow redistribution mechanism using the CLM model [16].

In Chapter 4, we defined certain settings for our simulations as described in section 4.1. Further in section 4.2 we talk about the performance measures that we use to quantify the vulnerability of power grids in case of cascading failures. In section 4.3 we talk about the results of our simulations based on the CLM model [16] and our cascading failure model (section 3.3). A detailed analysis of the two models is shown with illustrative experimental results throughout the text of this chapter.

Chapter 5 marks the end of the report with a short summary and focal points of the project. It also explains future potential that emerges from this work.

[1]Measures from CNA
[2]In the context of power grids

Chapter 2

Preliminaries

2.1 Context

Power grids are vulnerable to cascading failures. For a decade this has been an extensive topic of study owing to so many power outages across the world [39]. *Vulnerability* is quite the opposite of *Robustness*. Robustness is a measure of the resilience of a system in face of attacks and failures. Vulnerability dictates the lack of robustness. Robustness characterises different properties of a system. For instance, in a social network, more connections mean stronger relations between people. On the other hand, in case of an epidemic, more connections will contribute to the spread of the virus. It is a measure that differs from one system to another depending on the kind of property that is being highlighted by the system.

The main goal of vulnerability analysis is to design systems that are more robust. Vulnerability analysis of complex systems is necessary in order to find out the critical components of such networks. *Centrality* measures play a significant role in the identification of such critical components (links, nodes) [7] since they are an indication of the importance of nodes and links within a network.

We will describe the aspects of a complex network in great detail. A complex network is completely defined with the following points [34],

- *System*
  A system is defined as a network together with the protocols and services used and offered by it respectively

- *Properties*
  A system constitutes of certain properties to identify its capabilities. These may include connectivity, throughput, quality-of-service, etc.

- *Perturbations*
  A system can break down partially or completely in the face of perturbations like failures, targeted attacks or virus spreads

From the above discussion some questions about the context, characteristics and perturbations of a system arise [20],

1. What is the *Goal* of the system?

2. What are the perturbations it faces?

3. What strategies will preserve the goals of the system in face of those specific perturbations?

4. What are the costs of these strategies?

Answering these questions helps us in identifying the role of a network under study and also understand how the network fulfils its desired role. Based on the above notes it is important that the study of network robustness/vulnerability takes into account all the aspects addressed by the above questions and the models are closely related to the underlying network properties.

Power grids are considered as small-world networks [17] but some other studies also segregate them as scale-free [3, 21]. Scale-free networks have low connectivity but some nodes are highly connected and are the main hubs of the network [6]. If these hubs malfunction, it could disrupt the functioning of the network considerably. Thus, scale-free networks are robust to random failures but can be highly vulnerable to targeted attacks.

Every network has a different behaviour to disruption. For instance, if we remove a node from an epidemic network, it will help in reducing the spread of the virus. However, in a road network, blocking a main highway would result in severe traffic congestion in and around that area. These observations are very trivial and yet very complex in nature because each network has a distinct function and different dynamics of information flow to achieve that function.

We attempt to briefly answer the above questions about power grids to carry out a more informed vulnerability analysis from a realistic point of view. However, the outcome of these questions is only to present a more logical environment within which power grids can be studied. The questions have been mentioned before and are repeated again for clarity,

1. What is the *Goal* of the system?
   The goal of power grids is to transport electricity from generators to distribution sub-stations via the transmission links. Electricity flows in one direction and takes multiple paths depending on the system properties like impedance, admittance, etc.

2. What are the perturbations it faces?
   A power grid may have component failures, sagging links, overloaded sub-stations, human errors, solar flares and terrorist attacks disrupting the normal functioning of the grid. These perturbations may be modelled as a removal from the network based on the importance of the node/link being removed.

3. What strategies will preserve the goals of the system in face of those specific perturbations?
   Certain strategies like cut-off mechanism, load reduction techniques [31] and distributed renewable sources of energy may save power grids from power outages.

4. What are the costs of these strategies?
   These strategies protect power grids from cascading failures for a comparatively lower cost than the billions of dollars worth of damage caused by power outages to the business sector.

Through this discussion we aimed at understanding power grids better. We attempt to model them using tools from CNA and the methodology is described in the next chapter.

2.2 Related Work

This section is dedicated to the relevant work carried out in the field of power grids and cascading failures. We describe summaries of some relevant papers in detail and focal points of each work that might be useful for expanding our model.

Let us begin with the analysis of Van Eeten et al. (2011) [38]. They carry out empirical investigations in the interdependencies of critical infrastructures. Critical infrastructures are assumed to have a deep impact on each other due to minor faults in one of the networks. Their work involves analysing data from news reports and correlations between cascading failures across different infrastructures. Their findings were rather focused on only two domains: telecommunications and energy. Most failures have propagated from energy networks to telecommunications and there is no sign of a bidirectional relationship. As a result, they undermine the involvement of public sector to monitor all critical infrastructures. Moreover, they suggest a more elaborate involvement in the direction of energy and telecommunication networks instead of focussing on mitigating cascading failures across all critical infrastructures.

Researchers initially focussed on static failures in complex networks [1, 2, 14, 15]. These include the removal of certain components of a network and thereby evaluating the drop in performance of such a network. However, many infrastructure networks like the transport grid/gas pipelines may lose a sizeable part of their performance after a static failure and cause an overload in other parts of the network leading to a partial or complete disruption of facilities. This redistribution of flows led to development of dynamical approaches [16, 27] and their application to more real-world networks.

Next we talk about a dynamic model of cascading failures for complex networks: CLM model, by Crucitti et al. (2004) [16]. They build a simple model to show how small triggers could lead to catastrophic events in infrastructure networks. The model explains redistribution of flows within the network upon break down of a single node. They show that the collapse of a single node with the highest load is sufficient to considerably lower the efficiency of the entire network. For our model we use the basic structure of flow redistribution of loads around the network and reiteration of the CLM model on the new network to get a temporal perspective of the cascade. We briefly explain this model in section 2.3 and describe our cascading failure model in Chapter 3.

The CLM model has been adapted to the North American high-voltage power grid by Kinney et al. (2005) [23]. The North American power grid is one of the most complex technological networks and transmits electricity but also failures to many parts of the grid. They modelled the grid with certain assumptions about how the load is transferred from one transmission substation to another. Their results suggest that a breakdown of transmission sub-stations could lead up to a loss of 25% of efficiency by spreading flows across the network. Furthermore, they found out that disruption of 40% of the transmission sub-stations lead to cascading failures.

Considerable work has been carried out by researchers in the field of cascading failures in the direction of discrete network load models. The above mentioned studies have that feature in common. Simonsen et al. (2007) [35] have taken a step further and studied the transient dynamics of power systems caused by an initial shock in the network. They have justified their work by the fact that loads have oscillations while adjusting to the new structure of the network and some parts may breakdown before the network stabilises again. The lack of a dynamic load model greatly overestimates the robustness of the network. Furthermore, they conclude that using stationary network load models might predict a failure sequence for the network far from reality.

Until now we have discussed literature that marks direct application of concepts and measures from CNA to power grids. Power grids cannot be encapsulated by generic properties of a complex network. Bompard et al. (2009) [10] have done major work in exploring the shortcomings of generic application and designing new methodologies to study power grids more efficiently.

In [11] they analyse the structural vulnerabilities of power grids by extending the topological approach to include the physical operative state of the grid in terms of the flows distributing in the links. The authors designed new metrics such as *entropic degree* and *net-ability* and compared them with the traditional purely topological metrics. Entropic degree takes into account not only the number of connections a node has but also the distribution of weights around these connections. It provides a fairly quantitative measurement of the importance of sub-stations in a power grid. Net-ability describes how efficiently power flows over the links of a transmission grid from supply to demand side. They further extended the topological approach to incorporate several other features of a power grid such as line flow limits, flow paths and demand/supply distribution [9]. *Path redundancy* quantifies the available redundancy of paths in transmitting power from generation to distribution sub-stations using the concept of entropy. In case of a cascading failure, *Survivability*, based on path redundancy and net-ability, measures the effectiveness of a grid to match demand with the generation.

On the lines of extending topological approach to analyse power grids, Bompard et al. (2010) [12] formulated an electrical betweenness based approach that captures the specific properties of power grids in more detail. This metric focusses on the impedances of the transmission links. They compared this with the traditional betweenness metric for IEEE bus networks and found that traditional centrality[1] (section 2.3) measures underestimate the vulnerability of power grids.

Centrality is the backbone for identifying critical components in a complex network. Tutzauer (2007) [36] proposed a centrality measure based on entropy that characterises networks with flow along multiple paths. His motivation lies in the idea that centrality measures must be related to the underlying network properties. Very often, researchers use centrality measures to validate models for a particular type of network that may have flow characteristics different from the ones assumed by the measure. Borgatti [13] points out that this implicit assumption about the flow characteristics of networks may give rise to misleading results.

Bompard's seminal work may have inspired optimisation of distributed flows in power grids from a design perspective. Asztalos et al. (2011) [5] used an edge weighting scheme to optimise the flow efficiency and robustness of scale-free networks to cascading failures. They studied models where flow is distributed and initiates from all paths between a pair of source and sink nodes.

Power grids have been around for almost two centuries and the continuing complex interconnection patterns have made them more liable to disruptions. Pahwa et al. [31] have explored mitigation strategies against cascading effects and agreed upon the use of distributed renewable energy sources for decreasing the load on power grids.

2.3 Basics of Complex Networks

High-voltage grids serve power from generators (supply-side) to distribution sub-stations (demand-side) through a vast network of transmission sub-stations. This network can be modelled as a directed graph G with N nodes (generation/transmission/distribution) and L links (transmission) where power flows only in one direction under normal functioning. The high-voltage power grid is represented by an $N \times N$ adjacency matrix $\{a_{ij}\}$. The element a_{ij} is 1 if there is a connection present between nodes i and j and 0 otherwise.

[1] Betweenness in this case

Vulnerability analysis of a power grid requires quantification of the performance of the grid under certain conditions. We can slightly change the adjacency matrix and define *efficiency* [16] as a measure of robustness of power grids. We replace a with e in the definition of the adjacency matrix changing its connotation from adjacency relation to efficiency relation. Now, the element e_{ij} is 0 if there is no connection present between nodes i and j, otherwise it is a value between $(0,1]$ depending on the change in efficiency of the link between i and j. Initially at time $t = 0$, e_{ij} is set to 1 for all the existing links signifying normal flow of power from node i to j and it represents the efficiency of that link.

In the CLM [16] model, the efficiency of a path between two nodes i and j is defined as the harmonic composition of the efficiencies of the links that belong to that path. The harmonic composition of N numbers $x_1, x_2, x_3, \ldots, x_N$ is defined as $\left[\sum_i^N 1/x_i\right]^{-1}$ and is useful in a variety of different fields [22]. It is important for our discussion to note that harmonic composition is more suitable than using the mean[23]. Let us replace the numbers x_i by links in a path. In this case, harmonic composition takes into account the number of links traversed and if a link has efficiency 0, then the harmonic composition is also 0 as there is no link between the nodes. Whereas, for the mean, a 0 efficiency link would not render the path as invalid.

Average efficiency of the network [25] is used to quantify the performance of the network. This measure is adapted to the North American power grid [23], as follows:

$$E = \frac{1}{N_G N_D} \sum_{i \in G_G} \sum_{j \in G_D} \epsilon_{ij} \tag{2.1}$$

where ϵ_{ij} is the efficiency of the most efficient path between generator i and distribution substation j and is calculated as the harmonic composition of the efficiencies e of constituent links of that path. N_G and N_D are the number of generators and distribution sub-stations in the power grid respectively. G_G and G_D are sets of generators and distributions sub-stations respectively.

A quantity called the capacity C_i is associated with each node i based on the initial load(betweenness) [40] it carries,

$$C_i = \alpha . L_i(0) \qquad \forall\, i = 1, 2, \ldots N \tag{2.2}$$

where α acts as a tolerance parameter and depicts the ability of a node to withstand large inflows of power. This value is greater than 1 but not excessively large since it is limited by the cost of infrastructure. $L_i(0)$ is the load flowing through each node when the power grid is functioning under normal conditions (time $t = 0$).

If an attack or failure occurs at a sub-station of the power grid, the power flowing through this sub-station is redistributed to other parts of the network. Sometimes this power is not very large in quantity and the effect is absorbed by the neighbouring nodes without affecting much part of the power grid. But if the overload caused by this breakdown is huge then power gets redistributed in the entire network and cascades through nodes degrading the efficiency of links along its path, eventually making the entire grid degrade in performance. The efficiency of a link represented by $\{e_{ij}\}$ mimics the flow of redistribution over a period of time. The degradation model is represented as follows,

$$e_{ij}(t+1) = \begin{cases} e_{ij}(0)/\frac{L_i(t)}{C_i} & \text{if } L_i(t) > C_i \\ e_{ij}(0) & \text{if } L_i(t) \leq C_i \end{cases} \tag{2.3}$$

where j extends to all the direct neighbours of i. If a node i fails then the efficiency of the transportation of power from(to) i to(from) its neighbours will follow a linear degradation cycle based on the overload $L_i(t)/C_i$.

Flow redistribution is caused by removing a node from the network. Naturally, removal should be significant enough to cause a network to break down. How do we define the significance of a node? *Centrality* is a measure to quantify the importance of a node within its network. For example, the relevance of a leader in a social network, a road in a transport network, or a cell in a brain network. But before we go deeper into the concept of centrality we need to define the notion of *distance* between two nodes in a network.

Geodesic Distance

The distance between two nodes of a graph is the number of links in the shortest path connecting them. If there are no paths connecting the two nodes, the distance between them is infinite. The *distance* d_{ij} is the length of the shortest path between i and j. The diameter of the network is the maximum d_{max} over these distances and the characteristic path length is defined as,

$$\bar{d} = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \tag{2.4}$$

In the context of pure distance, the length of a path in the case of an unweighted graph is equal to the number of links in the path. In the simplest case of a weighted graph, when the link weights denote the length of the link, the length of the path is the sum of the weights of its links.

Centrality

There are 3 generic ways to define centrality and these are focussed only on the topology of a network,

- *Node degree* [7] is the number of links having a connection to a node. If the graph is directed then there are two forms of degree centrality, *indegree* and *outdegree.* Indegree is the number of links that are directed towards a node and outdegree is the number of links that are directed away from a node.
  In general, for a graph with N nodes, the degree centrality for node v is,
  $$C_D(v) = \frac{deg(v)}{N-1} \tag{2.5}$$
  where N is the number of nodes in the network.

- *Betweenness* [7] of a link or node v is defined as the number of shortest paths between any pair of nodes (set V) passing through v. It is defined as follows,
  $$C_B(v) = \sum_{s,t \in V \setminus \{v\}} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{2.6}$$
  where $\sigma_{st}(v)$ is the number of shortest paths between s and t that are passing through v and σ_{st} is the total number of shortest paths between s and t. It is a measure for determining which nodes or links occupy central positions in a network.

- *Closeness* [7] can be regarded as a measure of how long it will take information to spread from any node to other reachable nodes in the network. It is defined as the mean geodesic distance $d_G(v,t)$ between a node v and other reachable nodes t from that node,
  $$C_C(v) = \frac{\sum_{t \in V \setminus \{v\}} d_G(v,t)}{N-1} \tag{2.7}$$

A contextual centrality measure for power grids is defined as follows,

- *Node Significance* [24] is the amount of outgoing power from a node v in a power grid normalised by the total amount of outgoing power flowing through each node.

$$N_S(v) = \frac{P_{out}(v)}{\sum_{t=1}^{N} P_{out}(t)} \tag{2.8}$$

  Since the flow in power grids is distributed, node significance gives a good estimate of how important a node is in transferring power to other parts of the grid.

Chapter 3

Methodology

This chapter describes our *Cascading Failure* model and its constituent parts in great detail. We describe the flow characteristics of power grids in section 3.1. To study cascading failures differently from the already existing efforts [16, 23, 31, 35, 38], we design a new flow redistribution mechanism and implement performance measures to compare different removal strategies (both theoretical[1] and context[2] based). Section 3.2 explains the basis for our proposed flow redistribution mechanism. In section 3.3 we build the model around the new flow redistribution mechanism using the CLM model [16].

3.1 Flow in Power Networks

As previously explained in section 1.3, in a power grid, flow between a pair of source and sink nodes does not follow the shortest path between them. Power follows a distributed path based on many system properties (impedance, conductance, capacitance) that impose certain limits on the capacity of these paths [11]. Every path that constitutes the flow from source A to sink B is made of several links each. These links have different parameters and each path may contribute a different proportion of power to the flow from source A to sink B.

The amount of information flowing between a pair of nodes is reflected by the shortest path between the nodes [2, 16, 33]. Contrary to this assumption, in power engineering, electric current (the information flowing in power grids) follows multiple

[1] Measures from CNA

[2] In the context of power grids

paths between a pair of nodes. Merely using simplified measures from the theory of CNA for defining flow redistribution in a power grid is not enough. The flow in a power grid can be represented in essence by the PTDF.

3.2 PTDF

The Power Transfer Distribution Factor (PTDF) [18] describes the sensitivity of each transmission link to power injection at a particular node and withdrawal at a reference node. This reference node is also known as *Slack*. If the slack node is missing, then a generator node with the highest real power is chosen as slack for the system.

PTDF of a component T for a flow between source A and sink B reflects the percentage of power that flows through that component T. For instance, when component Z (part of a path between A and B) has a PTDF of 0.2 for a transfer from A to B, then a transfer of 100 MW from A to B would result in an increase of 20 MW at component Z. On the other hand, if the same factor is -0.2, then this transfer would result in a decrease of 20 MW at component Z. This implies that for calculating the PTDF of a transmission link, the source and sink nodes must be specified. In the absence of a sink, the reference node acts as the sink that demands the injection at source in the first place. PTDFs are very useful for analysing the change in load at a particular transmission link and identifying the links that might overshoot their capacity.

For a network with N nodes and L transmission links, the PTDF matrix is given by,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ & & \vdots & \\ a_{L1} & a_{L2} & \dots & a_{LN} \end{bmatrix} \tag{3.1}$$

where A is an $L \times N$ matrix and a_{ij} is the change in power of link i when 1 unit of power is injected at node j and withdrawn at the slack node.

The PTDF of a transmission link i for a flow between a pair of nodes g and d can be defined using Eq. 3.1 and the slack node as follows,

$$a_i^{gd} = a_{ig} - a_{id} \tag{3.2}$$

where a_i^{gd} is the change in power of link i when 1 unit of power is injected at node g and withdrawn at node d.

Computation of the PTDF matrix requires a model to calculate load flows and in most cases a Direct Current (DC) model of power flow [18] is used. DC power flow models are widely used by the power engineering sector to compute load flows for high-voltage transmission grids. Their reliability has been long questioned and not completely accepted for all application purposes but it is nonetheless a good approximation of active power flows and usually gives values close to $\pm 5\%$ [32] as long as the conditions forming the basis of the method are met with. For a high voltage transmission grid, the conditions for a DC model are satisfied [32].

3.3 Model

The cascading failure model that we have designed for simulating a high-voltage power grid has three distinct parts, namely input, redistribution and output illustrated in Fig. 3.1. Before explaining all the constituent parts of the model we point out certain assumptions made during the course of this work.

3.3.1 Assumptions

1. Nodes of the power network are considered to shut down irreversibly, i.e. once they are switched off during the simulation of a cascading failure (whether due to a failure being propagated or a safeguarding strategy to cut-off a part of the network) they are assumed to be dead for the remaining period of the simulation. We made this choice because of the fact that this work aims at a short interval of time which is conceived from a fault in the network, resulting in a power outage in a matter of seconds.

2. The high-voltage power grid network is assumed to be loaded homogeneously, i.e. each link is loaded to exactly the same percentage of its maximum capacity.

3. At each time step of the simulation, a few nodes may get overloaded. The transition between simulation steps is assumed to be static as compared to a dynamic transient analysis where power flows mimic the real time taken for power to flow through a link with certain physical properties. Overloaded nodes are kicked out of the simulation all at once for any intermediate state of the simulation.

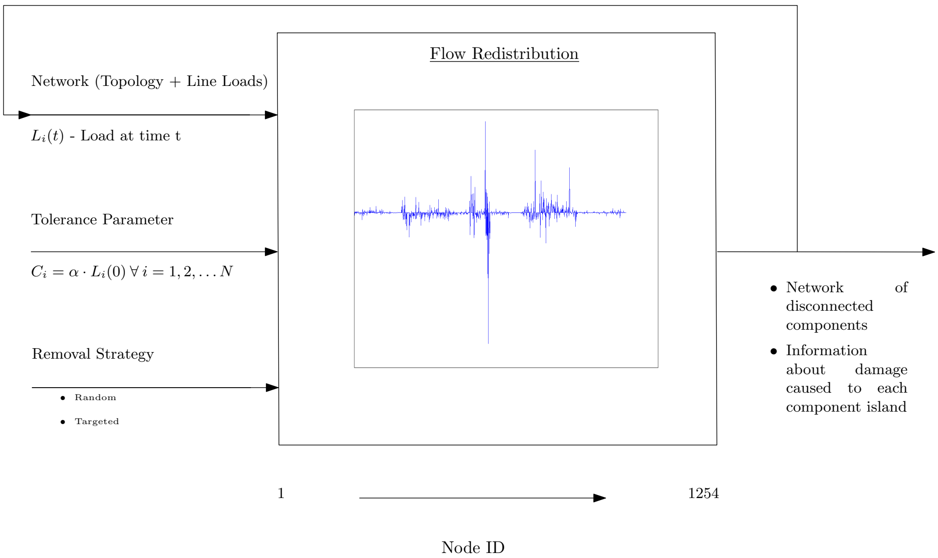


Figure 3.1: Constituent parts of the cascading failure model

3.3.2 Input

The input settings for our model consists of,

Tolerance parameter α

Capacity is a characteristic property of each node/link in any network and it the dictates the maximum load a node/link can withstand. Naturally, we will assume that capacity C_i of a node i is directly proportional to its initial stable load $L_i(0)$ [27] at the start of the simulation (time $t = 0$). It is defined as follows,

$$C_i = \alpha.L_i(0) \qquad \forall\, i = 1, 2, \ldots N \tag{3.3}$$

where $L_i(0)$ is the current power flowing through each node at the moment when a DC power flow solution was calculated. This proportionality coefficient $\alpha \geq 1$ is the tolerance parameter that marks an upper bound on the load that can flow through each node at any time during the simulation. α is technology dependent, i.e. it cannot be an unreasonably high value.

From equation 3.3, it is evident to state the following relation between the tolerance parameter and loading level which is a more inherently natural concept to grasp,

$$\alpha = \frac{100}{ll} \tag{3.4}$$

where,
α is the tolerance parameter and ll is the loading level expressed as a percentage.

The relation in equation 3.4 implies that when the loading level is say 50%, the tolerance parameter will be $\alpha = 2$, i.e. the network can handle twice its initial load at each node.

Network Model

High-voltage transmission grid data is used [37]. This network has 1254 nodes and 1944 links. This network data consists of the topology of the network, the demand and supply parameters at each generation and load station, and node voltages. A DC power flow solution with all these parameters gives an approximate amount of power flowing through all the high-voltage transmission links of the grid. This is a stable snapshot of the system that we use for future modelling and simulation purposes.

The DC power flow solution comes from a MATLAB package called MATPOWER [42]. The calculation of PTDF matrices is also a part of this package. It is used for modelling the power grid as a network to solve power flow equations and carry out other power flow computation problems.

Removal Strategy

The cause for redistribution of flows in a power grid is due to a component failure or shut-down. In order to simulate the network for cascading failures we have to remove a node from the network to trigger the redistribution of flows. This removal can either be *Random* or *Targeted.* By removing the most *important* node from the network we can estimate the worst case damage caused to a power grid due to cascading failures. Importance (centrality - section 2.3) is a vital concept from the theory of CNA.

We included three measures from the theory of CNA that are applicable to all generic network topologies. These are as follows,

- Betweenness Centrality

- Degree Centrality
- Closeness Centrality

And one measure specifically designed for power grids called *Node Significance* [24].

These measures provide us with information about the importance of a node in a network with respect to the network itself.

On the basis of the above discussion we form the following five strategies for node removal which trigger a cascading failure,

1. Betweenness based removal
2. Degree based removal
3. Closeness based removal
4. Node Significance based removal
5. Random removal

1-4 belong to the targeted removals and are based on the importance of a node within a network. A random removal signifies failure or shut-down of a component.

3.3.3 Flow Redistribution

Once a node has been removed, flows start to redistribute in the network. When a node is removed from the network, the incoming and outgoing links have certain loads that redistribute to the neighbours of the removed node.

Fig. 3.2 illustrates the flow redistribution mechanism that we designed for re-routing power in case of a failure or an attack on a power grid. When node N is removed, all the incoming loads at node N are injected to its neighbouring nodes at the other end of the transmission links. Similarly, all the outgoing loads from node N are rejected from its neighbouring nodes at the other end of the transmission links. It can be seen from Fig. 3.2 that the link carrying load l_{iN} will be removed from the network after node N has been removed and the load l_{iN} will be injected into node i. This procedure is applied to all the incoming neighbours i of node N. Similarly, for all the outgoing neighbours j of node N, a rejection of load l_{Nj} occurs at each node j.

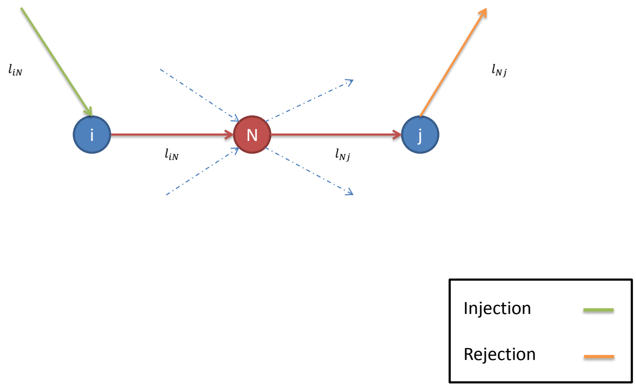


Figure 3.2: Node N is removed from the network and flows redistribute to its neighbours

An injection makes sure that the load that was going into node N will now be redistributed to other parts of the network and a rejection allows for removing the excessive load from a part of the network that doesn't actually receive it any more.

A cumulative shift in load at each node is graphically shown in the flow redistribution part of the model in Fig. 3.1. Due to the numerous injections and rejections, loads at each node either increase or decrease and some might even change the direction of flow of power.

Using (3.1) we can deduce equations formulating injections and rejections for removal of node N as follows,

$$I = \sum_{i \in IN} l_{iN} \cdot A[:, i] \tag{3.5}$$

$$R = \sum_{j \in ON} l_{Nj} \cdot A[:, j] \tag{3.6}$$

where IN and ON are sets of incoming and outgoing neighbours of N and $A[:, i]$ is an $L \times 1$ column vector specifying all rows of the i^{th} column of the PTDF matrix. I and R are $L \times 1$ column vectors of cumulative injections and rejections at every transmission link of the power grid due to removal of node N.

Following equations. 3.5 and 3.6, we can state the following,

$$L_{NEW} = L_{OLD} + I - R \tag{3.7}$$

where L_{NEW} and L_{OLD} are $L \times 1$ column vectors specifying the loads at each transmission link.

After the initial trigger for simulating cascading failures, wherein we remove a node and all the incoming and outgoing links of this node from the network, the flow dynamics of the power grid change. We mark these nodes and links as out-of-service. As the next step, the model calculates the new loads for all the in-service transmission links. The loads flowing through each of the remaining nodes of the network are calculated by simply summing up all the incoming or outgoing loads of each node.

Each node may or may not have exceeded its capacity due to the possibility of extra incoming load. In Eq. 3.8 if the boolean statement x is true then at time t node i has exceeded its capacity and will be made out-of-service together with all its incoming and outgoing links.

$$x : L_i(t) > C_i \tag{3.8}$$

As a result of rendering some nodes and their links out-of-service, some parts of the network may get disconnected and disintegrate into separate islands. Each island of the disintegrated network may or may not have generation substations. If an island is free of any generating source then it is dead, marked red in Fig. 3.3. However, if an island has a generating source, then we iterate the model on all such islands by removing all the nodes (together with their connected links) with the value of x as *true*, ie. overloaded nodes. Each island may result in an additional group of islands and the model is iterated on each group of islands belonging to the same time step simultaneously. This procedure can be encapsulated by a tree structure illustrated in Fig. 3.3.

Some islands may not be affected completely by flow redistribution, i.e. they still have a power generating source. In this case they are still functioning and are marked green in Fig. 3.3. Cascading effects have subsided in such islands and they are still providing energy for a part of their consumers. All the intermediate islands that could break down completely due to cascading failure are marked blue. The tree structure

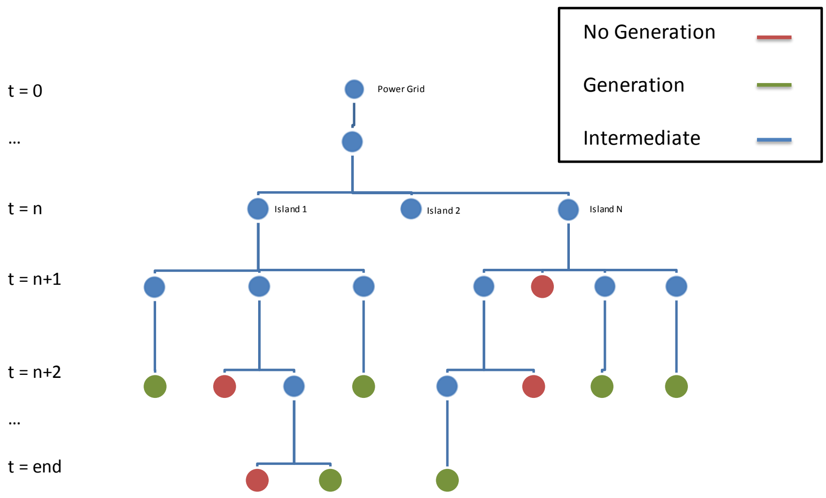


Figure 3.3: A tree structure depicting the logical depth of the model and iterating procedure

shows how an island (original power grid under consideration) can undergo changes in the flow of power and disintegrate into several islands some of which are functioning and some are deeply affected by the cascade.

3.3.4 Output

Simulations (Chapter 4) are carried out to observe the evolution of a cascade. The output of our model consists of several islands that may or may not be able to deliver power to their customers. At each time step a network is produced consisting of one or more disconnected components (islands). Every island contains information about the increase or decrease in load of each node belonging to that island. This group of disconnected islands is fed back as input to the model for the next time step after removing the overloaded nodes and connected links from the network. The flow redistribution model is iterated again on these group of islands. After processing this information from each time step we are able to assess the damage caused to the power grid.

Chapter 4

Experiments

We initiated the research with a question about the vulnerability of power grids to cascading failures. Let us state it again,

How to identify the most vulnerable nodes in a high-voltage electricity grid?

We followed a simulation based approach to solve our research problem. We defined certain settings for our simulations as described in section 4.1. Further in section 4.2 we talk about the performance measures that we use to quantify the vulnerability of power grids in case of cascading failures. In section 4.3 we talk about the results of our simulations based on the CLM model [16] and our cascading failure model (section 3.3). A detailed analysis of the two models is shown with illustrative experimental results throughout the text of this chapter.

4.1 Scenario

The scenario for carrying out our simulations is composed of three parts,

- Network Model

  We deal with three network models. The main focus is on the high-voltage European power grid. This data consists of the topology and load settings of the grid and is available on-line from the database of UCTE [37]. Apart from this we have also considered BA [6] and ER [19] network models of the same size for comparison of real power grid data to synthetic network topologies.

- Node Removal Strategies

  These strategies are based on centrality measures (section 2.3) for nodes in a network. We focus on betweenness centrality, closeness centrality and degree centrality [30] as theoretical measures from CNA to compare with node significance [24], a context-based measure that captures power flows. We also analyse the network using an average of a hundred random removals. The distinction will be made wherever necessary.

- Tolerance Parameter
  We have taken a realistic range of the tolerance parameter, $\alpha \in [1.01, 2.8]$ (Eq. 3.3), in various settings. In reality this is different for each link/node of a network model that represents a power grid but for simplicity we have assumed a homogeneous loading level of the network as previously explained in section 3.3.1.

4.2 Performance Measures

The following performance measures are used to assess the vulnerability of power grids,

- Efficiency
  Efficiency of a network is proportional to the sum of the reciprocals of the shortest path lengths between all pairs of nodes, normalised for that network [16, 23, 25]. This metric is based on distances and does not represent the functioning of power flows. However, it has been used by Crucitti et al. [16] and we retain it for a small part of our simulations for comparison purposes.

- D_1
  The fraction of nodes being overloaded at any time during a cascading failure represents the cumulative damage caused to a power grid at that point in time. This metric as a performance measure is very relevant to our work because a cascading failure propagates as a result of *overloading.*

- D_2
  The fraction of energy demand that cannot be matched at any time during a cascading failure represents the cumulative damage caused to power grid users (consumers) at that point in time. This metric is relevant to our work because cascading failures directly affect the *consumers* of a power grid.

4.3 Simulations

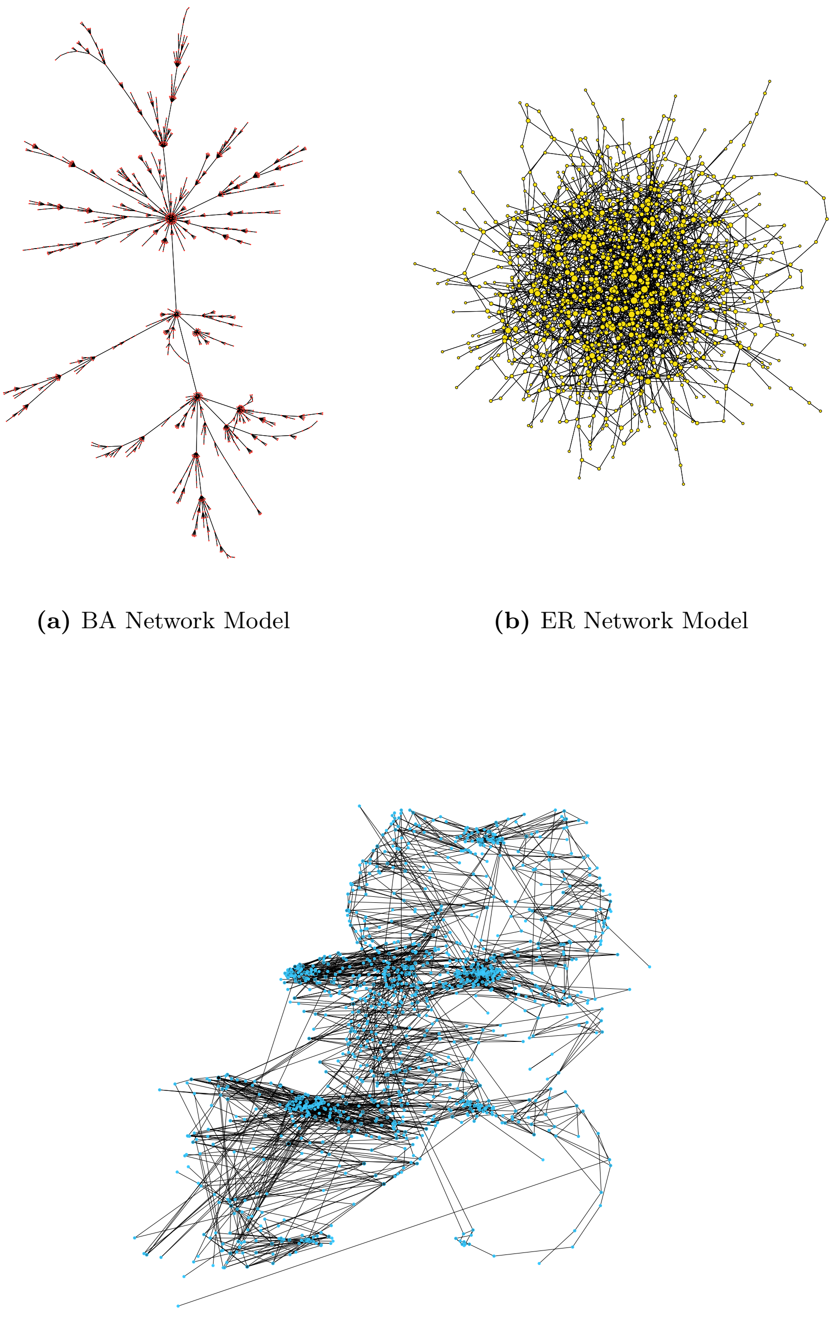
4.3.1 CLM Model

We implemented the CLM model [16] from scratch and investigated cascading failures in two synthetic network topologies[1] and one high-voltage power grid network model which was not a part of the CLM analysis: *a*) BA model [6] in Fig. 4.1a; *b*) ER random graph model [19] in Fig. 4.1b; and *c*) the European high-voltage model [37] in Fig. 4.1c. All the above network models are of the same size (N = 1254, L $\approx$ 1944).

The results were in agreement with those published in [16]. In Fig. 4.2 we show the evolution of network *efficiency* [16, 23] for BA, ER and UCTE network models. Flow redistribution is achieved by a random removal of a single node at time $t = 0$. This simulation has been carried out for three values of the tolerance parameter, namely $\alpha = 1.01$, 1.05, 1.08 and for each value of α the plot is averaged over 10 random removals following the original authors' work [16]. In the first case ($\alpha = 1.01$) the network efficiency drops down by 15% for a BA, 16% for an ER and 16% for a UCTE network model: due to a low tolerance parameter the networks lose a considerable part of their efficiency. In the second case ($\alpha = 1.05$) the network efficiency stabilises at a lower level than before the removal of the node. The third case ($\alpha = 1.08$) shows a constant value of network efficiency, except in a UCTE network model in Fig. 4.2c where the tolerance parameter had to be increased to 2.25 to witness no change in efficiency throughout redistribution of flows across various parts of the network. The above results depict that the UCTE network model is less robust in terms of *efficiency* than the synthetic BA and ER network models of the same size.

The oscillating effect of network efficiency is attributed to the reversibility of overloaded nodes, i.e. nodes that get overloaded can get back to their normal functioning level after some time and redistribution of flows starts the cycle all over again. Suppose there are two paths from generating station g to distribution station d (A and B) and path A is more efficient during normal functioning of the power grid (before any failure). After a failure has occurred, and nodes in path A break down, the flows from path A are redistributed to path B making it more efficient. Due to redistribution of flows, in time, the nodes in path A recover and path A surpasses the efficiency of path B, thus restoring the initial conditions and switching back to the more efficient path

[1] This part is a repetition of [16] and carried out only for comparison purposes

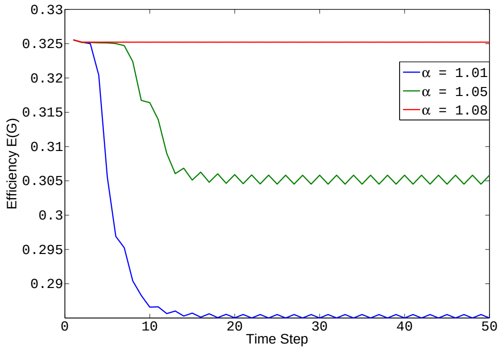


(a) BA Network Model

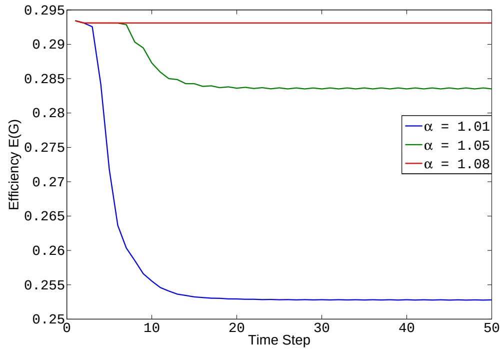
(b) ER Network Model

(c) UCTE Network Model

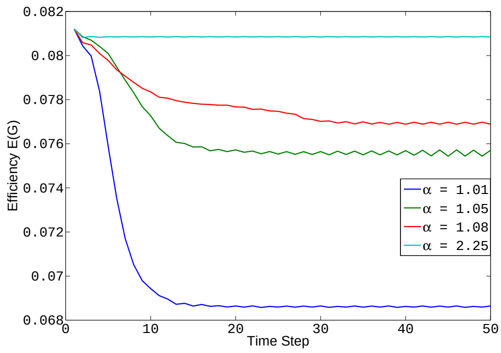
Figure 4.1: Network visualisations of BA, ER and UCTE models of size $N = 1254$ and $L \approx 1944$



(a) BA



(b) ER



(c) UCTE

Figure 4.2: Drop in Efficiency of BA, ER and UCTE network models over time for three different values of the tolerance parameter α. This is caused by the initial removal of a node chosen at random. The plot is an average of 10 such random removals

A. This switching between alternate paths causes the oscillatory effect in efficiency of the network [23]. The oscillations will reoccur continuously if simulation time tends to infinity since at each time step the less efficient path will heal (nodes are reversible in nature) and takeover the previous more efficient path.

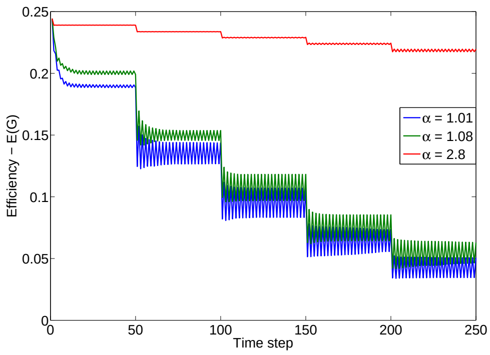
In the next experiment we differed from the CLM approach and changed the simulations to incorporate more than one removal phase and analyse targeted removals. Specifically, whenever the network stabilised (recurring oscillations in network efficiency over time) after a cascading failure subsided, we removed the current most important node[1] of the network and iterated the cascading failure model of CLM on the remaining network. The results are shown in Fig. 4.3. The changed CLM model was simulated on all three network models for five removal stages and three values of the tolerance parameter, namely $\alpha = 1.01$, 1.08, 2.8. The wide gap from 1.01 to 2.8 is because we wanted to test the behaviour of the network for higher values of the tolerance parameter as the network was undergoing multiple stages of redistribution. We changed α considerably from 1.08 and stopped at 2.8 due to the unnaturally high loading capacity that accompanied it.

Apart from the drop in network efficiency for lower values of tolerance parameter we also witnessed an abrupt drop in efficiency at each removal phase, namely at $t =$ 0, 50, 100, 150 and 200. In each removal stage the most important node is removed. After the first node removal, the effects of a cascading failure follow a similar pattern to what we observed in Fig. 4.2. The remaining network is subjected to another node removal thereby causing more damage and lowering the network efficiency even further. It can be seen in Fig. 4.3 that in the worst case (fully loaded network - $\alpha = 1.01$), multiple cascading failures could allow a network to lose 80% of its efficiency in a BA, 65% in an ER and 95% in a UCTE model.

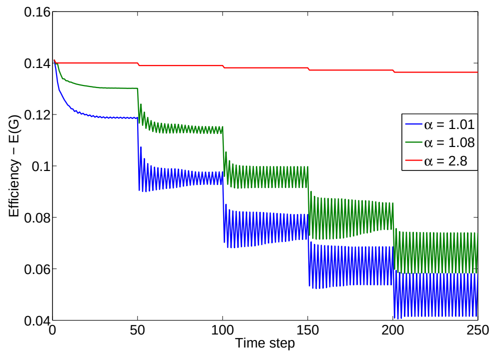
4.3.2 Enhanced Cascading Failure Model

The above results have flows redistributing based on the shortest path between two nodes in a network. We designed a flow redistribution mechanism to accommodate for distributed flows [10], a characteristic property of power grids. We simulated our cascading failure model defined in section 3.3 using proposed flow redistribution mechanism on the UCTE network model. Cascading effects were observed for several removal strategies under different input parameter settings. To encapsulate a wide domain of removal strategies we removed nodes with the highest betweenness centrality, degree centrality, closeness centrality and node significance separately and let the network mimic redistribution of flows that lead to a power outage if at all. For random failures and faults in the network, we removed a random node and averaged the damage caused by a hundred such removals. We used already defined measures of *damage* to quantify the vulnerability of a power grid to cascading failures. We varied the tolerance parameter α and observed the evolution of D_1 and D_2 caused to the network. Our results will

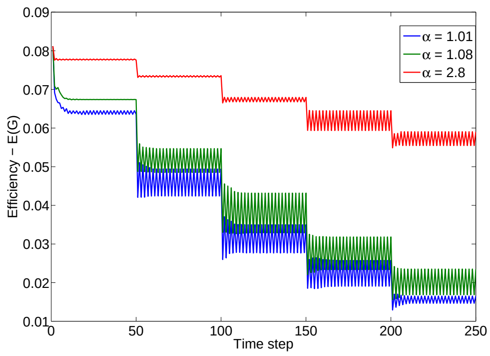
[1] Targeted removal: node with the highest betweenness centrality



(a) BA



(b) ER



(c) UCTE

Figure 4.3: Drop in Efficiency of BA, ER and UCTE network models over time for three different values of the tolerance parameter α. This is caused by the initial removal of a node having the highest betweenness centrality.

not show any oscillations unlike the results from the CLM model because overloaded nodes are not placed back in the system.

4.3.3 Discussion

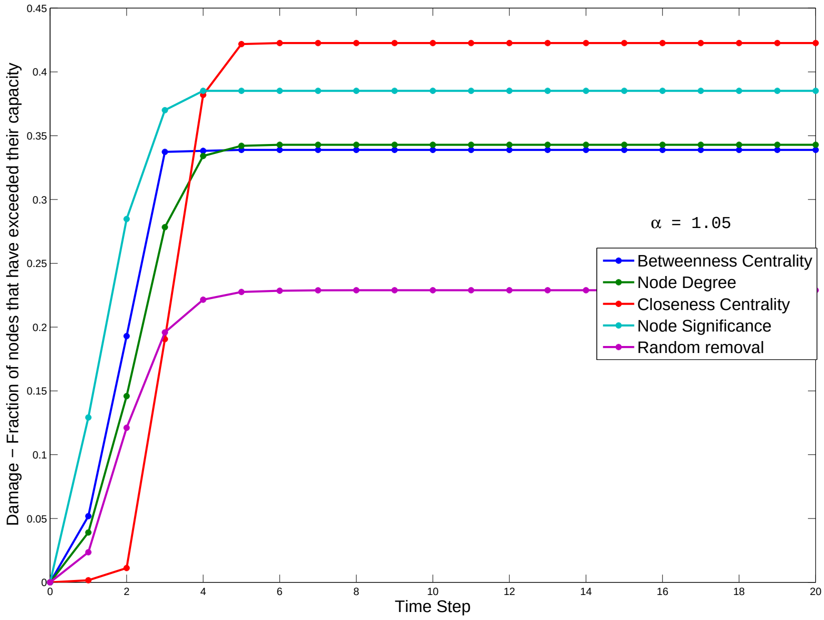
The results are plotted in figs. 4.4 to 4.9. The different figs. correspond to different loading levels of the network. Damage increases with time[1] and after a certain point in time the effects stabilise. We observe in Fig. 4.4 ($\alpha = 1.05$) that within each damage plot the damage caused by any of the removal strategies follows a similar trend and strategies do not require any distinction. This can be attributed to the loading level of the nodes being very close to the maximum or tolerance parameter being close to the minimum (Eq. 3.4); they are bound to break down or cut-off in the event of even the slightest extra incoming power. In Fig. 4.5 ($\alpha = 1.2$), we observe that node significance based removals cause more damage than other removal strategies. Moreover, as illustrated in Fig. 4.6 ($\alpha = 1.5$), the damage caused by a removal strategy based on node significance results in approximately 20 times more damage compared to removal strategies based on the theory of CNA (section 3.3.2). As we come close to $\alpha = 1.8$, in Fig. 4.7, node significance based removals stand out separately from the theoretical measures. In Fig. 4.9 the distinction is still visible but the damage itself is not so considerable any more owing to an even higher value of the tolerance parameter ($\alpha = 2.3$).

Measures from CNA underestimate the vulnerability of nodes in a power grid. Heuristically speaking, *Node Significance* may be an upper bound for the worst case damages caused to a power grid due to cascading failures.

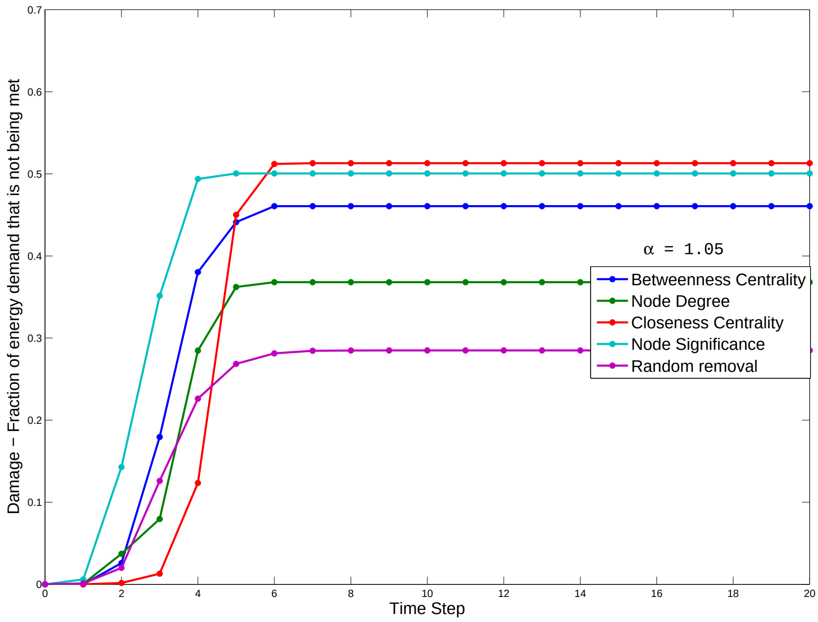
Since both D_1 and D_2 plotted over time for a particular value of α are comparable to each other, we use only one of them for expressing further results, i.e. fraction of unavailable energy. We summarise D_2 caused by removing the node with the highest node significance. In Fig. 4.10, we see that the final damage after a cascading failure has subsided decreases with an increase in the tolerance parameter.

In Fig. 4.11 we plot *damage* as a function of the tolerance parameter. It can be seen that the damage caused by cascading failures decreases by increasing the tolerance parameter, however, it is not a monotonically decreasing function. It behaves erratically for a small change in the tolerance parameter.

[1] 1 iteration of simulation is 1 time step

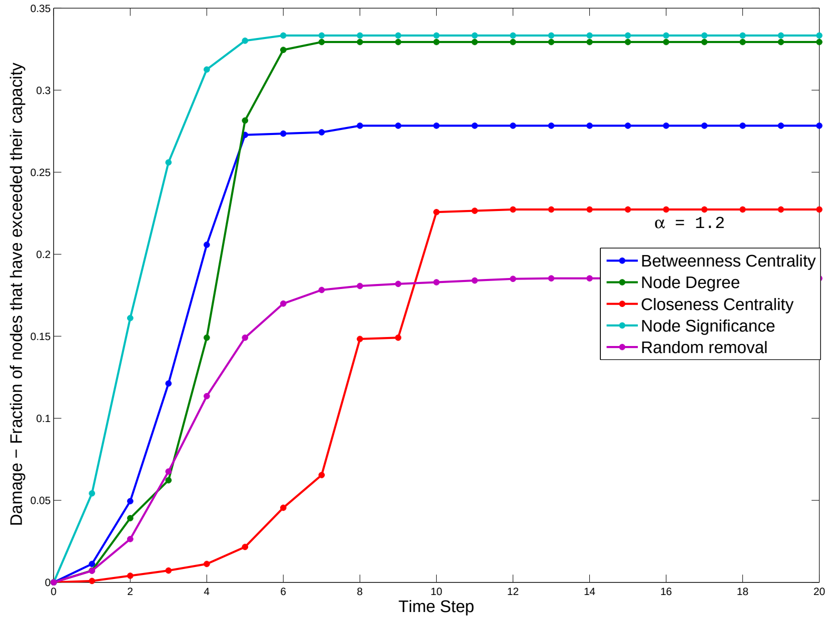


(a) Fraction of overloaded nodes

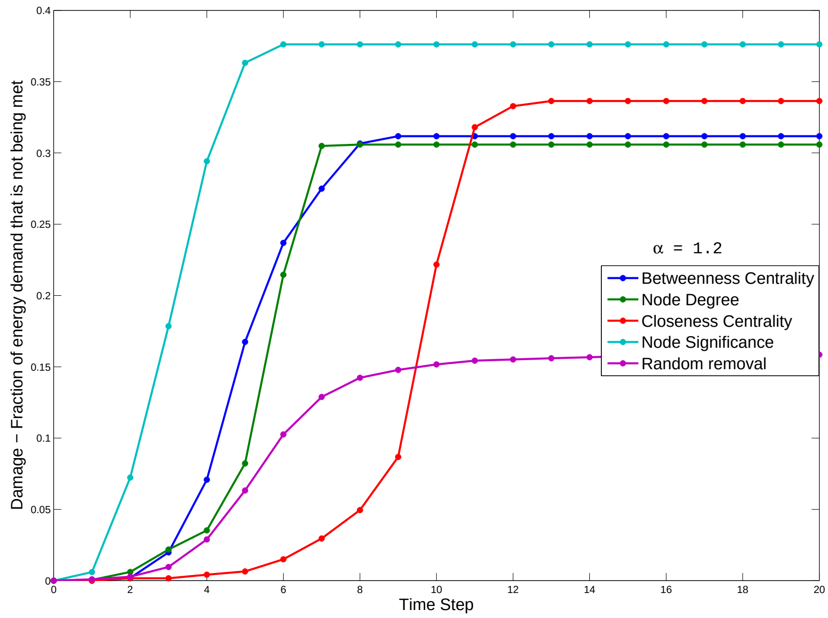


(b) Unmet energy demand

Figure 4.4: Comparison of damage caused over time for different removal strategies due to a cascading failure for $\alpha = 1.05$

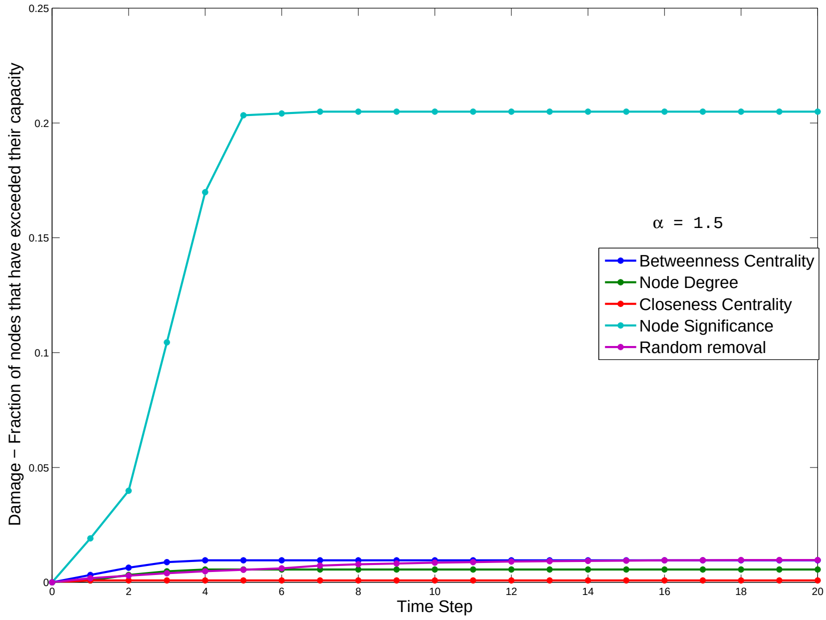


(a) Fraction of overloaded nodes

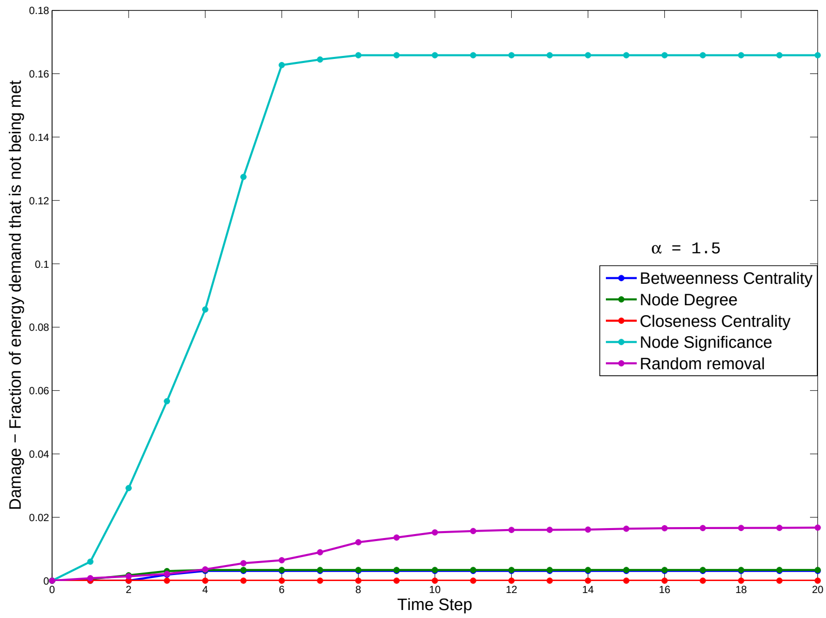


(b) Unmet energy demand

Figure 4.5: Comparison of damage caused over time for different removal strategies due to a cascading failure for $\alpha = 1.20$

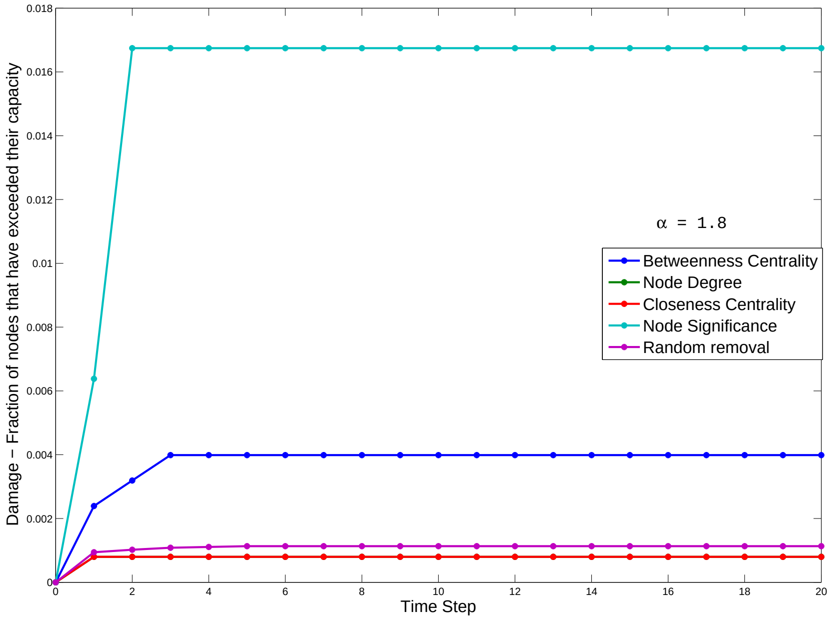


(a) Fraction of overloaded nodes

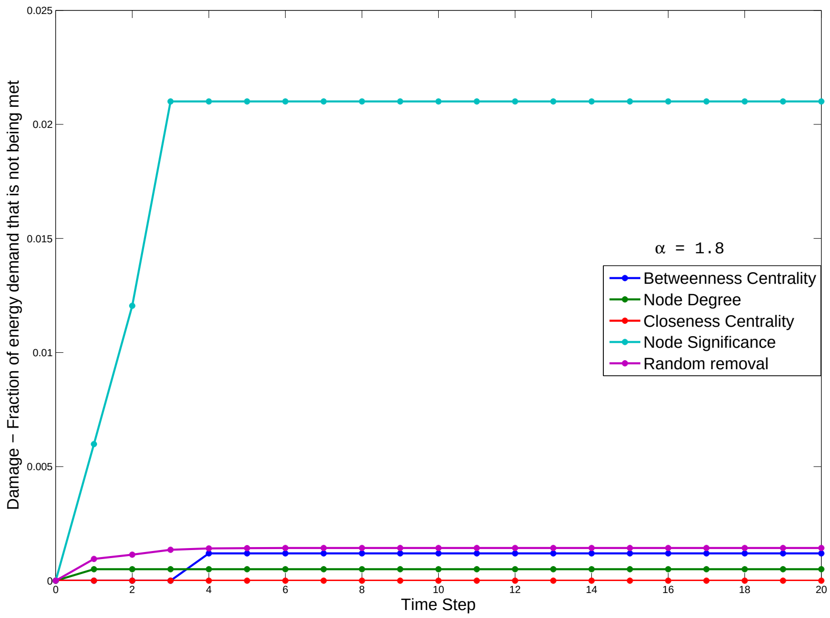


(b) Unmet energy demand

Figure 4.6: Comparison of damage caused over time for different removal strategies due to a cascading failure for $\alpha = 1.50$

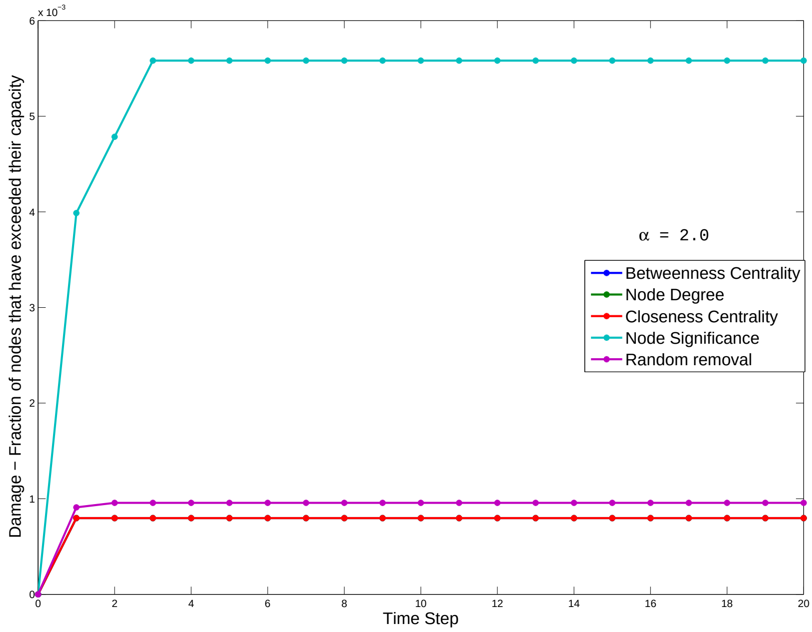


(a) Fraction of overloaded nodes

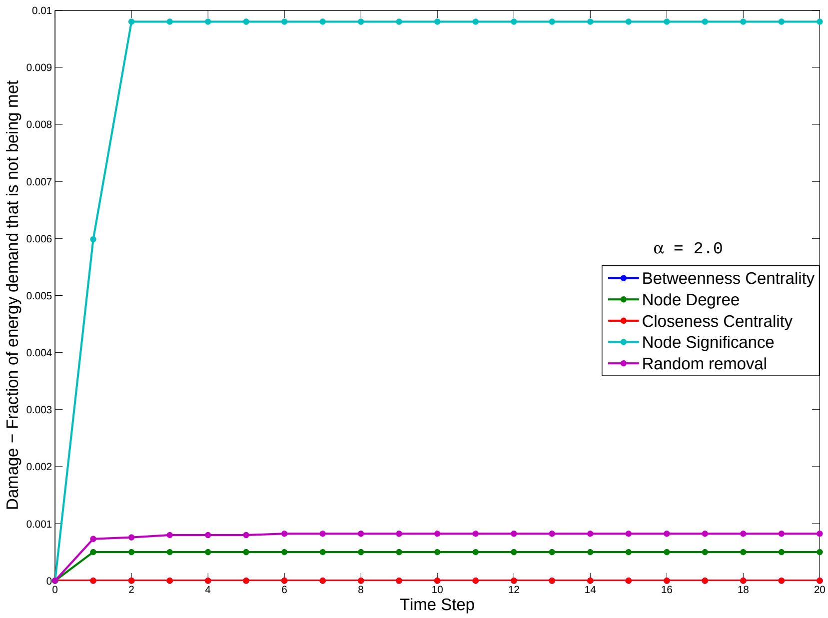


(b) Unmet energy demand

Figure 4.7: Comparison of damage caused over time for different removal strategies due to a cascading failure for $\alpha = 1.80$

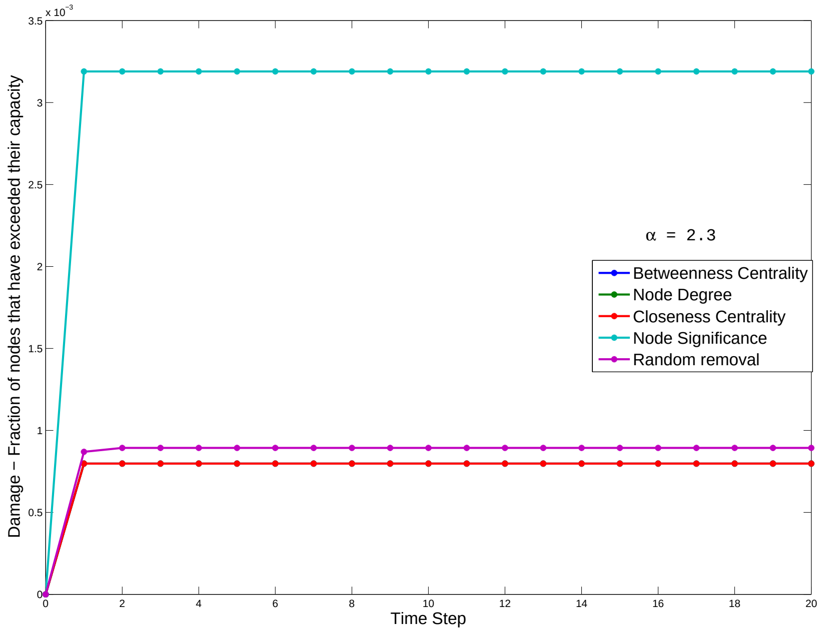


(a) Fraction of overloaded nodes

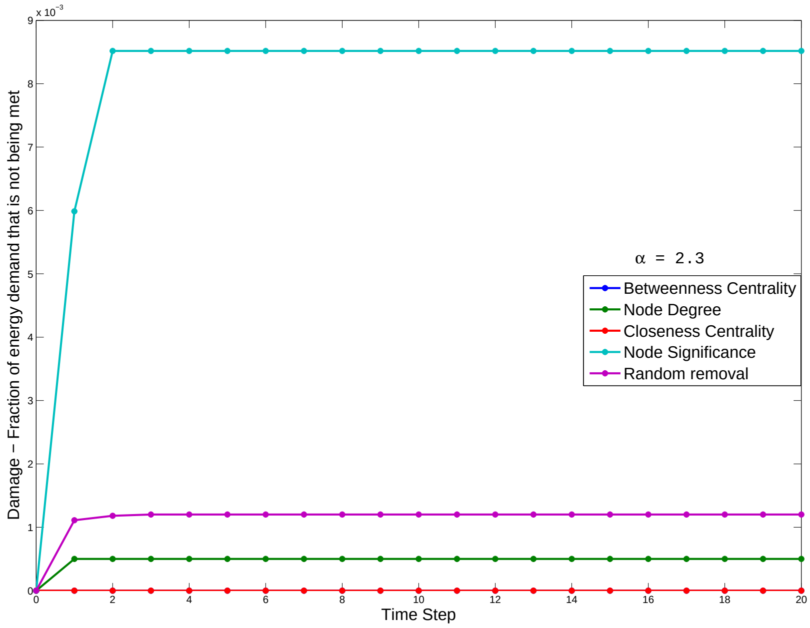


(b) Unmet energy demand

Figure 4.8: Comparison of damage caused over time for different removal strategies due to a cascading failure for $\alpha = 2.00$



(a) Fraction of overloaded nodes



(b) Unmet energy demand

Figure 4.9: Comparison of damage caused over time for different removal strategies due to a cascading failure for $\alpha = 2.30$

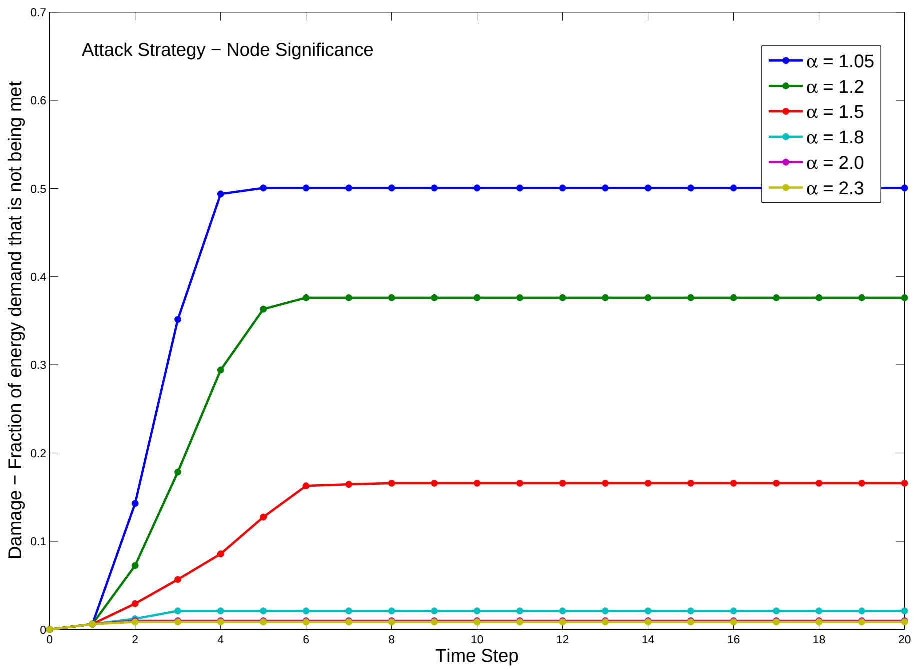


Figure 4.10: Damage caused to consumers over time by removing the most significant node in the network

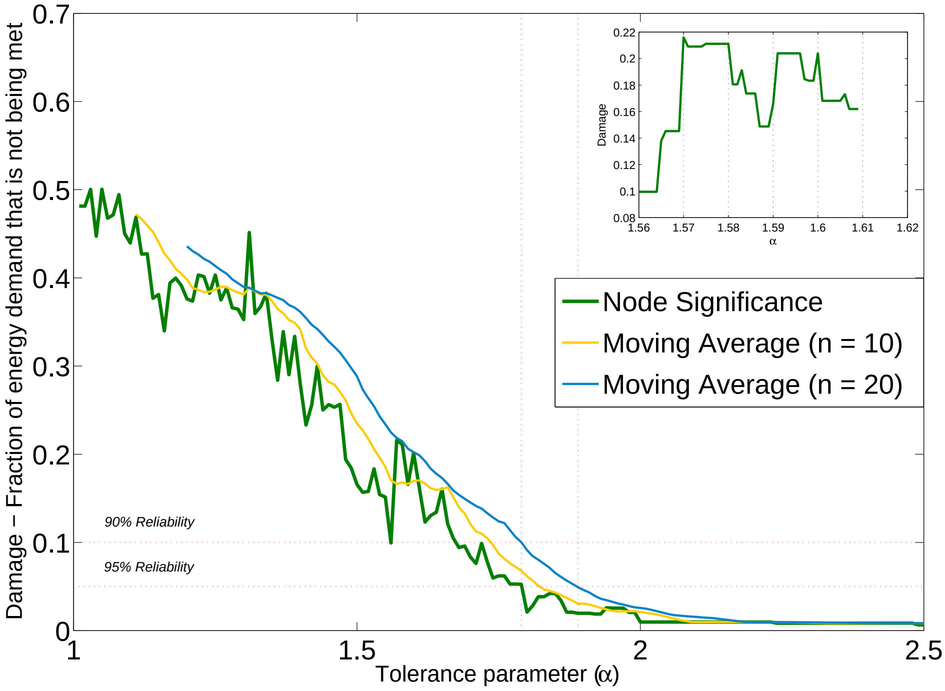


Figure 4.11: Final D_2 caused by removing the most significant node in the network for different values of tolerance parameter

The reason for this erratic behaviour lies in the origin of flow redistribution. Each sub-station of the power grid is assumed to be homogeneously loaded at the beginning. Loads at each substation change after a failure or shut-down causes flows to redistribute. As a result each sub-station has a different threshold for breaking down. To illustrate this we take $\alpha = 1.56$ and remove a node with the highest node significance from the network. There are 21 nodes that get overloaded in the first time step. At $\alpha = 1.57$ there are 20 nodes that get overloaded in the first time step. Since the flows start redistributing after the first time step, this change in the number of overloaded nodes restructures the generation of islands and the redistribution takes different paths in different islands for these two values of α. Since we do not have reversible nodes (section 3.3.1), islands are generated because connectedness of the network cannot be kept intact after overloaded nodes go out-of-service. Therefore, the dynamics of flows change considerably for a very small change in the value of α (step increments of .01) as shown in Fig. 4.11. We change the step increments to .001 and the inset of Fig. 4.11 shows that the damage curve becomes more erratic.

To show the long-term trend we plotted a *simple moving average* of the damage caused at each step of the tolerance parameter. Simple moving average [41] is the mean of previous n discrete data points. The curve represents the general long-term trend of reducing damage to a power grid by increasing the tolerance parameter. In Fig. 4.11 we have taken $n = 10$ data points and $n = 20$ data points. If we have more data points in the window of moving averages, we get a smoother relation between D_2 and α owing to the definition of moving averages itself. Hence, moving average is a good numerical approximation to show the intuitive relation between the destruction caused by a power outage and maximum capacity of power system components.

For a $90 - 95\%$ reliability of the European high-voltage power grid, a value of $\alpha = 1.79 - 1.89$ for the tolerance parameter would suffice. This implies that in the worst case if each sub-station of the UCTE power grid is loaded to at most 55.8% (Eq. 3.4) of its total capacity, the power grid will provide energy for its consumers in the above mentioned reliability zone shown in Fig. 4.11 (intersecting with the blue line, moving average with $n = 20$).

Chapter 5

Conclusion

5.1 Summary

We started by reviewing some traditional cascading failure models in Chapter 2. One particular model, the CLM model, used geodesics for flow redistribution. We reproduced the model to understand the impact of a failure on a power grid. For this we also used the European high-voltage network data to test the model on a real network and contrast the outcome with two synthetic topologies: ER random network and BA scale-free graph. Our results show that power grid studies may not be approximated by studying synthetic topologies of the same size.

Further, we used a simulation based approach to implement a flow redistribution mechanism (Chapter 3) that takes into account the underlying flow characteristics of power grids. We continued the research by analysing the vulnerability of the European high-voltage power grid using our enhanced model. This analysis involved an initial trigger to set redistribution of flows for different values of the tolerance parameter α. The trigger was in the form of a removal based upon centrality measures and α regulated the capacity of each node in the grid. We selected 4 measures, betweenness, closeness, node degree and node significance. We carried out a comparative analysis (Chapter 4) based on these four removal strategies and an average of a hundred random removals. As expected, the removal of a random node caused less damage than a targeted attack. Our studies show that theoretical centrality measures (betweenness, closeness and node degree) underestimate the vulnerability of power grids to cascading failures in comparison with node significance which is a context based measure. Removing a node with the highest node significance causes more damage than removing a node with the highest centrality (betweenness/closeness/degree). This shows that node significance is more suitable than the traditional centrality measures for finding out the most vulnerable nodes in a power grid.

Node significance is valuable for the vulnerability assessment of power grids and also gives insight into the amount of capacity required by the nodes of the grid to provide certain reliability to its consumers.

5.2 Future Work

We examined cascading failures in power grids and wrote a detailed comparison between different centrality measures. A logical next step is to expand the typography and include more centrality measures related to link capacities and optimised flow distribution. A detailed investigation in entropy based centrality measures will also be helpful in understanding the vulnerability of power grids.

Furthermore, transient dynamic analysis of power flow can be explored for a more real-time behaviour. So far we have assumed that transition from one simulation stage to the next is discrete. Using the temporal behaviour of power flow through transmission links will be a step closer to reality.

We also assume that the links are loaded homogeneously. An effort towards using real capacities of power components instead of estimates with tolerance parameter would provide for a more accurate study.

Concerning applications to the industry, with minor changes to input parameters and a suitable replacement for power distribution this model can be applied to low voltage segments to investigate their structural vulnerabilities as an initiative towards *smart-grids*.

References

[1] R. Albert, H. Jeong, and A.L. Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794): 378–382, 2000. 10

[2] Réka Albert, István Albert, and Gary L. Nakarado. Structural vulnerability of the north american power grid. *Phys. Rev. E*, 69:025103, Feb 2004. doi: 10.1103/PhysRevE.69.025103. 10, 17

[3] L. A. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley. Classes of small-world networks. *Proc Natl Acad Sci U S A*, 97(21), October 2000. ISSN 0027-8424. 8

[4] S Arianos, E Bompard, A Carbone, and F Xue. Power grids vulnerability: a complex network approach. Technical Report arXiv:0810.5278, Oct 2008. 2

[5] Andrea Asztalos, Sameet Sreenivasan, Boleslaw K Szymanski, and G Korniss. Distributed flow optimization and cascading effects in weighted complex networks. Technical Report arXiv:1110.3832, Oct 2011. 12

[6] A. L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999. 8, 27, 29

[7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwanga. Complex networks: structure and dynamics. *Physics Reports*, 424:175–308, 2006. 7, 15

[8] E. Bompard, C. Gao, M. Masera, R. Napoli, A. Russo, A. Stefanini, and F. Xue. Approaches to the security analysis of power systems: Defence strategies against malicious threats. 2007. ISSN 1018-5593. 1

[9] E. Bompard, R. Napoli, and F. Xue. Extended topological approach for the assessment of structural vulnerability in transmission networks. *Generation, Transmission Distribution, IET*, 4(6):716 –724, june 2010. 11

[10] Ettore Bompard, Marcelo Masera, Roberto Napoli, and Fei Xue. Assessment of structural vulnerability for power grids by network performance based on complex networks. volume 5508, pages 144–154. Springer Berlin / Heidelberg, 2009. ISBN 978-3-642-03551-7. 11, 32

[11] Ettore Bompard, Roberto Napoli, and Fei Xue. Analysis of structural vulnerabilities in power transmission grids. *International Journal of Critical Infrastructure Protection*, 2(1-2):5 – 12, 2009. 2, 11, 17

[12] Ettore Bompard, Di Wu, and Fei Xue. The concept of betweenness in the analysis of power grid vulnerability. In *COMPENG'10*, pages 52–54, 2010. 11

[13] Stephen P. Borgatti. Centrality and network flow. *Social Networks*, 27(1):55–71, January 2005. 12

[14] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Network robustness and fragility: percolation on random graphs. *Physical Review Letters*, 85:5468–5471, 2000. 10

[15] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Physical review letters*, 85(21), 2000. 10

[16] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. *Phys. Rev. E*, 69:045104, Apr 2004. doi: 10.1103/PhysRevE.69.045104. 2, 4, 5, 10, 13, 17, 27, 28, 29

[17] Ming Ding and Pingping Han. Reliability assessment to large-scale power grid based on small-world topological model. In *Power System Technology, 2006. PowerCon 2006. International Conference on*, pages 1 –5, oct. 2006. 8

[18] Ian Dobson, Scott Greene, Rajesh Rajaraman, Christopher L. DeMarco, Fernando L. Alvarado, Mevludin Glavic, Jianfeng Zhang, and Ray Zimmerman. Electric power transfer capability: Concepts, applications, sensitivity, uncertainty. 2001. 18, 19

[19] P. Erds and A Rnyi. On the evolution of random graphs. In *PUBLICATION OF THE MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES*, pages 17–61, 1960. 27, 29

[20] S. Fox. How can we enhance the robustness via interconnectivity?, 2008. 8

[21] Paul Hines and Seth Blumsack. A centrality measure for electrical networks. In *Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, HICSS '08, pages 185–, 2008. ISBN 0-7695-3075-8. 8

[22] Raj Jain. *The Art of Computer Systems Performance Analysis: techniques for experimental design, measurement, simulation, and modeling.* Wiley, 1991. 13

[23] Ryan Kinney, Paolo Crucitti, Reka Albert, and Vito Latora. Modeling cascading failures in the north american power grid. *The European Physical Journal B*, 46:101, 2005. 2, 4, 5, 10, 13, 17, 28, 29, 31

[24] Yakup Koç, Martijn Warnier, Robert E. Kooij, and Frances M.T. Brazier. A robustness metric for cascading failures by targeted attacks in power networks. Submitted, 2012. 16, 22, 28

[25] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Phys. Rev. Lett.*, 87:198701, Oct 2001. doi: 10.1103/PhysRevLett.87.198701. 13, 28

[26] Melanie Mitchell. *Complexity : A Guided Tour.* Oxford University Press, Oxford, New York, 2009. 2

[27] Adilson E. Motter and Ying-Cheng Lai. Cascade-based attacks on complex networks. *Phys. Rev. E*, 66:065102, Dec 2002. doi: 10.1103/PhysRevE.66.065102. 10, 20

[28] Ring Power Multiplier. Damage caused due to cascading failures. `http://rerateme.com/?page_id=8`. 1

REFERENCES

[29] A. B. M. Nasiruzzaman and H. R. Pota. Transient stability assessment of smart power system using complex networks framework. *CoRR*, abs/1103.3339, 2011. 2

[30] Mark Newman. *Networks: An Introduction.* Oxford University Press, Inc., New York, NY, USA, 2010. ISBN 0199206651, 9780199206650. 28

[31] S. Pahwa, A. Hodges, C. Scoglio, and S. Wood. Topological analysis of the power grid and mitigation strategies against cascading failures. pages 272–276, April 2010. doi: 10.1109/SYSTEMS.2010.5482329. 2, 5, 9, 12, 17

[32] K Purchala, L Meeus, D Van Dommelen, and R Belmans. Usefulness of dc power flow for active power flow analysis. *IEEE Power Engineering Society General Meeting 2005*, 1:2457–2462, 2005. 19

[33] Mart Rosas-Casals, Sergi Valverde, and Ricard V. Sol. Topological vulnerability of the european power grid under errors and attacks. *I. J. Bifurcation and Chaos*, pages 2465–2475, 2007. 4, 17

[34] Caterina Scoglio, Robert Kooij, Phillip Schumm, Ali Sydney, and Mina Youssef. Metrics for robustness in complex networks. 2009. URL http://eece.ksu.edu/sunflower_wiki/images/f/f5/DelftJan09.pdf. 7

[35] Ingve Simonsen, Lubos Buzna, Karsten Peters, Stefan Bornholdt, and Dirk Helbing. Stationary network load models underestimate vulnerability to cascading failures. Technical Report arXiv:0704.1952, Apr 2007. Comments: 4 pages Latex, 4 figures. 2, 5, 11, 17

[36] Frank Tutzauer. Entropy as a measure of centrality in networks characterized by path-transfer flow. *Social Networks*, 29(2):249 – 265, 2007. ISSN 0378-8733. ¡ce:title¿Special Section: Advances in Exponential Random Graph (p*) Models¡/ce:title¿. 12

[37] UCTE. European high-voltage dataset. https://www.entsoe.eu/the-association/history/ucte/. 21, 27, 29

[38] Michel Van Eeten, Albert Nieuwenhuijs, Eric Luiijf, Marieke Klaver, and Edite Cruz. The state and the threat of cascading failure across critical infrastructures: The implications of emperical evidence from media incident reports. *Public Administration*, 89(2):381–400, 2011. ISSN 1467-9299. doi: 10.1111/j.1467-9299.2011.01926.x. 2, 5, 10, 17

[39] Wikipedia. List of power outages. http://en.wikipedia.org/wiki/List_of_power_outages, . 7

[40] Wikipedia. Centrality. http://en.wikipedia.org/wiki/Centrality, . 13

[41] Wikipedia. Moving average. http://en.wikipedia.org/wiki/Moving_average, . 43

[42] R.D. Zimmerman. Matpower simulation package. http://www.pserc.cornell.edu/matpower/. 21